Between-Word Assimilation and Syntax in Child Language Development: a Case Study

Sarah C. Bryan

Thesis submitted for the degree of Doctor of Philosophy

To
The Department of Human Communication Sciences
University of Sheffield

May 2012

Abstract

Previous developmental research has shown that interactions exist between the emergence of connected speech processes and the acquisition of syntax. This study is the first to have investigated these interactions in detail, using a dense data corpus collected for one child, Thomas, over a two-year period, from the age of two to four years. Investigations focused on the emergence of between-word assimilation in constructions containing the auxiliary verbs *can* and *can't*. The methods of research included impressionistic phonetic transcription, quantitative syntactic measures and qualitative syntactic analysis.

The results showed striking parallels between advances in Thomas's syntactic development and the establishment of between-word assimilation as a phonological phenomenon in his speech. It appears that the development of assimilation as a connected speech process (CSP) was directly dependent on Thomas's acquisition of those constructions which provide potential phonetic environments for assimilation to occur. A clear developmental trajectory for the acquisition of assimilation in constructions containing *can* and *can't* was found. This trajectory can be expressed as a phase model, comprising assimilation emergence, establishment and reduction phases.

The impact of maternal input on Thomas's assimilation was also investigated. It was found that cumulative exposure to assimilation over time was important in Thomas's acquisition of assimilation. Thomas's developmental patterns of assimilation and syntax are interpreted within the framework of a usage-based, constructivist approach to language acquisition.

Acknowledgments

I would like to thank the following for their support and assistance throughout the progression of this work:

Professors Bill Wells, Sara Howard and Mick Perkins for their roles as project supervisors; Professor Elena Lieven and Ms. Jeannine Goh for providing me with access to the Dense Database;

Thomas and his mother for their effort and commitment to the compilation of an unprecedentedly rich child language corpus;

My personal assistants, Lucy Sutton, Ruth Hurle, Stephanie Spencer, Sasha Davies, Emily Bennett and Laura Klabis for their help with reading and administrative tasks;

Jayne Woodward, Suzanne Russell, Alison Wheeler and all the staff working at the Disability and Dyslexia Support Service;

Mr. Jose Garcia and Mr. Jonathan Holroyd for technical support;

Dr. Sandra Whiteside and Miss Barbara Bassa for the valuable mentoring, which helped to keep my spirits up;

My family, especially my mother, Mrs. Kerrie Gray and husband, Dr. Matthew Bryan for their love and support through thick and thin;

My guide dog, Quanda, for transporting me safely to and from the university in almost all winds and weathers!

Conventions Used in this Thesis

Italics

These have the following uses:-

1. Used to mark abstract examples of words and constructions, for example, the verb *can* or the construction *can be*;

2. Used to mark grammatical morphemes, for example, the past *ed* morpheme;

3. Used to mark the names of theories, for example, *Universal Grammar*;

4. Used to mark titles of books, tables, columns in a table etc., for example, *Gimson's Pronunciation of English.*

Inverted Commas (double quotation marks)

These are used to enclose utterances actually spoken by a child, either Thomas in the current study, or another child reported in previous research. For example, "Mummy, you go".

Asterisks (*)

These have the following uses:-

1. Used as bullets in bulleted lists;

2. Used in the appendices to mark orthographic transcriptions, which the current author has interpreted differently from the glosses provided in the original data corpus (Lieven, Salomo, and Tomasello, 2009).

Parentheses

These are used to mark uncertainty or omission in the orthographic transcriptions:-

1. If the current author has identified that one of two possible words has occurred, but cannot identify which, then both words are given in brackets with a forward slash between them.

For example (a/the) means that either the word *a* or *the* has occurred, but the current author is uncertain which;

2. If a sound has occurred, but it is uncertain whether this relates to an inflectional morpheme, it is entered in brackets. For example "come(d)", means that the word *come* occurred with a final [d] sound , which might in fact have been *comed);*

3. If a sound or syllable is missing from a word, then the equivalent letters in the orthographic transcription are given in brackets. For example "(th)em" means that the word *them* was pronounced as [ðm], and "(to)matoes" indicates that the initial syllable was absent;

4. Used to note aspects in the transcription which cannot easily be transcribed orthographically or phonetically, for example (babble) or (noise).

Pauses

The following conventions are used to mark pauses in both orthographic and phonetic transcriptions:-

1. (.) indicates a pause lasting approximately one second;

2. (..) indicates a pause lasting approximately two seconds;

3. Pauses of longer or shorter durations are shown by the approximate duration in seconds enclosed in parentheses. For example (0.5) indicates a pause which is approximately 0.5 seconds in length, while (4) indicates a pause which is approximately four seconds in length.

Chat Conventions

The following conventions in the appendices are taken from the Codes for the Human Analysis of Transcription (Chat) system (MacWhinney, 2000) and feature in the original data corpus on which the current study is based:-

1. xxx indicates portions of speech considered to be unintelligible;

2. a@sc marks a schwa vowel, occurring when a target word (such as a pronoun, preposition, or adverb) is not fully pronounced and is therefore not always identifiable;

3. The & symbol occurs when an attempt has been made to approximate a sound orthographically. For example "&na" refers to a sound approximating [næ] and &ne refers to a sound approximating [nə].

Indeterminacy Conventions

The following conventions have been used from the *Extensions to the International Phonetic Alphabet* (extIPA) (Duckworth, Allen, Hardcastle, and Ball, 1990), to mark perceptible features of speech segments considered to be indeterminate and therefore not easily classifiable:-

1. (C) = a consonant of indeterminate place and manner of articulation;

2. (C,Vls) = a voiceless consonant of indeterminate manner of articulation;

3. (C,Vd) = a voiced consonant of indeterminate manner of articulation;

4. (Pl,Vls) equals a voiceless plosive of indeterminate place of articulation;

5. (Nas) = a nasal of indeterminate place of articulation;

6. (Fr,Vls) = a voiceless fricative of indeterminate place of articulation;

7. (App) = an approximant of indeterminate place of articulation;

8. (V) = an indeterminate vowel.

Introduction

Formal theories of phonological development have focused on speech phenomena occurring within individual words (Bernhardt and Stöl-Gammon, 1994; Grunwell, 1987; Stampe, 1979). These theories have provided useful accounts of such phenomena. However, they do not explain the phonological challenges specific to the production of connected speech faced by typically developing children and those with speech difficulties. It has been observed that children with speech impairments may produce single words with high intelligibility, but may be much less intelligible in connected speech (Faircloth and Faircloth, 1970; Howard, Wells, and Local, 2008).

These issues have led to recent research on the typical development of phenomena specific to connected speech, including between-word assimilation, elision and liaison (Bryan, Howard and Perkins, 2010A; Howard, Methley and Perkins, 2008; Newton and Wells, 1999, 2002; Thompson and Howard, 2007). Parallel investigations into the connected speech of children with speech impairments have also been conducted (Howard, 2004, 2007; Howard, Perkins and Raine-Killeen, 2008). These studies have shown that there exist individual differences in patterns of connected speech process (CSP) development and that there exist interactions between these patterns and other aspects of language acquisition, including syntax. The current study aimed to investigate the nature of these interactions in detail in one child, Thomas (AKA Brian in previous research by Bryan et al. and Lieven et al.).

Chapters One to Three of this thesis comprise a review of the literature relevant to the current study. Chapter One discusses the existing polarity between generativist and constructivist theories of language acquisition, although it focuses mainly on the usage-based approach relevant to the current study. Firstly, Chomsky's *Universal Grammar* is outlined as the most influential Nativist approach. A discussion of analytic and holistic learning then follows, leading to a review of the usage-based constructivist approach of Lieven and Tomasello; this

theory accounts for many aspects of language acquisition, including formulaicity in children's speech, which nativist approaches do not explain.

Chapter Two discusses the phonology and acquisition of auxiliary syntax, which sets the scene for the current study of *can* and *can't*. The first section discusses phonological theories of auxiliary syntax, which have mostly been proposed by nativist theorists in the seventies and eighties. The second section discusses patterns of auxiliary acquisition, which have been researched empirically by usage-based theorists in more recent years.

Chapter Three discusses aspects of phonological development, both at the single word level and in connected speech. The distinction between phonological versus phonetic theories is discussed. *Natural Phonology* is outlined as an example of the former and gestural phonology is outlined as an example of the latter. There then follows a discussion of the phonology of multi-word utterances and research on the development of connected speech processes, which sets the scene for the current study.

Chapters Four and Five comprise the methodology section. Chapter Four outlines the strengths and limitations of impressionistic phonetic transcription, including issues of reliability and validity. An entire chapter is devoted to this, as the main methodological tool employed in the current study. Chapter Five describes the aims and methodology in detail.

Chapters Six to Eight comprise the results of the current study. The relevant topics are examined in reverse order to that occurring in the literature review, starting with the phonetic assimilation data, followed by aspects of syntax relevant to assimilation development. This is because the assimilation data form the primary focus of the study, while syntactic aspects form a secondary focus. Also, understanding the results of the syntactic analyses reported in Chapter Seven relies on the reader's prior knowledge of the assimilation data.

Chapter Six contains the results of the phonetic investigations, including control analyses of target alveolar and velar plosives, the focal assimilation data and information on other phonetic phenomena occurring at potential assimilation sites. Chapter Seven contains both quantitative and qualitative data on syntactic development. The quantitative analyses include frequency counts of *can* and *can't*, mean length of utterance and maximum length of utterance. The qualitative data include details of the emergence patterns of *can* and *can't* and their gradual progression to more complex constructions. Parallels between acquisition of syntax and assimilation development are also described. Chapter Eight returns to phonetic data, in a comparison of the assimilation patterns of the child studied and his mother, both globally and in the local context of specific portions of interaction. This chapter also contains a brief comparison of syntactic phenomena produced by the child and his mother.

Chapter Nine is the discussion, in which the findings of the phonetic and syntactic analyses are interpreted in the light of linguistic theory. The necessary connections are made in order to account for the interactions found between the development of assimilation and syntax. The strengths and weaknesses of the current study are evaluated and directions for further research are recommended. This chapter ends with the conclusion, which brings together the major findings and theoretical implications.

Chapter One

Language Acquisition and the Development of Grammar

1.1. Introduction

The means by which human beings acquire language are widely debated. Some linguists believe that innate factors are primarily responsible, while others place greater emphasis on learning and the environment. These two positions are continuous rather than discrete schools of thought, although arguments made at extremes of this continuum are strongly opposing. At one end of the continuum, Noam Chomsky's *Universal Grammar* is an approach which focuses on the acquisition of language via innate, underlying linguistic competencies. This chapter begins with a brief introduction to this theory. At the other extreme, more recent theories, such as the usage-based constructivist approach of Elena Lieven and Michael Tomasello (2008), view language as acquired through a combination of global cognitive processing and linguistic experience. This theory forms the primary focus of this chapter and the current study.

1.2. What is Grammar?

The term grammar refers to the set of relationships which structure language. The acquisition of grammar requires the child to master both the morphology and the syntax of the language which he is learning (Karmiloff and Karmiloff-Smith, 2001). The morphology of a language involves the formation of words from subparts which have meaning, but many of which cannot stand alone as words. Examples of grammatical morphemes are the past *ed* ending which is added to verbs to signify the past tense and the *s* ending which when added to nouns

signifies a plural. Syntax is the way in which words combine to form sentences. The extent to which grammatical features are expressed through morphology or syntax varies greatly across languages. One of the most fascinating features of grammar is its creativity, that is the seemingly infinite morphological and syntactic structures which people can produce to express thoughts and feelings.

The acquisition of grammar begins in the second year of life, when the child has acquired between 100 and 150 words. However, there are considerable individual differences in patterns of grammar acquisition; some children demonstrate grammatical knowledge as early as 14 months of age, while others do not combine words until after two years of age (Karmiloff and Karmiloff-Smith, 2001). Formal theories of language acquisition have emerged over the last 70 years, which offer different explanations of the language acquisition process. The Nativist Approach of Chomsky and the constructivist approach of Lieven and Tomasello are discussed below.

1.3. Chomsky's Universal Grammar

Noam Chomsky's *Universal Grammar* (*UG*) is the best-known linguistic theory. It was originally proposed in relation to adult language (Chomsky, 1957) and was later specifically applied to language acquisition (Chomsky, 1965). *UG* has undergone many revisions since then, the most recent of which is the *Minimalist Program* (Chomsky, 1995). Chomsky argues that the genetic basis of the human language faculty explains several phenomena. Firstly, it explains the fact that language is unique to human beings and cannot be acquired even by genetically similar species such as apes (Chomsky, 1967A). Secondly, it explains how humans can use linguistic elements creatively to produce novel, grammatically correct sentences without prior experience of the relevant utterance (Chomsky, 1964, 1965, 1967A, 1976). Thirdly, it explains the universality of language among human beings, hence the term *Universal Grammar*.

UG aims to describe the nature of humans' genetically in-built linguistic capacity. It states that humans have innate knowledge of the principles of syntactic structure, and that as the child learns to speak, various parameters are adjusted, so that the child understands and produces those grammatical structures relevant to the language which he is learning (Chomsky, 1965, 1966, 1967A, 1985, 1995). During the period when the child is learning to speak and parameters are being set, Chomsky argues that there is a discrepancy between the child's linguistic competence and their linguistic performance, the former being more advanced than the latter. He therefore argues that observation of a child's comprehension and production skills, that is, their performance, is not reliable as a stand-alone measure of language acquisition, because it may lead the observer to under-estimate the child's linguistic abilities, that is, their competence. He argues that this discrepancy is most evident when a child shows sudden increases in progress, for instance a sudden progression from producing no auxiliary verbs to producing all of them (Chomsky, 1964).

According to Chomsky, (1966, 1985), the language faculty consists of several different modules, each of which has a specific linguistic function. These modules operate at different linguistic levels and work in parallel to produce the rules which generate semantic representations, syntactic structures and phonetic representations. The conversion between linguistic levels is known as transformation (Chomsky, 1967A). When a person has mastered the rules for the language they are learning, they can interpret any linguistic input, provided that it comprises basic linguistic units with which they are familiar and conforms to the rules which they have acquired.

1.4. The First Empirical Study of Language Acquisition

Chomsky did not consider it necessary to test UG empirically. The first empirical, longitudinal study of language acquisition was conducted by Roger Brown (1973). Three

children (known as Adam, Eve and Sarah) were observed from the onset of their first multi-word utterances until they were four years of age, at which point most linguistic skills had been acquired. This resulted in a large corpus of linguistic data, on which many different linguistic analyses were performed. Brown advocated the investigation of a child's level of linguistic performance as a means of determining their underlying linguistic knowledge and competence.

The children were matched for two quantitative measures of grammatical development devised by Brown: mean length of utterance (MLU) and upper bound (length of longest utterance) rather than age at the start of the study, to account for the fact that children learn language at different rates. MLU is calculated in morphemes rather than in words. The advantage of this is that MLU typically increases with the acquisition of almost any kind of new grammatical knowledge, including the addition of new clause and phrase elements, morphological development and overall complexity. However, MLU loses validity over time with the growing variety and complexity of the child's utterances, and syntactic structure and length are increasingly determined by interactional context (Brown, 1973).

In a detailed analysis of the children's syntax, Brown found that the acquisition of specific grammatical elements including pronouns and inflectional morphemes followed a similar developmental pattern across the children, although there was individual variability in the exact age at which specific forms were acquired. Brown interpreted his findings as evidence for innately pre-specified linguistic abilities, in support of UG. However, these findings could equally be attributed to universality in the general cognitive mechanisms which underpin language acquisition.

1.5. The Role of Formulaicity in Language Acquisition

A major difficulty with *UG* arises from research which has shown that a young child's linguistic competence does not necessarily exceed their early performance. It appears that

children can often produce relatively complex sentences in advance of their underlying grammatical knowledge. The first documented examples of such utterances came from Clark's (1973) case study of her son, Adam. Clark observed that Adam combined utterances which he had previously used or heard in order to construct new utterances for the purposes of elaboration, self-correction and constructing more complex sentences. One instance of elaboration occurred when he said, "Mummy, you go"; when his mother asked, "Where?", he elaborated with, "Mummy, you go swings". He produced imitations containing grammatical errors in syntactic structures which he had previously produced correctly, for instance, "sit my knee". He produced novel complex sentences by embedding two sentences which he had previously produced separately, for instance "I want you get a biscuit for me". He also substituted syntactic elements to produce novel sentences, for instance when he modified "wait for it to cool" (routinely used at mealtimes) to produce "wait for it to dry". Clark concludes from these findings that her son employed a strategy of using previously formed linguistic plans to construct novel utterances, thus showing linguistic performance without grammatical competence. In response to her son's ability to embed previously used clauses, she argues that familiarity of constituent phrases and clauses influences resultant syntactic complexity more than does sentence length alone. She proposes that grammar is acquired through the gradual analysis of utterances which the child initially learns as wholes, leading to increased syntactic creativity (Clark, 1973).

Clark's conclusions have been substantiated by further research, which has shown that children's early language comprises many formulaic utterances (also referred to in the literature as unanalysed, under-analysed, frozen, gestalt, stereotype, item-learned or prefabricated utterances) and relatively few productive or creative utterances. Peters (1983) defines a formulaic utterance as a multi-morphemic phrase or sentence which has become a single lexical item, either on an individual level or through social interaction. Wray (2002) defines a formulaic utterance in more cognitive terms as a sequence of elements (such as words) which appears to be stored in and retrieved from memory as a whole unit. A child may thus extract a phrase or sentence from the speech which they hear and process it along

with surrounding contextual information as a single linguistic unit or lexical item, even though it comprises several lexical items in the adult language (Peters, 1983; Wray, 2002). Cruttenden (1981) similarly defines unanalysed strings as sequences of words or morphemes, which are learned as single items and which the adult understands to have more complex structure than the child does.

Peters (1983) advises that it is theoretically misleading to consider a child's linguistic units as equivalent to the minimal units used by linguists to describe adult language. Instead, she suggests that formulaic phrases and sentences may initially function as operational linguistic units for the child. On these grounds, she cautions that mean length of utterance (MLU) may not be a valid measure of children's development, because this measure is based on morphemes, which are not necessarily psychologically real units for the child. The analytical process by which children segment larger units into smaller ones is gradual and prone to error. Plunkett (1993) proposes that one of three things may happen when a child first segments an utterance. Correct segmentation occurs when the child's segments match those in the adult language. Overshoot occurs when a child's segment contains several segments in the adult language, such as several words forming a complete phrase or sentence. Under-shoot occurs when the child's segments are smaller than those in the adult language, such as a single syllable which forms part of a disyllabic word (Plunkett, 1993). Initial rudimentary syntactic analysis may involve the child substituting only one element, such as the formulation of "wait for it to dry" from "wait for it to cool" (Clark, 1973). Peters (1983) explains this phenomenon in terms of utterance templates or frames with one or more substitutable slots.

The child's increasing ability to manipulate smaller syntactic units results in syntactic productivity. Productivity is defined as the ability to use the structural system of language analytically, through the segmentation of sentences into their composite words and morphemes, and through the combinatorial construction of sentences from words and morphemes (Wray and Perkins, 2000). Wray and Perkins (2000) and Perkins (1999) argue that once children have achieved segmentation and syntactic productivity, they may again

synthesise and store some larger linguistic units from small ones, in order to retrieve frequent collocations more efficiently and in a more adult-like way.

Peters (1983) suggests several criteria for determining whether a specific utterance used by a child comprises one unanalysed or partially analysed linguistic unit and is formulaic, or whether it comprises several juxtaposed units and is therefore productive. An utterance may be considered formulaic if:-

- The utterance is idiosyncratic to the child and is used repeatedly without alteration to the form;
- The utterance contains grammatical forms which are not yet otherwise present in the child's productive language;
- The utterance is sometimes inappropriate or erroneous in the context in which it is used;
- The utterance is phonologically coherent, in that it has no pauses and has a smooth intonation contour;
- The utterance is used in conjunction with specific contexts and events;
- The utterance is a formula used by others in the child's community.

Cruttenden (1981) distinguishes between two different developmental stages, namely item-learning and system-learning. These stages correspond very closely to Peters' (1983) concepts of initial formulaicity followed by gradual analysis and segmentation. Item-learning occurs when a child learns specific items in conjunction with the surrounding context. This leads the child to produce sequences at any linguistic level (phonological, syntactic or semantic) which are in advance of their developmental stage. Similarly to Peters' second criteria for formulaic utterances, Cruttenden (1981) argues that an utterance can be considered item-learned if it contains elements which are not used elsewhere in the child's productive language. Learning continues on an item by item basis until the child recognises part of an item in a novel context. This leads to generalisation and extraction of the linguistic system. This is known as system-

learning and corresponds to Peters' concept of segmentation and consequent productivity (Cruttenden, 1981).

Research on the role of formulaicity in children's language has also shown that the language acquisition process is not universal and that individual differences exist. Instead, there appear to be differences in the language learning strategies adopted by different children and employed by the same child at different developmental stages (Peters, 1977, 1995). On one end of the continuum, there is the analytic strategy, in which children appear to assemble utterances from their constituent linguistic units, such as words and morphemes. This type of acquisition can be accounted for by traditional nativist theories such as that of Chomsky (1965), owing to the focus of such theories on the formal linguistic units used by adults to describe language. On the opposite end of the continuum is the holistic strategy, by which children appear to learn utterances formulaically and undergo a gradual process of linguistic analysis in order to develop syntactic productivity. The distinction between analytic and holistic learning is continuous, rather than discrete. This reflects the finding of parallel analytic and holistic learning within the same child (Peters, 1995). The holistic strategy accounts for Clark's (1973) data and is discussed in more recent, non-nativist construction-based theories of language acquisition.

1.6. Research on Analytic Versus Holistic Learning Styles in Language Acquisition

Martin Braine (1976, 1963) was one of the first theorists to identify that children's early multiword utterances are not as creative as proposed by Chomsky. In his longitudinal study of three children's first two-word utterances, he observed that a small number of words were repeatedly used by individual children in the same position (either initial or final) across a variety of utterances (Braine, 1963). He referred to these words as pivots, examples of which include *more* and *allgone* occurring prior to a noun. Braine referred to the remaining open

class words constituting the majority of the child's vocabulary as the X-class. In contrast with pivots, X-words can occur in more variable utterance positions. Braine proposed that children's first two-word utterances are constructed using a pivot grammar, whereby the child uses knowledge of both the meaning and utterance positions of pivots to combine them with X-words, as appropriate for the physical and social context. In this way, Braine argues that children's syntax develops as the number of pivots in their vocabulary increases, while vocabulary increases when words are added to the X-class.

Following a more extensive, cross-linguistic analysis of child language corpora, Braine (1976) concluded that his previous proposal of pivot grammar had been too simplistic and did not account sufficiently for the variety of construction types encountered in his current analysis. Although he observed some two-word utterances which could be accounted for by pivot grammar, he also encountered utterances which did not appear to contain pivots, and in which word order varied across different productions of the same utterance. Because these findings were not predicted by pivot grammar, Braine proposed instead that children begin acquiring language by learning formulae of limited scope, along with specifications for the addition and substitution of open class words to create novel utterances (Braine, 1976). He also comments on the sharp contrast of these findings with those of Brown (1973), which support the innate capacity for syntactic productivity proposed by Chomsky.

Lois Bloom (1970) criticises pivot grammar, because it cannot account for situations in which children use the same pivot word or whole utterance in different contexts and with different context-dependent meanings. She described two instances in which her participant Kathryn said, "Mummy sock". On one occasion, this referred to Kathryn's mother's sock, while on the second occasion, this referred to Kathryn's mother putting on Kathryn's sock. She also demonstrated instances in which the word *no* (apparently a pivot) had different meanings in different contexts. "No fit" meant *it doesn't fit*, "no pocket" meant *there isn't a pocket* and "no dirty soap" meant *I don't want the dirty soap*. When comparing the language usage of her three participants, Eric, Kathryn and Gia, Bloom (1970) observed that Eric's language showed

more characteristics of pivot grammar, whereas Kathryn's and Gia's language was more productive and could not be accounted for by pivot grammar (Bloom, 1970). In her observations of this and other individual differences between the three children, Bloom (1970) was one of the first linguists to contradict the conclusions of universality drawn from research by Brown (1973) based on Chomskyan theory. She suggests instead that children use different strategies to acquire language, and that the strategy employed by the individual child results from interactions between their perceptual and cognitive function and their linguistic and non-linguistic experience.

Katherine Nelson (1973) observed the language acquisition of 18 children from the age of one to two years. She found considerable individual differences in the children's language learning strategies, but concluded that the strategies employed could be broadly categorised either as referential or expressive. She defines the referential strategy as involving a tendency towards higher intelligibility, single word production, a high proportion of nouns in early utterances and rapid increase in vocabulary during the second year of life. In contrast, children employing an expressive strategy attempted to approximate whole phrases (such as questions and comments) and were consequently less intelligible. However, these children acquired vocabulary at a slower rate than the referential children, producing a higher proportion of pronouns and a lower proportion of nouns. Children who employed a predominantly referential learning strategy formed the majority of Nelson's participants, while those employing a predominantly expressive strategy formed a large minority. Nelson additionally concluded that the learning style adopted by individual participants reflected the input which they received from their caregivers. She observed that referential children's caregivers often used language to name objects, while the expressive children's carers used more multi-word utterances for social purposes (Nelson, 1973).

In addition to the evidence which Nelson found for two distinct language learning strategies, she identified a number of demographic, social and environmental factors which were associated with the children's overall rate of language acquisition. Referential children were

more often first-born and/or came from highly educated families, whereas expressive children were more often later-born and/or came from less well educated families. She found that girls' linguistic development progressed more rapidly than that of boys. She also found that children who experienced more outings and who were exposed to more adults progressed faster than those who experienced fewer outings and who were exposed to fewer adults. In contrast, increased television watching, exposure to more children relative to adults and a commanding style of maternal interaction contributed to a lower rate of language acquisition. Although Chomskyan linguistic theory was most influential at the time of this study, Nelson interprets her findings in terms of a cognitive, construction-based approach to grammatical development. She views language acquisition as resulting from a combination of factors including the child's underlying cognitive abilities and understanding of concepts, social interaction and the physical environment.

Following the findings of Nelson (1973), Ann Peters (1977) suggests that these different language learning strategies can occur at different periods within an individual child. She points out that research into language acquisition has traditionally been interpreted according to the expectations of Chomskyan theory, which proposes an innate capacity for syntactic creativity. She specifically questions a number of assumptions which have underpinned language acquisition research. She questions Brown's (1973) assumption that language acquisition progresses steadily and consistently from simpler to more complex universal stages. She also questions the Chomskyan assumption that child language can be analysed in the same units as are applied to adult speech (including phonemes, morphemes and words). She argues that data which does not fit these expectations has been too hastily dismissed. This includes data from unintelligible children, owing to the inherent difficulties involved in the phonetic transcription of such data. This means that the language acquisition data used to support nativist approaches has only been obtained from a selected portion of the total population, that is those children who are talkative, non-imitative and intelligible.

In order to gain some insight into these issues, Peters (1977) studied the verbal development of a child (Minn) from seven months until 2;3 years of age. From eleven to twelve months, most of Minn's speech comprised imitative utterances. From 14 months, he developed a repertoire of single words which were produced with relatively high segmental accuracy. She interprets this as evidence for an analytic learning strategy at this stage, because Minn's utterances appeared to have been constructed from the individual contrastive phones found in adult speech. From 17 months, Minn's language style changed radically, as he began to produce strings which appeared to approximate multi-word utterances in terms of prosody. However, these utterances were extremely unintelligible owing to a great reduction in segmental accuracy which often rendered them indistinguishable from babble. Peters was consequently reliant on Minn's mother and contextual cues in order to understand Minn's speech. Filler syllables comprising isolated schwa vowels often occurred in these utterances and appeared to have a phonological or morphosyntactic function, although their exact meaning on any given occasion was often difficult to determine. Peters interprets these changes as evidence for a gestalt or holistic learning strategy employed at this stage, because the presence of syntagmatic fluency and relative absence of paradigmatic accuracy indicates that these utterances were being produced as one linguistic unit, rather than being formed from series of units such as phonemes, morphemes or words. From this point onwards, Minn produced a combination of analytic and gestalt speech. Analytic speech occurred in referential contexts, for instance when naming objects during activities such as reading, whereas gestalt utterances were produced during interaction and performed social functions such as requests, summonses and commands.

This interpretation of the analytic and holistic strategies can be equated with Nelson's referential and expressive strategies respectively, both in terms of the form taken and function performed by utterances resulting from each strategy. Peters suggests from her findings that the analytic and gestalt learning strategies exist on a continuum with individual variability in the extent to which the two strategies are employed. She concludes that the ultimate goal is

analytic language processing, which is much harder to attain for those children with a predominantly holistic style.

Peters also considers the type of input provided by caregivers as a possible influence on language learning strategy. Minn's mother spoke to him using longer and more complex sentences than is typical of child-directed speech. It is therefore possible that Minn's approximations of whole sentences was a direct consequence of the type of input which he received (Peters, 1977). Peters (1983) elaborates her argument for the importance of the input which the child receives for their language learning style. She points out that the child is not exposed to a dictionary of words and morphemes in the input, but to intermittent streams of speech. From these streams, the child extracts chunks of sounds with some information about their meaning. These chunks are then stored and made available for the child to use. However, because the chunks extracted may comprise several words in the adult language, this gives rise to formulaic utterances (Peters, 1983).

In a later review of the analytic and holistic strategies, Nelson (1981) acknowledges in the light of Peters' (1977) findings that the individual differences between children which she previously observed may in fact be differences occurring within the same children at different stages and communicative contexts. She concludes in agreement with Peters (1977) that the strategies exist on a continuum, but summarises a number of functional psychological distinctions between them. She argues that the high proportion of single word utterances and nouns characteristic of the analytic or referential strategy indicates that this strategy serves a cognitive function. In contrast, the production of formulaic multiword utterances and high proportion of pronouns relative to nouns characteristic of a holistic or expressive strategy reflects a social, pragmatic function (Nelson, 1981).

In a longitudinal study of four children from the onset of single-word utterances until an MLU of 2.5 was achieved, Bloom, Lightbown and Hood (1975) identified two distinct language learning strategies similar to those identified by Nelson (1973). They identified that before

reaching an MLU of 2, the girls produced a relatively high proportion of specific nouns. They identified these children as using a nominal style. In contrast, the boys appeared to use a higher proportion of pronouns and spatial deictics (Bloom et al., 1975). The authors identified this as the pronominal style. Once an MLU of 2 had been reached, these two distinct systems appeared to converge and become integrated. Once an MLU of 2.5 had been reached, the girls had learned to substitute many different pronouns for nouns and vice versa.

The referential and expressive styles proposed by Nelson (1973) have been likened to the nominal and pronominal styles of Bloom et al. (1975) respectively. However, Bretherton, McNew, Snyder and Bates (1983) point out that although these two distinctions are superficially similar, their underlying conceptualisations are very different. The distinction between nominal and pronominal styles made by Bloom et al. has a grammatical basis and the two styles are viewed as different linguistic methods of communicating, but ultimately with the same communicative aims. In contrast, the referential and expressive strategies proposed by Nelson (1973) are viewed as having different communicative functions. The referential strategy is characteristic of children who tend to label objects and events, whereas the expressive strategy is characteristic of children who engage in more social communication (Bretherton et al., 1983).

Bretherton et al. (1983) assessed the early grammatical usage of 30 children in an investigation of the nominal/pronominal styles proposed by Bloom et al. (1975). They categorised the behaviours which they observed into two grammatical clusters. The multi-word referential cluster was considered equivalent to the nominal style and was characterised by a high proportion of nouns and telegraphic utterances. The grammatical morpheme cluster was considered equivalent to the pronominal style and was characterised by higher proportions of pronouns, inflectional morphemes, articles, prepositions and auxiliaries. As predicted, Bretherton et al. found that labelling correlated highly with the referential cluster. However, they found that pivot utterances and imitations also correlated highly with the referential cluster, when they would have expected these behaviours to correlate more highly with the

grammatical morpheme cluster. Behaviours characteristic of both clusters were observed within individual children. They therefore concluded that although the two clusters represent a degree of distinction between the two styles, they are only partially dissociable and show some continuity. They therefore argue that children should be placed on a continuum which represents the relative contribution of each style to their learning, rather than discretely categorised as employing one style exclusively.

Nelson's interpretation of the expressive strategy has been criticised on the grounds that it has mainly been defined in terms of the absence of features which characterise the referential strategy, such as reduced intelligibility, slower rate of vocabulary development and fewer than 25 nouns among the first 50 words (Lieven, Pine and Barnes, 1992). Lieven et al. (1992) argue that this negative definition has led to a view that the expressive strategy is an inferior or atypical route to language acquisition compared with the more commonly occurring referential strategy. They propose instead that the two routes are equally valuable in learning, although qualitatively different. Their study of twelve children's early language acquisition therefore aimed to identify the positive characteristics of the expressive strategy. They also compared children's linguistic strategies at the same stage of vocabulary development, in contrast with the age comparisons conducted by Nelson (1973).

Lieven et al. (1992) and Lieven, Pine and Baldwin (1997) longitudinally studied twelve children from approximately one year of age (the point at which 20 words or utterances had been acquired) until three years of age, using a combination of maternal diaries and audio recording. Single words were classified according to formal grammatical categories. Multi-word utterances were classified either as frozen (indicating a formulaic strategy), semi-frozen (partially analysed) or constructed (indicating an analytic strategy). The latter two categories indicated a lesser or greater degree of productivity respectively. The three categories were determined by the extent to which individual words in the utterances had been observed to occur independently of each other and in the same position across different multi-word utterances. Thus, the words occurring in frozen utterances were not found to occur

independently, while the words occurring in constructed utterances were also found to occur independently and in a wider range of word positions.

Lieven et al. (1992) found individual differences in both vocabulary composition and the rate of vocabulary growth at the 50 and 100 word stages. Although Nelson (1973) defined proper nouns and social words as positive features of the expressive strategy, Lieven et al. (1992) found that the proportions of such words in children's vocabularies decreased significantly between 50 and 100 words. This suggests that the predominance of such words is a feature of very early vocabulary, rather than a more general feature of some children's learning styles. They found significant negative correlations between proportions of nouns and frozen phrases. They therefore agreed with Nelson that a high proportion of nouns is characteristic of the referential strategy, but conclude, in contradiction to Nelson in their assertion that frozen phrases are a better indicator of the expressive style than are social words. They also found positive correlations between the proportion of frozen utterances and potentially productive utterances at both the 50 and 100 word stages. This finding was replicated at the 100-word stage (Pine and Lieven, 1993), who additionally observed qualitative relationships between unanalysed and productive utterances in terms of the specific lexical items carried over. Lieven et al. (1992) therefore conclude in contradiction with Nelson (1973) that an expressive learning strategy which employs unanalysed utterances is an alternative route to the development of syntactic creativity, rather than a less efficient strategy than the referential strategy.

Pine and Lieven (1993) elaborate this conclusion, suggesting that more referential children derive multi-word utterances by combining their knowledge of single words. In agreement with Peters (1983), they suggest that more expressive children discover that the unanalysed utterances known to them have variable slots for lexical items, and that they derive productive multi-word utterances by gaining control over these variable slots (Pine and Lieven, 1993). Lieven et al. (1992) further conclude in agreement with Peters (1977) that previous research has focused exclusively on the analytic or referential strategy and that further research is

needed to understand the contribution of frozen utterances to the development of syntactic creativity.

Following evidence from bilingual child language data, Vihman (1999) argues that some apparently frozen phrases may in fact involve syntactic creativity. Vihman observed that similarly to the participants of Lieven et al. (1992, 1997), her son produced utterances in which the composite individual words did not occur independently in other contexts. However, six out of the 31 apparently frozen phrases contained words from both English and Estonian. She argues that such mixed language utterances provide evidence that her son was actively processing and combining units from separate sequences in the input which he received.

The concept of referentiality has also been questioned (Pine, 1992). In a comparison of maternal report and observational measures of vocabulary composition, Pine found that although nouns formed the most frequent grammatical category in children's vocabulary at the 50- and 100-word stages, they constituted only 30% of word tokens in the children's utterances. He also found that maternal report was biased towards common nouns compared with observational measures, and attributed this to reduced sensitivity to other word types. A possibility not suggested by the author is that mothers were better able to identify their children's nouns than other word types owing to the contextual cues provided by concrete referents. Pine concludes from these findings that the distinction between referentiality and non-referentiality may be of limited use in the description of language learning strategies. The concept of a continuum of analytic and holistic language learning therefore appears to be more useful.

In a more recent review, Peters (1995) expands on her interpretation of analytic and holistic learning. She argues that the analytic strategy occurs when children begin by learning open class lexical items and then learn to fill the gaps between these using grammatical words. In contrast, she views the gestalt strategy as involving reliance on formulaic syntactic frames,

using filler syllables as protomorphs before the words or morphemes which could optionally replace the filler syllables have been acquired. Similarly to Nelson (1981) she concludes that both analytic and holistic learning may be simultaneously evident in the same child across different linguistic domains (Peters, 1995).

In a case study of her bilingual daughter, Virve, Vihman (1982) reports differences in the usage of formulae both across the two languages being acquired and within the same lexical item in different constructions. Virve's first language was Estonian, but she was then regularly exposed to English at a daycare centre from 21 months of age. Many of her early English productions were multi-word utterances, learned as unanalysed lexical items; these included "happy birthday to you" and "that's mine". Vihman compares these findings with those for her son, who was exposed to English from six months and who acquired only one English multi-word utterance in his first 50 words. She explains Virve's high usage of formulas as the result of acquisition without prior exposure to and comprehension of the internal structure of English.

Vihman (1982) also observed parallel analytic and holistic learning of the lexical item *want*. At 24 months of age, Virve produced *want to* either as *wanna* or *wannu*. However, in the construction *I want more*, she pronounced *want* as [vɒnt]. From 25 to 26 months of age, she produced many constructions with *wanna/wannu*, for example, "I wanna down" and "I don't wannu potty". There were no further productions of *want* until 28 months, when it re-emerged as [ɯɒnt]. Vihman interprets these findings as evidence that parallel analytic and holistic learning may occur in the same child for the same lexical item in the adult language.

In a review of studies which have found evidence of analytic and holistic language learning strategies, Bates, Dale and Thal (1995) summarise fundamental differences in the language development profiles of children with a tendency towards one or other of these strategies. Children employing a more analytic language learning strategy are described as having the following linguistic characteristics:-

- Word orientated, with many nouns present in early vocabulary;
- Consistent in their application of syntactic rules;
- Likely to produce novel, creative multi-word utterances;
- Showing rapid syntactic development and vocabulary growth;
- Segment-focused in their phonology, producing intelligible speech with consistent pronunciation.

In contrast, children employing a more formulaic language learning strategy are described as having the following linguistic characteristics:-

- Likely to produce imitative formulae rather than novel word combinations;
- Producing a high proportion of pronouns;
- Inconsistent in their application of syntactic rules;
- Showing relatively slow development of vocabulary and syntax;
- Intonation-orientated in their phonology with relatively low segmental accuracy and variable pronunciation (Bates et al., 1995).

Bates et al. (1995) summarise a number of factors which may contribute to these individual differences. These include demographic and environmental variables such as gender, socio-economic status (SES) and caregiver input style, and endogenous differences such as temperament and intelligence. They also argue that the analytic/holistic distinction may be too broad and simplistic a generalisation which is made when children are showing different developmental profiles in terms of speed and accuracy across linguistic levels. They propose an alternative explanation that all children may learn language analytically, but that they may differ in the size of linguistic units which they are able to process, owing to individual differences in memory and speech processing capacity at different stages.

More recent research has moved away from a binary distinction between expressive versus referential learning or analytic versus holistic learning. A recent study by Fernald and Marchman (2012) tracked the vocabulary outcomes in typically developing (TD) children

versus late-talking children, using similar measures to those used by Nelson (1973). However, they report a range of individual differences between children, rather than a dichotomous distinction between groups of children. They measured the children's vocabulary at 18, 21, 24 and 30 months of age. Late-talkers were defined as those children who showed vocabulary scores below the normal range at 18 months. The researchers also administered a spoken word linguistic processing task at 18 months, in which the child was asked to identify the picture which matched the spoken word stimulus. Both accuracy and speed of linguistic processing were measured. The researchers' aim was to determine whether linguistic processing and vocabulary measures at 18 months of age predicted vocabulary outcomes at 30 months.

Fernald and Marchman (2012) found that the late-talkers had significantly lower vocabulary scores than the TD children at all four points in time. However, there were considerable individual differences among the late-talkers. Whereas 14 out of the 26 late-talkers still showed significantly delayed vocabulary development at 30 months of age, 26 showed periods of accelerated vocabulary growth, achieving scores within the normal range by 30 months. The late-talkers showed significantly slower and less accurate performance on the linguistic processing task than the TD children. Those late-talkers who performed faster and more accurately showed significantly higher vocabulary outcomes at 30 months and a more accelerated developmental trajectory than those who performed less well. Delayed vocabulary at 30 months was thus largely predicted by low vocabulary scores and less efficient linguistic processing at 18 months. However, nine of the late talkers who showed slower linguistic processing at 18 months achieved vocabulary scores within the normal range at 30 months. These findings demonstrate that individual differences existed within an apparently homogeneous group of late-talkers. While some showed persistent delays in vocabulary development throughout the study, others showed variable periods of slow and accelerated learning, eventually achieving vocabulary scores within the normal range at 30 months.

Perkins (1999) also argues that the analytic/holistic distinction is too simplistic, acknowledging that individual differences in language acquisition result from the combination of individual cognitive and language processing abilities. Whereas Peters (1977) argues that the child's ultimate goal is towards an analytic strategy, Perkins argues that an adult formulaic strategy is ultimately attained. He points out that adult language contains many formulaic utterances, which are used on multiple occasions in order to reduce effort and processing load on the part of both speaker and listener. He therefore hypothesises that children progress from a formulaic strategy to an analytic and creative one and then back to a formulaic strategy. The analytic strategy is thus viewed as a stepping stone between the immature and mature formulaic strategies. However, unlike the initial formulaic strategy evident in some young children during early language development, the adult formulaic strategy requires creative and analytic proficiency and is therefore adopted only when the child has mastered syntactic productivity (Perkins, 1999). Perkins emphasises that research is needed into the acquisition of adult formulaicity.

Wray and Perkins (2000) present a modification of this view of adult formulaicity, proposing that adult language processing involves a balance between the more frequently used holistic processing strategy and the less frequently used analytic strategy. The interlocutor (unconsciously) adopts the strategy appropriate for each communicative situation, depending on the requirements of the social interaction and memory limitations on linguistic processing. Whereas the more frequent holistic strategy is suitable for most occasions and serves to reduce processing load, the analytic strategy is sometimes needed for communication in difficult communicative situations, such as background noise, or when compensating for the differences in accent and grammar produced by a non-native speaker (Wray and Perkins, 2000). This concept of balance and compromise between the analytic and holistic strategies corresponds with Peters' (1983) suggestion that the units of adult language may comprise both small units such as morphemes and some large units such as frequently used utterances.

Joan Bybee (2000, 2002) also supports the storage of frequent or formulaic utterances as single lexical items. However, her accounts of these phenomena take a more biological, motoric perspective, focusing on the neuromotor processes which underpin speech production. She greatly emphasises the influence of formulaic utterances on language processing at both lexical and phonological levels. She observes that high frequency words and multi-word utterances are produced with greater phonological reduction than those of lower frequency. For example, she notes that high frequency words such as *memory* and *summary* are subject to schwa deletion, whereas phonetically similar low frequency words such as *mammary* and *summery* are not. She refers to this lexically-dependent sound change as lexical diffusion and proposes that it results from the impact of frequency of usage on lexical storage and cognition (Bybee, 2000).

Bybee also notes that high frequency multi-word utterances are produced with more connected speech processes (CSPs) than those of lower frequency. She explains this phenomenon in terms of exemplar storage, whereby high frequency utterances are stored in memory as single units with unique neuromotor routines and are therefore retrieved as articulatory gestalts during speech production (Bybee, 2002, 2006). This happens when repeated practice in producing high frequency utterances leads to increased articulatory speed and fluency, as the extent of articulatory movement decreases and the degree of gestural overlap increases (Bybee, 2002; Bybee and McClelland, 2005). Bybee places the concepts of lexical diffusion and exemplar storage in the context of a usage-based model, in which grammar is viewed as "the cognitive organization of one's experience with language" (Bybee, 2006, p. 711). The following section discusses usage-based models in more detail. (Usage-Based Phonology is discussed specifically in Chapter Three).

1.7. Usage-Based Models of Language Acquisition

There are several different usage-based models, but they all have similar underlying principles (Bybee, 2006; Langacker, 2000; Lieven and Tomasello, 2008). Such models are non-nativist

and focus instead on the child's ability to acquire language through domain-specific and domain-general cognitive mechanisms in response to input from other speakers and their own usage. Linguistic competence and performance are therefore viewed as equivalent. These models have been greatly influenced by the research on formulaicity discussed above and they seek to account for the individual variability which has been found. Language is viewed as similar to other forms of learning and requires the same cognitive mechanisms, such as memory and motor planning (Sosa and Bybee, 2008). When a child detects similarities between linguistic forms or exemplars which they hear, they gradually extract general linguistic schemas from which linguistic knowledge emerges. Learning is data-driven and linguistic knowledge develops both as a cause and a consequence of the child's usage (Sosa and Bybee, 2008). The relative frequency of specific constructions which the child experiences impacts directly on the child's cognitive representations of language. This is how high frequency utterances come to be stored as single units and are used as formulae (Bybee, 2006).

Sosa and Bybee (2008) extend this usage-based model to phonological development. In their model of cognitive phonology, they propose that phonological development is derived from similarity relationships between items which the child hears. They take a holistic view of language development and emphasise the importance of relationships between different linguistic levels for the learning process, including phonology, grammar and the lexicon.

1.8. A Usage-Based, Constructivist Approach to the Acquisition of Grammar

The usage-based approach which forms a major focus of the current study is a constructivist theory proposed by Lieven and Tomasello (2008). This theory has directly resulted from the earlier evidence of frozen phrases reported by Lieven et al. (1992, 1997) and other associated research on formulaicity in child language. The principle concept of this theory and other

usage-based approaches is that the child acquires grammar from the specific utterances which they hear using domain-general cognitive abilities (Lieven and Tomasello, 2008).

Tomasello argues that the uniqueness of linguistic ability to human beings does not necessarily mean that language is genetically pre-specified. He points out that many other activities such as cooking are also specific to humans, which have developed over centuries and are not the result of genetic endowment. Instead, he suggests that linguistic commonalities between people result from the ability to symbolise and from mutual experience of the world (Tomasello, 1995). One issue not considered by Tomasello is the evolutionary changes which have also been essential for the development of human speech. According to Fitch (2000), two major evolutionary contributions to the human speech capacity have been the development of the current human vocal tract configuration and the ability to produce vocal imitations of the sounds we hear.

Instead of learning abstract grammatical rules by setting parameters as suggested by Chomsky, children draw on their innate domain-general cognitive abilities in learning concrete linguistic constructions, hence the term *construction-based approach* (Tomasello, 1998). The child constructs their own grammatical knowledge from the linguistic constructions which they hear in the context of specific events. These instances are known as usage events and this approach is therefore known as a usage-based approach (Lieven and Tomasello, 2008). The constructions learned are initially lexically-specific and are not generalised to similar syntactic structures. Constructions also vary in terms of unit size, from whole utterances such as declarative, imperative and interrogative clauses, to smaller units such as clause elements and individual morphemes (Lieven and Tomasello, 2008; Tomasello, 2000B). The constructions are learned as wholes, and children's early utterances are believed to be approximations of multi-word constructions, rather than single words. The internal structures of early constructions are gradually analysed as the child learns the relationship between a word's form and function. The child identifies common patterns between constructions and parts of constructions and learns which elements can be substituted within a specific construction.

Thus, the child learns that utterances can act as pivots or frames with variable slots for the substitution of elements, for instance *It's a (noun phrase)* (Lieven and Tomasello, 2008; Tomasello, 2000B). These processes enable the gradual abstraction of grammatical knowledge such as tense, agreement, clause and phrase elements and thematic roles. The child's grammatical knowledge thus extends from lexically-based constructions to more general, abstract schemas for more creative use. The schemas are strengthened over time and experience, a process known as entrenchment (Lieven and Tomasello, 2008). The units of language which the brain can process at any one time are not pre-specified by this theory, thus accounting for variability across different syntactic skills within an individual child and across different children (Tomasello, 2000B). Lieven and Tomasello (2008) acknowledge that humans have an innate capacity for language learning, but emphasise the unique combination of cognitive mechanisms which have developed in humans to make this learning possible.

Evidence for the lexically-specific nature of early constructions comes from a number of studies which have shown that children demonstrate restrictions on the usage of specific constructions, which gradually reduce, leading to more generalised usage as a function of age. Tomasello (2000A, 2000B, 1992) observed that his daughter's early use of individual verbs was restricted to specific syntactic constructions. Semantically similar verbs were not found to be related in terms of the numbers and types of constructions in which they were used. This led Tomasello to propose that language acquisition is initially centred around the learning of verb-specific constructions with open slots for object noun phrases, a proposal known as the *Verb Island Hypothesis* (Tomasello, 1992, 2000A).

This observation has been supported by a number of experimental studies which compared children's ability to use verbs creatively as a function of age. Akhtar (1999) found that when children were exposed to novel verbs (nonsense words) in transitive constructions with ungrammatical word orders, four-year-olds were significantly more likely than two-year-olds and three-year-olds to correct the word order to subject, verb object (SVO) when reproducing the verbs. The two younger groups were equally likely to produce sentences with correct and

incorrect word orders. The author interprets this finding as evidence against the strong version of Tomasello's *Verb Island Hypothesis*, which would predict that the younger children would only reproduce the novel verbs in sentences with incorrect word orders (Akhtar, 1999).

In a similar study which compared younger two-year-olds with a group of three- and four-year-olds, Abbot-Smith, Lieven and Tomasello (2001) found that the younger children were more likely to correct the incorrect word order of constructions containing familiar than nonsense verbs. In contrast, the older children corrected the word order of constructions containing either verb type. The authors conclude from these results that two-year-olds' knowledge of word order is linked to specific lexical items, whereas three- and four-year-olds have acquired more general knowledge of word order which they can apply to novel verbs (Abbot-Smith et al., 2001). They therefore argue that general syntactic schemas develop gradually from lexically-based schemas.

In a later study of word order comprehension, Gertner, Fisher and Eisengart (2006) presented 21-month-olds and 25-month-olds with a series of transitive constructions. The test phase involved sentences containing familiar verbs and the experimental phase involved novel verbs. The children's looking preferences to one of two videos were observed for each sentence as a measure of word order comprehension. The children looked for significantly longer at the video which matched the word order of each sentence, a finding which remained constant across familiar and novel verb conditions. The authors interpret their findings as evidence which refutes construction-based accounts such as that of Lieven and Tomasello (2008) and Tomasello (2000A). Instead, they argue that their findings demonstrate that the children applied innate, abstract syntactic knowledge in order to interpret novel verbs through syntactic bootstrapping (Gertner et al., 2006).

There appears to be considerable conflict between the findings of Gertner et al. (2006) and those of Abbot-Smith et al. (2001) and Akhtar (1999). The former appear to demonstrate that two-year-olds can apply general, abstract syntactic knowledge in the interpretation of novel

verbs, whereas the latter demonstrate that sensitivity to correct word order increases and generalises to a wider range of novel verbs as a function of age. One reason for these differences is that there may be a discrepancy between two-year-olds' comprehension and production abilities, consistent with the finding that comprehension develops ahead of production (Bates et al., 1995). In terms of the construction-based approach of Lieven and Tomasello (2008) and Tomasello (2000A), children may go through a stage when they have abstracted general schemas for word order at the level of comprehension, while their production remains constrained by the lexically-specific verb constructions which they have experienced most frequently (see below for a discussion of frequency effects). Gertner et al. (2006) employed a practice phase with familiar verbs before testing the children on novel verb comprehension. It is therefore also possible that the practice phase primed the children to a certain response in advance of the experimental phase. It would be interesting to observe the results in a control comparison wherein the novel and practice phases are presented in reverse order. A further possibility is that the children in the study of Gertner et al. (2006) were employing syntactic bootstrapping as suggested by the authors, but that syntactic bootstrapping is one of the processes by which construction-based knowledge is incorporated into general schemas, thus rendering the two theories mutually compatible in explaining these findings.

A further prediction of the usage-based approach is that high frequency words play a greater role in the development of schemas than do low frequency words, because they are present in a greater number of usage events (Matthews, Lieven, Theakston and Tomasello, 2005). Following the findings of Abbot-Smith et al. (2001) and Akhtar (1999), Matthews et al. (2005) tested this prediction by presenting two- and three-year-olds with sentences containing either, high, medium or low frequency verbs in constructions with an incorrect SOV word order. They found that two-year-olds corrected the word order of sentences containing high frequency verbs more than those containing low frequency verbs. In contrast, the three-year-olds produced many corrections across frequency conditions, demonstrating a reduced effect of the frequency of specific lexical items on their knowledge of word order. The researchers conclude that the acquisition of word order and transitive constructions occurs through the

entrenchment of schemas and that schemas for high frequency words are entrenched earlier than those for low frequency words, giving rise to more grammatically correct productions of utterances containing these words.

In another study which led to a similar conclusion, children aged two to five years were presented with sentences which contained either a correct or incorrect embedded clause and either a high or low frequency verb (Kidd, Lieven and Tomasello, 2006). The children were more likely to produce accurate repetitions of correct sentences and to correct incorrect sentences which contained high than low frequency verbs. In addition, more correct repetitions were produced by the older children. The authors conclude from these findings that early knowledge of permissible syntactic structures is related to specific lexical knowledge, and that early lexical knowledge is related to word frequency.

Another element of the usage-based approach is the emphasis on input received by the child to create usage events from which schemas are gradually formed. Childers and Tomasello (2001) found that training two-year-olds to use verbs in transitive constructions facilitated the independent production of trained verbs and to a lesser extent, the production of novel verbs in these constructions. These results show the importance of adult models for the formation of general schemas (Childers and Tomasello, 2001). These results were replicated, with the additional finding that four-year-olds were better able to use novel verbs following training than two-year-olds (Abbot-Smith, Lieven and Tomasello, 2004). The authors conclude that these age differences demonstrate how schemas for syntactic structure are acquired over time with increased exposure to exemplar utterances from the input.

In order to investigate the type of input on children's ability to learn complex constructions, three-, four- and five-year-olds were trained to produce sentences with object clefting using one of two training regimes. Massed exposure involved intensive training prior to the experiment, whereas distributed exposure involved less intensive training over five days prior to the experiment. A control group received very brief training prior to the experiment which

was sufficient for them to understand the task required of them (Ambridge et al., 2006B). The children exposed to the distributed training performed better than either the massed exposure or control groups. The authors conclude from these findings that language learning benefits from the distributed presentation of stimuli in common with other aspects of learning, indicating that linguistic and other skills may be acquired through some shared cognitive mechanisms. The temporally distributed input which appears to optimise linguistic learning matches the type of input to which the child is naturally exposed, thus explaining how a child may learn from the positive evidence of grammatically correct productions, despite the absence of specific parental correction of the child's erroneous utterances, as discussed by Pinker (1994) and Marcus (1993). The results of these studies therefore support the findings of Huttenlocher et al. (1991) and Barnes et al. (1983) mentioned earlier, in emphasising the important role of parental input in language acquisition.

As well as the controlled, experimental studies described above, Lieven and colleagues have conducted a number of studies based on densely collected spontaneous child language data. A major finding emerging from these studies is that children's early utterances are not as creative as suggested by Chomsky (1976; 1967A; 1964), thus providing further evidence for the role of formulaicity in language acquisition (see previous section). Lieven, Behrens, Speares and Tomasello (2003) analysed the syntactic creativity of a two-year-old girl (Annie) from six weeks of densely sampled data. All of the utterances which Annie produced in the final hour of data were syntactically compared with all those produced previously in the sample. They found that only four of the 232 single-word utterances produced in the final recording session were novel. Of the 295 multi-word utterances, 63% were not novel, comprising utterances which had been produced in previous recording sessions, self-repetitions and imitations of utterances produced by the mother. Of the 37% of utterances which were novel, 74% differed from previous utterances by only one operation, defined as a change in syntactic structure involving the substitution, addition or deletion of an element. More than one operation was required to derive the remaining 26% of novel utterances from previous utterances.

These findings were replicated in a comparison of the multiword utterances produced by four two-year-olds including Annie and Brian (Lieven et al., 2009). In addition, Lieven et al. (2009) found that the number of repetitions decreased and the number of utterances derivable by more than one operation increased as a function of MLU. Similar results were found in relation to the development of Annie and Brian's question forms (Dabrowska and Lieven, 2005). In addition, Annie produced many more creative questions than Brian, reflecting her generally superior linguistic performance to Brian at two and three years of age.

The authors of these studies conclude in support of Tomasello (2000A) that the young child initially stores the constructions which he/she has experienced in the form of frames with variable slots which allow changes to the syntactic structure (Lieven et al., 2003). The child's increasing creativity with age and MLU demonstrates a gradual decline in the child's reliance on lexically-specific, stored constructions as the child's working memory capacity and ability to use abstract schemas increases (Dabrowska and Lieven, 2005; Lieven et al., 2003).

The findings of these studies and the conclusions drawn from them are remarkably similar to the observations of Clark (1973), which were made at a time when Chomsky's theory was still greatly influential and cognitive-linguistic theories such as that of Lieven and Tomasello (2008) had not yet been proposed. They are also connected with previous observations of formulaicity in child language acquisition, which were not based on usage-based approaches (Bates et al., 1995; Nelson, 1973; Peters, 1977; Wray and Perkins, 2000). The evidence shows that children can initially produce language without detailed knowledge of its syntactic structure, using concrete constructions learned from usage events, prior to the abstraction of general linguistic schemas which give rise to creativity.

There is a considerable body of empirical evidence in support of the predictions made by the construction-based approach. The linguistic knowledge of two-year-olds appears to be initially linked to specific constructions, but is better generalised to novel contexts by older children. Two-year-olds appear to show limited syntactic creativity in their early productions

and employ the constructions which they have previously experienced in the production of novel, syntactically similar utterances. This approach also accounts adequately for linguistic learning in the absence of parental feedback, by explaining how children can learn from temporally distributed positive evidence of correct language usage.

1.9. Summary and Conclusions

Nativist theories of language acquisition have the advantage of accounting for the uniqueness of language to humans, its acquisition in the absence of explicit instruction and its creative capacity. However, the notion that caregiver input is relatively unimportant in the language acquisition process has been refuted by research from psychological and construction-based linguistic perspectives. The distinction between linguistic competence and performance and the extent of early syntactic creativity proposed by Chomsky have also been brought into question. Research has shown that children's syntactic creativity is limited and that many early multi-word utterances are either entirely formulaic or closely derived from formulae which the child has previously used or heard. There are considerable individual differences and intra-speaker variability in the extent to which children use language analytically or holistically at different developmental stages. These differences are associated with endogenous factors such as temperament and intelligence, along with a range of demographic and socio-economic factors.

Nelson (1973) has categorised individual differences in language learning strategy along a continuum of referentiality versus expressiveness. However, Lieven et al. (1992) have criticised Nelson's concept of expressiveness, owing to the lack of a positive definition and the assumption that this style of learning is inferior to the referential style. Pine (1992) has also criticised Nelson's definition of referentiality, following his observations that nouns constitute only a small proportion of tokens in early multi-word utterances, and that maternal report

often overestimates the extent to which children use nouns. For this reason, Peters' distinction between analytic and holistic strategies may be more useful.

Although there is clear evidence for the existence of analytic and holistic language learning strategies, it would be too simplistic to make a categorical, binary distinction between them. A number of suggestions have been made to account for the complex inter- and intra-speaker variability which has been observed. Bates et al. (1995) argue that all children learn analytically, but that cognitive differences affect the size of units which they are able to process. Perkins (1999) argues that the child's ultimate goal is to acquire an adult formulaic strategy, and that the immature formulaic and analytic strategies are both important in the developmental process. Wray and Perkins (2000) argue that a combination of analytic and formulaic language processing is necessary for communication, and that the interlocutor must achieve a balance between the two according to the communicative situation.

In conclusion, there is considerable debate as to how analyticity and formulaicity operate both in language acquisition and in adult communication. However, it is evident that formulaicity and individual differences exist in language acquisition and that these phenomena need to be accounted for. The construction-based approach of Lieven and Tomasello (2008) currently provides the most useful account, owing to its focus on learning, cognition and the input which the child receives.

Chapter Two

The Phonology and Acquisition of Auxiliary Syntax

2.1. Introduction

Following the exploration of theories of language acquisition in Chapter One, this chapter explores specific research on auxiliary syntax. The auxiliary verbs *can* and *can't* form the primary syntactic focus of the current study and are also of phonetic and phonological interest. It is therefore necessary to understand theories which have been proposed concerning the phonological representations of auxiliary verbs and their developmental trajectories of acquisition. This research is discussed in chronological order, starting with theoretical approaches to auxiliary phonology. Most of these accounts were proposed in the seventies and eighties, evolving from a generativist linguistic framework. There then follows a discussion of the empirical research on the developmental trajectories of auxiliary syntax. This research has been conducted in recent years and is evolving from the usage-based constructivist approach to language acquisition.

2.2. The Phonetic and Phonological Nature of Auxiliary Verbs

Many English words have several possible phonetic forms, depending on the utterance context in which they occur (Simpson, 1992). These are known as strong and weak forms and can be found in a range of high frequency function words including auxiliary verbs. Strong forms are those which occur when words are stressed in an utterance or are spoken in isolation, rendering them prosodically independent. They have full vowels (monophthongs or diphthongs) and can occur as the final word in an utterance. They can also combine with the ending *n't* to produce a negative form of the verb (Ogden, 1999). They are considered to be

stored as abstract linguistic representations in the lexicon and are known as citation forms (Cruttenden, 2001).

In contrast, weak forms are unstressed forms of words with altered word shapes and vowel characteristics. Initial consonants are often elided and vowels are neutralised to produce a schwa vowel; these processes are known as phonetic reduction (Cruttenden, 2001). Weak forms cannot stand alone as citation forms, but are clitics which are dependent on the prosodic context of surrounding syllables. For this reason, they cannot occur in final position in an utterance (Ogden, 1999). These differences result from prosodic characteristics of utterances including rate and stress patterns (Cruttenden, 2001) and the communication of previously given information (Shockey, 2003). Strong and weak forms are not always syntactically equivalent (Ogden, 1999).

An example is the verb *have,* which has the strong form [hæv] and the weak forms [æv], [həv], [əv] and simply the consonant [v], which then becomes a syllable coda for the preceding pronoun, e.g. *you've* and *they've*. A Similar range of strong and weak forms has been documented for many auxiliaries, including *is*, *am*, *are*, *has*, *does*, *will* and *would* (Mackenzie, 2012; Ogden, 1999).

The literature on the phonology of auxiliary development has seldom focus specifically on strong and weak forms of *can*, although Shockey (2003) mentions these forms from the perspective of the phonetic reductions which occur in connected speech. She cites the form [kn̩], which results from syllabification of the final nasal, as it overlaps with and further reduces the schwa. It is noteworthy that the types of elision occurring in auxiliaries with initial [w], [h] or vowels do not occur in *can*, which has an initial plosive.

Much of the existing literature on the phonetic and phonological relationships between auxiliaries dates back to the 1970s and 1980s and was greatly influenced by Generative

Phonology, which was the dominant phonological theory at the time (Chomsky and Halle, 1968).

There has been considerable debate regarding the processes which underlie the alternations between strong and weak forms of auxiliaries. Strong and weak forms are traditionally considered to constitute part of a language's phonology and to have separate phonological representations in the lexicon; this is a viewpoint which continues to be reflected in modern introductory phonetics textbooks (Simpson, 1992), for example, *Gimson's Pronunciation of English* (Cruttenden, 2001). More detailed theoretical accounts of strong and weak forms have focused on elaborate phonological and syntactic processes by which weak forms are phonetically derived from strong forms (Simpson, 1992). One argument is that they are derived from strong forms by means of specific generative phonological rules (Zwicky, 1970). Examples of these rules include deletion of initial /w/ in forms such as *will* and *would*, and deletion of initial /h/ in forms such as *have* and *has*. Zwicky argues that some rules are more likely to apply than others, for example, deletion of initial /h/ is more likely to occur than deletion of /w/. Some rules are restricted to specific lexical items, while others are less restricted, although they are still more likely to occur in some lexical items than others. Ogden (1999) criticises this approach on the grounds that specific rules, such as deletion of initial /h/ and /w/ apply only to a restricted set of auxiliaries and are not applicable to weak forms of auxiliaries.

In a more general discussion of weak forms occurring in English, Zwicky (1970) also describes the likely effect of speech rate on the occurrence of weak forms. He proposes that they do not occur at all in slow, careful speech, that weak forms of some auxiliaries and pronouns can be found in speech produced at a moderate rate and that in fast speech, the rules of weak forms are further extended to more unstressed words and unstressed syllables in multisyllabic words.

In contrast, Baker (1971) explains the stress patterns of auxiliaries in terms of their position within syntactic structure. Baker proposes the generative rule of *Auxiliary Shift*, which places the auxiliary to the left of the verb phrase and also to the left of any preverbal elements, such as adverbs. Baker points out that stressed auxiliaries are often preceded by adverbs, whereas this is not necessarily the case for unstressed auxiliaries, which are often followed by adverbs. The example which he gives is: "The sea has *never* been my element and never *will* be." (Baker, 1971) (p. 168).

A further view of strong and weak forms as rooted in syntax, rather than in phonological rules, is to treat strong and weak forms as suppletive allomorphs, on the grounds that they have different syntactic distributions (Kaisse, 1983). Suppletive allomorphs are grammatical inflections of the same word which are not phonologically related, e.g. *go* and *went* (Ogden, 1999). Kaisse argues that strong and weak forms of auxiliaries are stored in the lexicon as separate items, similarly to other suppletive allomorphs.

Whereas Zwicky (1970) argues that speech rate has a direct effect on the occurrence of weak forms of auxiliaries, Kaisse concludes that the rules governing strong and weak forms are independent of speech rate; this is on the grounds that weak forms of auxiliaries occur in relatively slow speech, whereas vowel reduction in general occurs only in rapid speech. She therefore argues that the generative rule of *Auxiliary Reduction* is not strictly a rule of phonological reduction, but rather a syntactic rule stating the circumstances under which each suppletive allomorph may occur.

In a critique of Kaisse's (1983) paper, Ogden (1999) points out that the account of strong and weak forms of auxiliaries as suppletive allomorphs does not explain the phonetic and phonological similarities which exist between strong and weak forms. For example, strong and weak forms of *was* have the same initial and final consonants ([wɒz] and [wəz]), whereas *go* and *went* have no phones in common.

In a more recent account of the phonology of auxiliaries, Simpson (1992) proposes that auxiliaries can be described in terms of a phonological system specific to this class, which emphasises the commonalities between different auxiliaries. He outlines phonetic features which occur with some regularity within the class of auxiliaries which have both strong and weak forms. These include an onset of only one consonant (if at all present), monophthongal nuclei (mostly short vowels, but occasionally long vowels) and simple rimes, consisting either of a single vowel, or a vowel and single consonant. For example, he describes phonetic features which are common to the modal auxiliaries *will*, *shall* and *can*. The present tense strong forms of these vowels all have unrounded vowels and their coda consonants vary, while their past tense forms *would*, *should* and *could* have rounded vowels and the common coda /d/. Phonological approaches such as this, which attribute different rules and characteristics independently to different classes of words, are known as polysystemic approaches (Ogden, 1999; Simpson, 1992).

Richard Ogden (1999) further develops the polysystemic approach in his declarative account of strong and weak forms of auxiliaries. He attempts to explain strong and weak forms of auxiliaries within a framework of *Declarative Phonology* (DP). This theory proposes that different linguistic structures are governed by different constraints or rules. This means that constraints affecting a specific word class within a language (for example auxiliaries) need not affect other classes. This view gives rise to polysystemic analyses, which focus on consistent phonological relationships within systems and interactions between separate systems within a language. Each system has its own phonological and syntactic characteristics. DP does not support constraints which are destructive, such as the deletion of phonemes from weak forms; this contrasts sharply with Zwicky's (1970) generative rules of /h/ and /w/ deletion.

In contrast with the purely phonological approach of Zwicky (1970) and the syntactic approaches of Baker (1971) and Kaisse (1983), Ogden (1999) argues that both phonological and syntactic relationships exist within the system of auxiliaries. He views strong and weak forms as having contrastive vowels, rather than schwa having secondary status and being

derived from the vowel in the strong form by means of reduction. He proposes that strong and weak forms are stored together in the lexicon and that the phonological constraints which determine their occurrence make reference to grammatical information. Ogden acknowledges that a remaining question is the extent to which the relationship between strong and weak forms of auxiliaries is determined by linguistic structure or connected speech processes (CSPs).

Laurel MacKenzie (2012) criticises some of the early literature on the phonological characteristics of strong and weak auxiliaries (such as Zwicky (1970) and Kaisse (1983), on the grounds that it was not drawn from specific evidence in speech data. She examined strong and weak forms of *has*, *have*, *is* and *will* from a corpus of 240 hours of telephone conversations between 542 adult speakers.

MacKenzie found evidence of full (strong) forms and contracted (weak) forms, as well as intermediate forms which showed partial contraction. The distinctions between full, intermediate and contracted forms were made in relation to syllable shape rather than segmental characteristics. Full forms were defined as having an audible initial consonant and a vowel of any quality. Intermediate forms were defined as having no initial consonant, but an audible vowel. Contracted forms were defined as having no audible initial consonant or vowel, but only a single consonant. She makes no distinction between those auxiliary verbs with a full vowel and those with a schwa, arguing that vowel reduction is a CSP characteristic of fast speech and therefore separate from the phonological processes governing auxiliary contraction. This is in direct contrast with Kaisse (1983), who argues that weak forms of auxiliaries occur in relatively slow speech, unlike other forms of vowel reduction; Kaisse therefore concludes that vowel reduction is linked to grammatical information stored with the individual lexical representations of strong and weak forms, specifying the contexts in which different forms are permissible.

MacKenzie (2012) found that intermediate forms frequently occurred following subject noun phrases. There appeared to exist an effect of subject length, whereby full forms were more likely to occur after longer subject phrases, and contracted forms ceased to occur following a noun phrase of eight words or more. She also concluded that different auxiliaries have different syntactic distribution of full, intermediate and contracted forms. She found that whereas both intermediate and contracted forms of *has* occurred after noun phrases (for example *John has* [ʤɒnəz] and [ʤɒnz], contracted forms of *have* and *will* did not (for example *three have* [θɹiv] and *Sue will* [sul]. Instead, only full and intermediate forms of these verbs were found following noun phrases. She concludes from these findings that auxiliaries have alternations between two or more allomorphs for full and contracted forms, and that intermediate forms result from lower level phonological or phonetic processes acting on these allomorphs. Thus, she argues that the intermediate forms of *has* which occurred alongside contracted forms had underlying full forms, affected by further phonological processes. In contrast, she argues that the intermediate forms of *have* and *will* had underlying contracted forms, which were affected by lower level processes to produce the intermediate forms. In summary, she concludes that intermediate auxiliaries are a heterogeneous group; some are derived from full forms, while others are derived from contracted forms.

One problem with Mackenzie's (2012) criteria for categorisation of full, intermediate and contracted forms is that they do not include vowel reduction, a phenomenon shown by Simpson (1992) to occur regularly within the auxiliary system. Moreover, forms identified by Mackenzie as containing only a coda consonant are much more likely to occur for verbs which have accepted contracted forms, such as *is* (*'s*) and *have* (*'ve*). Such contracted forms do not exist for *can;* it is therefore less likely that this verb would be realised with the single coda consonant [n]. It is therefore concluded that while Mackenzie's recognition of a third intermediate category contributes to current knowledge of the phonology of auxiliaries, the system of categories needs to be modified and elaborated, in order to account for phonetic behaviours observable in a wider range of auxiliaries, including *can*.

2.3. The Phonological Development of Auxiliary Verbs

There appears to exist only one study which has explored the phonological development of auxiliaries in detail (Dye, 2011). Dye investigated the possible causes of phonetic and phonological reduction of auxiliaries in French. She analysed the phonetic production of auxiliaries in relevant constructions in 28 children aged from 1;11 to 2;11, the period during which verbs including auxiliaries emerge and become established. She noted a continuum of realisations, from full phonetic forms at one extreme, through to forms showing various phonetic reductions, through to complete omissions of the auxiliary at the other extreme (referred to as deletions). Deletion was said to occur when the child produced the correct syntactic structure for an auxiliary, but without the auxiliary, for example, subject pronoun followed by the infinitive form of the main verb.

Previous researchers have interpreted auxiliary omission in children's speech as evidence that syntactic representations for auxiliaries have not yet been acquired. In contrast, Dye (2011) proposes that children as young as age 1;11 have the representations, but do not realise auxiliary forms owing to restricted phonological processing. In support of her argument, she reports that this continuum of auxiliary development was found within individual children. Full forms were found in the speech of the youngest children, aged 1;11. All children produced at least one form of phonetic reduction and most children produced multiple forms of reduction. Some auxiliary deletion was also found to occur in the adult speech of the children's caregivers. Dye therefore concludes that auxiliary omission is not linked with syntactic knowledge, but results from pressures on production which are evident when the child speaks in multiword utterances. (Dye's usage of the term *production pressure* refers to motor constraints which limit the child's articulatory capacity.) She suggests that children's relatively limited phonological processing capacity and therefore increased production pressure explains why more auxiliary omissions are found in children's speech than in adults' speech.

One difficulty with Dye's conclusion is that it is not made clear whether individual children showed a continuum of production types for all auxiliaries, or whether omissions were restricted to only specific auxiliaries for individual children. The account of reduced phonological processing capacity and production pressure only holds, if it can be shown that individual children produced different forms of the same auxiliary in the same construction. Otherwise, it may be that children had acquired syntactic knowledge of some auxiliaries, or had acquired specific auxiliaries within the context of individual constructions, but not others. Alternatively, the findings may be explained in terms of frequency of specific constructions. It may be that the extent to which an auxiliary is reduced or omitted depends on the frequency of the construction in which it occurs. In this case, one might expect the full phonetic forms of auxiliaries to be found in more frequent constructions and omissions to occur in less frequent constructions. On the one hand, reduced forms might be expected to occur in less frequent constructions, owing to reduced phonological processing capacity in these situations. On the other hand however, they might equally be found in highly frequent constructions, owing to increased gestural overlap and the application of CSPs in utterances which the child has mastered. Similarly, omissions might be found to occur as an extreme form of reduction in highly familiar utterances. This would explain the finding that omissions were found even in caregivers' speech.

A further observation made by Dye (2011) is that children with a slower speech rate produced fewer auxiliary omissions. Again, she interprets these findings as evidence for the involvement of phonological factors in auxiliary omission. However, an alternative explanation may be that slower speech rate and fewer auxiliary omissions are both independent characteristics of children with a more analytic language learning style. It is therefore concluded here that Dye's findings cannot be interpreted as firm evidence that auxiliary omission results solely from phonological factors. It may additionally result from variables concerning specific lexical items and constructions. This issue is returned to later, in the section below on the acquisition of auxiliary syntax.

It is noteworthy that these two most recent studies of Mackenzie (2012) and Dye (2011) are the only two which have, firstly, based their conclusions on empirical findings from actual language data and, secondly, reported that phonetic forms of auxiliaries in adults' and children's speech can be placed on a continuum. Mackenzie (2012) views full, intermediate and contracted forms to be categorically different, whereas Dye (2011) reports a gradient of different realisations. This finding that strong and weak forms exist along a continuum was not detected by previous reports, which were purely theoretical. These different findings of categorical versus gradient distinctions between full, intermediate and weak forms may result from cross-linguistic differences between the two studies. It may be that the differences between these forms are more categorical in English, but more gradient in French.

2.4. Summary of Research on the Phonology of Auxiliary Verbs

In summary, there are several different theoretical viewpoints concerning the phonological and lexical relationships between strong and weak auxiliary forms. Some argue that strong and weak forms have different phonological representations linked to the same lexical item (Cruttenden, 2001). Others propose that weak forms are derived from strong forms via phonological and/or syntactic rule-governed processes (Baker, 1971; Kaisse, 1983; Ogden, 1999; Zwicky, 1970). However, many of these accounts are based purely on linguistic theory, and make no reference to empirical data. In contrast, the recent studies of Dye, (2011) and Mackenzie, (2012) are based on substantial corpora of child and adult language respectively. These researchers have shown that the relationships between strong and weak auxiliary forms are more complex than first thought. A continuum of phonetic forms appears to exist, with strong forms at one extreme, fully reduced weak forms at the other and a range of partially reduced forms in-between. The exact roles of syntactic processing, phonological representations and CSPs in determining which form occurs in any specific context continue to be debated.

2.5. The Developmental Trajectory of Auxiliary Emergence

There exist two large-scale longitudinal studies, which have focused on the developmental trajectories of auxiliary verbs in young children (Richards, 1990; Wells, 1979). Wells (1979) conducted a longitudinal study of 128 preschool children, in order to determine the age at which specific auxiliary forms were acquired. Ten recordings were made of each child at approximately three-monthly intervals from age 15 months until 42 months (three-and-a-half years). The age of acquisition of a specific form was defined as the age at which at least 50% of the sample had used the form at least once. The primary auxiliaries *be* and *have* were the earliest to be acquired and were mastered by 100% of the children by 42 months of age. In contrast, the majority of modal verbs had not been mastered by 50% of the sample by the end of the study. Wells also found individual differences in the rate of auxiliary acquisition. He operationalised mastery of auxiliary syntax as the point at which children had used at least five different auxiliary forms. This ranged substantially from 21 to 42 months of age. Wells also found that auxiliary acquisition was a gradual process. In most cases, there was a period of three months or more between the emergence of the child's first auxiliary and the point at which five different forms had been acquired. The earliest auxiliaries to be acquired were the main auxiliaries *be*, *do* and *have*, the modals *can* and *will* and the quasi-auxiliary *be going to*.

Richards' (1990) study included a specific investigation of the emergence of the modals *can* and *can't*, which form the primary grammatical focus of the current study. He takes a usage-based perspective on auxiliary verb development. He proposes that the emergence of auxiliaries involves the attainment of several perceptual and cognitive skills. These include the ability to perceive unstressed grammatical words, the ability to perceive syntactic regularities in others' usage of auxiliaries and finally, the ability to recognise auxiliaries as a grammatical class and use them appropriately. Richards argues that rate and style of auxiliary acquisition depend upon environmental factors, such as the input which the child receives. Environmental influences may be facilitative if they increase the salience of unstressed forms,

demonstrate the relationships between contracted and full forms, clarify boundaries between auxiliaries and other syntactic elements and clearly illustrate the relationships between the linguistic and situational contexts in which auxiliaries are used.

Richards (1990) also observed individual differences both in the rate of auxiliary verb development and qualitative patterns of emergence. He found individual differences in the extent to which auxiliaries emerged in unanalysed constructions, depending on the individual child's relative caution or impulsiveness in using these forms without fully understanding their internal structure. In a study of 33 children's auxiliary development, Richards classified the children according to a four-cell model which aimed to reflect their different learning styles. Apparent slow language learners with early emergence were those children who initially used many unanalysed or partially analysed forms. Apparent slow language learners with late emergence were those children whose input may have been less facilitative, or who were slower to learn grammatical rules. One child was an apparent fast language learner with early emergence. Richards acknowledged that this was a rare category and suggested that children falling into this category may be relatively impulsive in their language usage and prepared to take risks. Apparent fast language learners with late acquisition were more analytic, showing more rule-governed usage. Richards suggests that this latter group spent longer analysing their input and internalising grammatical rules prior to usage. This therefore constitutes a more cautious learning style.

Richards (1990) investigated specific qualitative trends in auxiliary development by testing the complexity principle. This principle simply states that learning should progress from utterances of lower syntactic complexity to those of higher syntactic complexity (Brown, 1973). Richards predicted that any violation of this principle evident in his participants' utterances would result from holistic learning, which would lend more support to usage-based approaches than nativist theories. Richards (1990) conducted a longitudinal study of seven children's auxiliary development, which started either in the second or third year of age and ended in the fourth year. He found that the affirmative forms of most auxiliaries were learned

before the negative forms, as predicted by the complexity principle. However, there were specific exceptions for which most children acquired negative forms first, including *can't*. Declarative constructions containing auxiliaries were learned prior to question forms with subject verb inversion, confirming a further prediction of the complexity principle. Constructions containing auxiliary plus main verb emerged earlier than ellipsis, also confirming the complexity principle. Combinations of inversion and ellipsis such as those occurring in tag questions were some of the latest constructions to emerge. However, contracted forms (such as *wasn't* and *didn't*) emerged prior to full forms (*was not* and *did not*), contrary to the predictions of the complexity principle.

In his investigation of the emergence of *can* and *can't*, Richards reports that *can* was the earliest modal verb to emerge for six out of seven children, and that it was the most frequent modal verb for all of the children. *Can't* emerged prior to *can* in declarative constructions, but *can* emerged prior to *can't* in question forms with subject verb inversion. The pattern of emergence in declarative constructions therefore contravenes the complexity principle and indicates holistic learning, whereas the pattern of emergence in interrogative forms confirms it. Richards explains his findings in terms of rote-learning of *can't* in declarative constructions, a finding also reported by Bloom (1970).

A further finding which can be explained by rote-learning of *can't* is that *can't* was initially found to occur much more frequently than *can* in the children's vocabularies, twice as frequently in some cases and four times more frequently in one child. This pattern occurred consistently until *can* began to feature regularly in the children's vocabularies, except for two instances in which the frequency pattern was temporarily reversed, with *can* occurring more frequently than *can't*. The usage of *can't* was syntactically more restricted than that of *can,* occurring in stereotyped utterances and co-occurring with only a limited range of main verbs including *do*, *get*, *put*, *find*, and *see*. Richards attributes this to the relative syntactic complexity involved in using the negative form *can't* in question forms. In contrast, *can* co-occurred with less evidence of rote-learned, stereotyped utterances and a much wider range of

constructions and main verbs. The most frequent constructions in which *can* occurred were declaratives and *yes*/*no* questions. For three out of the seven children, there were identifiable points at which the range of contexts for *can* increased. This was found to coincide with the emergence of *yes*/*no* questions in two of the children (Richards, 1990).

In a more recent study of auxiliary development, Lieven (2008) summarises the stages of auxiliary acquisition within the context of the utterance. She states that the earliest multi-word utterances contain no overt auxiliaries. She proposes that the earliest auxiliaries in children's speech are unanalysed and occur in rote-learned utterances, such as *don't want it* and *can't do it*. When children begin to produce auxiliaries, they then go through a long period in which they omit auxiliary forms which have evidently become part of their vocabulary. Over time, this omission reduces and provision (production) of auxiliaries increases (Lieven, 2008). This issue of omission is returned to in the following discussion of investigations of auxiliary acquisition from a usage-based perspective.

2.6. The Acquisition of Auxiliary Syntax: Research within a Usage-Based Constructivist Framework

Whereas the previous section on the phonological characteristics of auxiliaries mainly discussed research from the 1970s and 1980s, the literature discussed in this section is much more recent. This is because much of the research which has explored the grammatical development of auxiliaries in detail has been carried out by researchers who adopt the currently popular usage-based approach to language acquisition. These two contrasting sections therefore reflect a major paradigm shift in the field of language acquisition, from early generative theories which focus on grammatical and phonological rules, to more recent constructivist theories, which focus on usage and the child's capacity to learn language through their global cognitive abilities. The research covered in this section has many of the same recurring themes as that already discussed in Chapter One, as an introduction to the

usage-based approach. A combination of cross-sectional laboratory experiments and longitudinal studies of natural language have revealed that the acquisition of auxiliary verbs is initially lexically specific and restricted to only a limited range of different constructions. This restricted usage reduces over time, as children learn to apply verb forms to a wider range of constructions. Input frequency has been established as a factor which influences age of acquisition. The earliest forms to be acquired are those which occur most frequently in the speech of the children's caregivers, with only a few exceptions.

Children begin to produce auxiliary verbs at two years of age, although they do not achieve adult-like competence until during their fourth year (Theakston, Lieven, Pine and Rowland, 2005). The most complex and therefore latest features of the auxiliary system to be acquired are the modals and the usage of auxiliaries in *wh* questions (Lieven, 2008). Auxiliaries have many grammatical functions, including the expression of tense, agreement, negation, modality and their role in questions. Mastery of the auxiliary system is therefore considered to reflect maturity in a child's grammatical development (Lieven, 2008). The main auxiliaries in English are *be*, *have* and *do*, while modal auxiliaries include *can*, *will* and *might* (Lieven, 2008).

Theakston et al. (2005) and Lieven (2008) contrast the generativist and constructivist approaches to auxiliary acquisition. According to generativist approaches, children have the innate abstract linguistic knowledge necessary to work out how their native language expresses different functions, including tense, agreement and negation. Auxiliaries are viewed as grammatical units within a syntactic category which constitutes part of the child's innate linguistic knowledge. It is argued that auxiliaries cannot possibly be acquired through learning, because they do not make reference to specific objects or concepts and therefore lack semantic content (Chomsky, 1965). Lieven (2008) argues that a problem with such approaches is that they do not specify the exact means by which the child's innate dispositions, performance constraints and semantic bootstrapping work together in order to resolve the child's difficulties and error patterns.

In contrast, usage-based constructivist approaches take the view that the child acquires linguistic knowledge through their ability to form abstractions from the language which they hear and use (Lieven, 2008). Grammatical units (including auxiliaries) are learned in the context of individual constructions with a specific usage function. Children learn and store an increasing inventory of lexically-based constructions and sub-constructions over time from the linguistic input provided by their caregivers. This enables them to develop an understanding of the linguistic relationships between similar constructions and, subsequently, to form abstract schemas for specific linguistic rules which underlie adult language (Lieven, 2008; Theakston et al., 2005; Rowland and Theakston, 2009). One example of such a rule is subject-auxiliary inversion.

High token frequency of specific words and phrases in the input strengthens (entrenches) the child's schemas for comprehension and usage of these words and phrases. High type frequency of different words in similar construction contexts in the input enables the child to make generalisations concerning the slots within constructions. The child thus learns that items can be substituted at various positions within an utterance, for example, *where's x gone?*. The child eventually acquires frames for longer and more complex constructions, which are more abstract and which have wider scope in terms of the slots which they contain and the items which can fill these slots (fillers) (Lieven, 2008).

This approach therefore emphasises the role of frequency and the distribution of exemplars in the input in the acquisition of auxiliaries (Theakston et al., 2005). According to this view, constructions which are highly frequent in the child's input are acquired early, stored as whole constructions, have strong representations and should be used frequently by the child. In contrast, those constructions which are less frequent in the input, or which have more variable constituent units should be acquired later, have weaker representations and be used less consistently (Theakston et al., 2005). Children are believed to have only partial representations in instances when they can produce specific forms, but not other related forms in the adult grammar (Lieven, 2008). If children have not yet acquired a specific auxiliary

through encounters with relevant constructions, or have only a partially specified linguistic representation of the auxiliary, they may then use a construction containing another auxiliary which is more familiar to them. Alternatively, they may combine two constructions learned from their input which do not contain an auxiliary, giving rise to auxiliary omission (Theakston et al., 2005). The issue of omission is returned to later in this discussion. While constructivists argue that the child's input plays a crucial role in language acquisition, they acknowledge its interaction with other factors, including the child's existing linguistic system and their communicative interests (Lieven, 2008).

Several studies investigating the acquisition of auxiliary verbs have been carried out by Anna Theakston, Elena Lieven and colleagues at the University of Manchester and the Max Planck Institute for Evolutionary Anthropology, Leipzig. This research has yielded results which indicate that the acquisition of auxiliary verbs appears to be lexically-specific and initially limited to specific constructions. These findings therefore support a constructivist, rather than a generative framework. Some studies have involved elicitation tasks in laboratory settings, while others have involved analyses of longitudinal corpora of naturalistic data, consisting of language occurring in spontaneous interactions between the children and their mothers. More recent studies have included a combination of cross-sectional and longitudinal methods.

The elicitation studies were conducted in a laboratory setting. An experimenter attempted to elicit the target auxiliary form from the child, either in declarative or interrogative constructions, using a range of toys and activities. For example, a declarative condition would involve the child answering the researcher's question: "what is happening?". An interrogative condition would involve the child asking questions in response to prompts from the researcher. For example, the child would ask a toy frog "is the fox cooking?" (Theakston and Lieven, 2005).

In contrast, the longitudinal studies involved the analyses of large audio and video corpora of naturalistic mother-child interaction data. These include twelve children from the Manchester

Corpus (Theakston, Lieven, Pine and Rowland, 2001) and the Thomas Corpus from the dense database (DDB) (Lieven et al., 2009). These studies have the advantage of being able to investigate linguistic phenomena as they emerge naturally in spontaneous language and are not subject to the confounds which can occur in experimental data.

Most studies have focused on the auxiliaries *be* and *have*. The rationale for exploring these verbs is that they have different forms to mark both tense and agreement in English (Theakston and Lieven, 2005). However, more recent studies have also focussed on the acquisition of modals, including *can*.

Theakston and Lieven (2005) examined the usage and error patterns of children's auxiliaries in declarative and interrogative constructions (*yes*/*no* questions). The study focused on singular and plural forms of *be* (*is* and *are)* and *have* (*has* and *have*). The children in the declarative condition were aged two to three years, whereas only three-year-olds participated in the more advanced interrogative condition.

The researchers found significant differences in levels of correct usage as a function of construction type, auxiliary verb and verb form. More errors occurred in the interrogative than declarative condition. The children showed significantly fewer correct usages of *have* than *be.* More errors were found in the usage of the plural form *have* than the singular form *has*, especially in the interrogative condition. Differences in specific error patterns were also found as a function of construction type and auxiliary. Agreement errors were more characteristic of interrogatives, whereas auxiliary substitution errors were more characteristic of declaratives. Utterances with *have* contained more agreement errors and auxiliary substitutions than utterances with *be*.

Theakston and Lieven (2005) conclude that the children were more familiar with the auxiliary *be* than *have*. Whereas the older children had acquired interrogative syntax with the more familiar verb *be*, they had not yet generalised this knowledge to interrogative constructions

with the less familiar verb *have*. Theakston and Lieven interpret their findings as evidence for the lexically-specific nature of auxiliary acquisition, as proposed in the usage-based approach. They argue that their results do not support a generative approach, which would predict that children's innate understanding of the relationships between auxiliaries would enable them to immediately transfer their knowledge to all auxiliary forms (Chomsky, 1965).

A further study investigated the error patterns occurring in two-year-olds' *wh* questions (Rowland, Pine, Lieven and Theakston, 2005). This involved a longitudinal analysis of naturalistic data for twelve children and a diary study of a further child. Similarly to Theakston and Lieven (2005), they found more errors for the plural forms *are* and *have* than for the singular forms *is* and *has*. Overall error rates were higher for *do* and modal verbs than *have* for the child in the diary study. *Do* and modal verbs occurred with significantly more inversion errors than either *be* or *have*. Omission was the most frequent error type and the plural *have* was the most frequently omitted form. Rowland et al. (2005) concluded that omission errors did not result from the children's lack of knowledge of specific forms, since they produced the omitted forms appropriately elsewhere.

In a further elicitation study involving children aged three and four years, Ambridge et al. (2006A) investigated inversion error rates in a range of *wh* words, auxiliaries and grammatical forms. They found significantly more inversion errors in questions with the plural form *do* than the singular form *does*. However, this finding was only applicable in questions with *what* and *who*, but not those with *how* and *why*. Significantly more double marking errors were found in questions with *do* than in those with *be* or *can*.

The findings of these two studies provide further support for the lexically-specific nature of auxiliaries. They also indicate that auxiliary usage may initially be limited to specific constructions (Ambridge, Rowland, Theakston and Lieven, 2006A; Rowland et al., 2005). Rowland et al. suggest that children may initially develop partially analysed frames for *wh*

questions, consisting of a *wh* word and an auxiliary. The child must then analyse these frames further in order to generalise the usage of auxiliaries to the full range of wh questions.

Following the findings of these studies, Theakston et al. (2005) investigated patterns of auxiliary usage (provision) and omission in a longitudinal study of eleven children aged two to three years. They predicted that there would be higher rates of provision for forms heard frequently in the input, and those which occur in constructions with more fixed slots for lexical items. In contrast, they predicted lower provision rates for forms heard less frequently and for those which occur in more variable constructions and cannot therefore be used prior to the development of abstract schemas for usage.

Overall, the children omitted auxiliaries in 65% of constructions which would contain an auxiliary in the adult grammar. Moreover, the average provision rate was only 60% at the end of the study. It was evident that children were continuing to omit specific auxiliaries which had emerged in their language. Provision rates of approximately 50% were even found for the most frequent and earliest verb forms to emerge: *am, is, has* and *have*. The authors conclude from this finding that auxiliary omission does not result from a lack of lexical and grammatical knowledge of the target auxiliaries.

There were significant differences in provision rates as a function of verb form, with more provision of *is* than *am* and more provision of *has* than *have*. There were also differences in levels of provision as a function of lexical subject, with more provision of *he's* and *it's* than *I'm,* and more provision of *he's* and *it's* than *I've* and *we've*. In addition, there were higher levels of provision for constructions with fixed subjects, such as *he's* and *it's* than for constructions with more variable subjects, such as proper noun plus contracted auxiliary *'s*.

There were significant correlations between the maternal input frequencies of specific forms and the ages at which the children acquired these forms. The exceptions were *you're* and *you've*, which were highly frequent in the mothers' speech, but some of the latest forms to

emerge in the children's speech. The authors argue that additional cognitive development is required for these forms to be used appropriately, namely reversing the usage context from that heard in the input. The authors predicted that input frequency and age of acquisition would be positively correlated with provision rates. While this was largely the case, there were specific examples for which this prediction did not hold. The forms *I'm* and *I've* had high input frequency and were acquired early, but had relatively low provision. The authors could not immediately account for these results, and concluded that the usage-based approach needs to develop further in order to explain such anomalous findings.

Once again, this study provides evidence for the lexically-specific and construction-specific nature of auxiliary acquisition, with a particular emphasis on provision rates. There is also support for the authors' prediction that forms occurring in fixed constructions would be easier to learn than those occurring in more variable constructions, for which the child must develop a more abstract schema for usage in different contexts.

It is interesting to compare the findings of this study of auxiliary omission with that of Dye (2011), reported above in the section on phonological characteristics of auxiliaries. Dye (2011) and Theakston et al. (2005) agree that omission does not indicate a lack of grammatical and lexical knowledge, on the grounds that full forms and omissions can be observed concurrently within individual children's language. Theakston et al. provide further clarification that this is also observable for individual verb forms. However, the authors differ greatly in their explanations of auxiliary omission. In the absence of evidence that children lacked lexical representations of the verbs which they were omitting, Dye assumed the omissions to result from restrictions on articulatory capacity. However, more careful scrutiny of omission patterns by Theakston et al. revealed that omission was more likely to occur for some lexical items than others and in constructions with more variable slots than those with more fixed constituents. These findings show that it is necessary to consider the child's knowledge of specific verbs and constructions, rather than to assume that emergence of a verb form in a child's speech constitutes complete knowledge of its usage (Theakston et al., 2005).

Theakston and Lieven (2008) investigated the impact of input and prior discourse context on provision and omission rates of singular and plural forms of the auxiliary *be*. This research involved both an elicitation study of 96 children aged two to ten and a longitudinal study of Brian from age 2;8 to 3;2. In the elicitation study, the children were primed either with declarative statements or *yes*/*no* questions, either with the auxiliary form *am* or *are*, prior to elicitation of the target auxiliaries in declaratives. The children showed significantly higher provision of *am* than *are,* as predicted. The children primed with declaratives showed higher auxiliary provision rates than children primed with questions. Provision rates were higher for utterances containing familiar main verbs (real words) than those containing novel main verbs (nonsense words).

In the longitudinal study, those utterances in Brian's speech which either contained or required an auxiliary in the adult grammar were identified. The five preceding lines of transcript were then scanned for utterances which similarly either contained or required an auxiliary, occurring either in Brian's or his mother's speech. Auxiliary provision rates were found to increase over time. Brian was significantly more likely to omit auxiliaries following his mother's questions than following her declaratives, or if there was no prior context in the input. There were no significant differences in provision rates following a maternal utterance with the target auxiliary form of *be*, compared with a maternal utterance containing a different form of *be*.

Provision was highest in contexts in which Brian had previously produced an auxiliary, reduced in contexts in which there was no prior auxiliary, and lowest in contexts in which he had previously omitted an auxiliary. This was significant for *are*, but not for *is*. Provision was higher when Brian had previously produced the target form of *be*, compared with another form of *be* or no prior context. Provision was reduced following a prior context in which Brian produced a different form of *be*, compared with provision rates in utterances with no prior auxiliary context.

Prior auxiliary provision rates in Brian's speech had an effect on the target utterance which was independent of input context. Maternal input only affected provision in instances when Brian produced no prior auxiliary context or omitted the auxiliary in the prior context. Provision rose to 61.7% in contexts in which Brian had produced an utterance with the same lexical subject, and 70% in instances in which Brian had previously produced an utterance with both the same lexical subject and auxiliary verb. In contrast, only 38.1% of utterances in the prior maternal input contained the same lexical subject as the focal maternal utterance under investigation.

Theakston and Lieven conclude from their findings that auxiliary provision initially relies on the child's knowledge of specific subject-auxiliary combinations. Furthermore, they suggest that children's knowledge of auxiliaries may be tied to specific lexical items and constructions, as shown by the higher provision rates for familiar than novel verbs in the elicitation study. In both studies, questions in the input appeared to have an inhibitory effect on auxiliary provision, while declaratives in the input appeared to have a facilitatory effect. The longitudinal study showed that provision was maximised by Brian's use of the same auxiliary and lexical subject in a preceding utterance. The findings support the view that children are able to acquire abstract knowledge of linguistic rules from the lexically-based exemplars to which they are exposed. However, Theakston and Lieven acknowledge that their findings cannot be explained solely in terms of the impact of maternal input. It is evident from their findings that the child's own usage in prior contexts also influences provision rates. The authors therefore suggest that further research should focus on the possible interactions between the role of input and the child's existing linguistic knowledge in the acquisition process.

The studies discussed above have shown that children's knowledge of auxiliaries is initially lexically-specific. Lieven (2008) attempted to track the process by which children's auxiliary knowledge develops from a stage of lexical specificity, to partial productivity and finally to full abstraction of schemas for auxiliaries. Lieven analysed the emergence of auxiliaries in

novel constructions in six children from age 2;0 to 3;0. She found both commonalities and individual differences in the children's patterns of auxiliary acquisition. All children began producing frames for auxiliaries within the first four months of their third year. Three children added frames to their inventories at a steady rate throughout the study period, while the other three showed periods of sudden increased usage in different frames either at age 2;6 or 2;8. There also existed individual differences in the order in which specific auxiliaries were acquired. Auxiliaries emerged in frames with slots for main verbs and lexical subjects. Contracted forms such as *it's* were some of the first frames to emerge; in these instances, the subject and verb did not have separate slots in the frame, but were learned as a single lexical item. A single frame could be highly productive if it occurred with a slot which had many possible auxiliary fillers; Lieven concluded that this did not necessarily constitute full abstraction of the auxiliary.

Increasing abstraction of schemas was evident over time, as the number and scope of slots within frames increased, for example to include a wider range of subjects, verbs and other syntactic elements. Lieven interprets the substitution of an increasing range of auxiliaries within appropriate slots as evidence that the child was learning the connections between syntactically related forms. This process led to the emergence of novel construction types and grammatical forms, including tag questions, *wh* questions and morphological markers of tense and agreement. When children's usage of auxiliaries was initially tied to individual constructions, error rates were low. However, error rates increased during the gradual process of generalisation.

A positive correlation was found between the order of frame emergence in the children's speech and the usage frequency of frames in maternal speech. The mothers produced many *wh* questions with the auxiliary *do*, which the children learned to produce over time. The mothers also produced many questions with the conditional forms *could*, *should* and *would*, compared with the children's greater tendency to produce forms such as *can* and *can't*. The mothers also produced more instances of the subjects *you* and *we* in auxiliary frames

compared with more instances of *I* in the children's speech. This may be a reflection of the mothers' communicative interests, which would have differed to some extent from those of the children (Lieven, 2008).

Similarly to the omission study of Theakston et al. (2005), this study has the advantage that the results are based entirely on naturalistic, longitudinal data. However, it appears to have several methodological difficulties. Lieven defines the level of abstraction according to the number of main verbs with which an auxiliary was found to occur within a specific construction type. Thus, one main verb equals lexically specific, two verbs equals partially productive and three main verbs equals full abstraction of schemas. These numbers seem somewhat arbitrary, as they attempt to quantify and categorise a process which appears in reality to be gradual and continuous. A further disadvantage is that the maternal speech used in the correlational analysis came from a different data sample from the children's speech. The results of this part of the analysis are therefore only valid if one can assume that there exist general commonalities in the construction types occurring in maternal input.

The most recent two studies conducted by these researchers on auxiliary development aimed to investigate the extent to which children aged two and three years are able to generalise their knowledge of auxiliary usage across different forms and different construction types: positive and negative declaratives and questions (Rowland and Theakston, 2009; Theakston and Rowland, 2009). These studies took the form of a series of elicitation studies carried out longitudinally at six-weekly intervals, with twelve children aged from 2;10 to 3;6. Theakston and Rowland (2009) investigated the usage and error patterns of *be*, while Rowland and Theakston (2009) investigated *does* and the modals *can* and *will*.

Overall numbers of responses and levels of correctness were higher for declaratives than questions. As found in their previous studies, the form *is* occurred with significantly higher levels of correctness than *are*. There were similar levels of correctness for *is* across declarative and interrogative constructions, whereas *are* showed more errors in questions than

declaratives. Correct responses increased in declaratives as a function of age, but this was not found in interrogatives. The authors interpret these findings as evidence that the children understood the relationships between construction types better for the form *is* than *are*.

Theakston and Rowland also found interactions between age, auxiliary and construction type in the occurrence of error patterns. A significantly higher error rate for *are* than *is* was only found in *wh* questions at age 2;11 and only in declaratives at age 3;2. Omission errors occurred more often in declaratives, whereas more agreement errors were found in questions. Agreement errors were more likely to occur in questions with *is* than those with *are*. This pattern was not observed for declaratives.

Rowland and Theakston found substantially more correct responses for all three auxiliaries *can*, *will* and *does* in declaratives than in questions. *Can* was used more correctly than *does* in both positive and negative declaratives, although this only reached significance at ages 2;11 and 3;5. Similarly, children tended to perform better for *can* than *will,* although this only reached significance at age 3;5. Performance was significantly better for *can* than either *will* or *does* in positive questions, although similar correctness levels were found across the three auxiliary forms in negative questions. These findings show that at specific points in time, the children's knowledge of *can* usage in positive questions did not generalise to their usage of *will* and *does*. The usage-based approach accounts well for this finding. However, a finding which indicates a degree of generalisation is that there were no significant differences in overall levels of correctness between positive and negative declaratives, although positive question forms initially showed higher correctness levels than negative questions. It is evident that the children understood the relationship between positive and negative forms in declaratives and were able to generalise their knowledge of auxiliary usage to some extent. However, they could not initially make these generalisations in question forms, perhaps owing to the increased complexity of these forms (Rowland and Theakston, 2009). The usage-based approach cannot currently account for these specific early similarities in usage across positive and negative declaratives (Rowland and Theakston, 2009). The authors suggest that some

patterns of linguistic abstraction may occur earlier than others, and that the usage-based approach needs to develop predictions regarding the order in which different patterns of abstraction occur.

In order to investigate the role of maternal input on the children's usage of *can*, *will* and *does*, the results of this study were compared with maternal input frequencies of these forms, using samples of spontaneous conversation between the children and their mothers. The mothers produced significantly higher proportions of *can* than either *will* or *does*. This explains the children's earlier acquisition and generalisation of *can*. However, some findings could not be explained in terms of input frequency. The children produced similar numbers of correct responses for positive and negative declaratives, although there were much higher levels of positive declaratives in the input; this would predict more correct responses for positive declaratives. The children produced more correct questions with *will* than *does*, although their mothers more frequently used questions with *does* than *will*.

2.7. Summary of Research on the Acquisition of Auxiliary Syntax

Elena Lieven, Anna Theakston and colleagues have conducted a series of studies which have thoroughly explored children's usage and error patterns for auxiliaries. While the usage-based constructivist approach accounts for many of the patterns which have been found, there also exist some patterns which this approach cannot currently account for.

As predicted by the usage-based approach, children aged two and three years appear to have lexically-specific knowledge of auxiliaries. Initially, this knowledge also appears to be restricted to specific construction types. Evidence for these conclusions comes from higher levels of correct usage of specific forms and construction types than others. *Be* was more often used correctly than *have* and singular forms were more often used correctly than plural forms.

Different types of errors were also found to co-occur with specific auxiliary forms and construction types. Auxiliary substitution errors were more characteristic of declarative constructions, while agreement errors occurred more often in interrogatives. Inversion errors occurred more often in questions with *do* and modal verbs than those with *be* or *have*. Questions with the *wh* words *what* and *who* appeared to occur with more omission errors than those with *why* or *how*. Provision initially appears to rely on the child's knowledge of specific subject-auxiliary combinations. Provision rates were highest with the forms *is*, *has*, *it's* and *he's* and lowest for the forms *am*, *have*, *I'm, I've* and *we've.* In addition, there appear to be complex interactions between age, auxiliary form and construction type which determine correctness levels and error patterns in usage.

Much of this research has shown that children's knowledge of the usage of one auxiliary does not immediately generalise to other auxiliaries in their vocabulary. However, there appear to be some exceptions in which children appear to understand syntactic relationships and generalise auxiliary knowledge from an early age. Evidence for this comes from the finding that children produced similar numbers of correct responses across positive and negative constructions. Rowland and Theakston (2009) conclude that the usage-based approach needs to develop further in order to predict the order in which specific linguistic abstraction patterns occur.

Familiarity and experience with specific auxiliaries and constructions also appear to play a role in children's usage. Better performance for some auxiliary forms than others are explained in terms of differences in familiarity (Theakston and Lieven, 2005). Similarly, provision rates were higher in constructions with more fixed slots for specific subjects and auxiliaries, than those with more variable slots for different items, which were presumably less familiar to the children.

Maternal input frequency appears to play a role in children's acquisition and usage of auxiliaries. However, this is an area of research in which some findings remain unaccounted

for. There were positive correlations between maternal input frequency and age of acquisition for most auxiliary forms, with the exception of the highly frequent forms *you're* and *you've*; the acquisition of these forms requires additional cognitive development on the child's part (Theakston et al., 2005). Provision rates were also unexpectedly low for *I'm* and *I've*, which were highly frequent in the maternal input. Children's especially high correct usage of *can* compared with other verbs was correlated with the significantly higher frequency of *can* in the input (Rowland and Theakston, 2009). However, the mothers used significantly more positive than negative declaratives and more questions with *does* than *will*. These findings were not reflected in the children's relative proportions of correct usage.

Auxiliaries in declarative constructions in the input appeared to facilitate auxiliary provision in children's utterances, whereas auxiliaries in question forms in the input appeared to inhibit provision in the children's utterances more than if there was no prior auxiliary context. The analysis of discourse context in one child's speech showed that provision in a target utterance was more greatly facilitated by provision in the child's own previous utterances than by provision in the maternal input (Theakston and Lieven, 2008). The researchers conclude from these findings that further developments of the usage-based approach need to focus on possible interactions between input and the child's existing linguistic knowledge (Theakston and Lieven, 2008).

In conclusion, the usage-based constructivist approach accounts for children's lexically-specific and construction-specific usage of auxiliaries. There exist complex interactions between auxiliary verb, verb form and construction type which determine usage and error patterns over time. Familiarity, experience and maternal input frequency are also important factors affecting children's auxiliary acquisition and usage. Factors which remain to be explained are interactions between the child's existing linguistic knowledge and the input which they receive, and the child's earlier abstraction of some linguistic patterns than others.

2.8. Summary and Conclusions

This chapter has reviewed research on the phonology of auxiliary verbs from a generativist perspective and the syntactic acquisition of auxiliary verbs from a usage-based perspective. These two schools of research are strikingly different in terms of both the areas investigated and the methodologies employed. Generativist theorists have not yet explored in detail the developmental trajectory of auxiliary acquisition. Conversely, usage-based research on auxiliary syntax has not yet investigated the acquisition of strong and weak forms. This latter issue is directly addressed in the current study for the verb forms *can* and *can't*.

Chapter Three

The Development of Phonology

3.1. Introduction

This chapter provides an overview of phonological development. Firstly, phonology is defined and typical patterns of phonological development are summarised. This is followed by a discussion of phonological versus phonetic theories of phonological development. *Natural Phonology* is exemplified as a more phonological approach and *Articulatory Phonology* is exemplified as a phonetic approach. These approaches have contributed in different ways to current linguistic understanding of phonological development and impairment. However, a major, recurring short-coming of phonological approaches (such as *Natural Phonology* is that they focus on the single word and therefore do not provide accounts of phenomena specific to connected speech. This section is followed by an outline of research on early multi-word utterances in the eighties, which highlighted the need to further investigate the phonology of multi-word speech. This leads to the next, most detailed section of this chapter, which discusses the nature of connected speech processes (CSPs) and recent research into their developmental patterns. This section sets the scene for the aims and methodology of the current study. Because usage-based approaches are central to the current study, and because these approaches greatly emphasise the role of input in learning, it is also necessary to include a section on the phonological characteristics of child-directed speech. This discussion sets the scene for Chapter Eight of this thesis, which focuses on the role of maternal input on assimilation development.

3.2. What is Phonology?

Phonology concerns the organisation and function of speech sounds (phones) within language and is therefore described as the interface between phonetics and linguistics. Every language comprises a phonological system, in which phones are linked by paradigmatic relationships of contrastive meaning and syntagmatic relationships, which govern the combination of phones within syllables according to phonotactic constraints. During early speech development, the child's phonological system undergoes constant change, resulting in different and more variable usage and combination of contrastive phones compared with that of adults. As the child's phonological system matures and stabilises, productions become more consistent and more similar to adult forms (Grunwell, 1987).

3.3. Typical Phonological Development

3.3.1. The Acquisition of Segments and Syllable Shapes

Because the current study focuses on the period from two to four years of age, this summary of phonological development begins with the emergence of first words. Prelinguistic development is not relevant to the current study and is therefore not included. Children produce their first words at approximately twelve months of age, after which, words and babble co-occur for several months (Stöl-Gammon and Sosa, 2007). Children's first words usually contain the same phones as those occurring in their concurrent babble. English-speaking children's first words typically comprise plosives, bilabial and alveolar nasals, glides and a range of vowels (Dodd, Holm, Hua and Crosbie, 2003; Grunwell, 1987; Stöl-Gammon and Sosa, 2007). Syllable structures typically found in first words are the monosyllabic structures CV and CVC, and the disyllabic structure CVCV (where C equals consonant and V equals vowel) (Stöl-Gammon and Sosa, 2007). Individual children's patterns of phones, syllable structures and lengths of vocalisation are often extended from their babble to their first words (Stöl-Gammon and Sosa, 2007; Vihman, 1996). During the period from the

emergence of first words until the child has acquired a vocabulary of 50 words, a limited repertoire of sounds and syllable shapes is established in the child's phonological system, known as a phonetic inventory (Stöl-Gammon and Sosa, 2007). The latest consonants to be acquired are those which either occur relatively infrequently in the language (such as /v/ in English), or those which require a high degree of articulatory precision (including fricatives, affricates and liquids in English) (Grunwell, 1987; Stöl-Gammon and Sosa, 2007).

From the age of 18 months to two years, the child's vocabulary increases rapidly and they produce their first multi-word utterances (Bloom, 1970; Stöl-Gammon and Sosa, 2007). Parallel phonological advances are evident, as the child acquires a much wider range of phones (including fricatives) and syllable shapes (including disyllabic words (Stöl-Gammon and Sosa, 2007). It is estimated that by two years of age, a child learning English has acquired a productive vocabulary of 250 to 350 words (Stöl-Gammon and Sosa, 2007). The speech of two-year-olds is characterised by a number of phonological phenomena which affect the overall structures of words, such as the omission of weak syllables. Phenomena which affect the realisation of individual phones, such as alveolar realisation of target velar plosives (fronting), may persist into the child's fourth year. The majority of these immature phonological realisations are eliminated by the age of four years, although the child may still be mastering the most challenging consonants. The latest consonants to be acquired by English-speaking children are dental fricatives, affricates and /r/, as well as consonant clusters (Grunwell, 1987).

3.3.2. The Acquisition of Stress

In addition to acquiring the phones and syllable shapes of their native language, children must also learn the stress patterns of words and utterances. The stress patterns of children's early words differ according to the language to which they are exposed. Some English-speaking children show a tendency towards producing words with a trochaic stress pattern, with primary

stress placed on the first syllable (Allen and Hawkins, 1980; de Boysson Bardies, Vihman, Roug-Hellichius et al., 1992). However, this Trochaic bias does not appear to be universal among English-speaking children, with evidence of considerable individual differences (Kehoe and Stöl-Gammon, 1997; Vihman, 1998). In contrast, the first words of French-Speaking children often have an Iambic stress pattern, with primary stress placed on the second syllable (de Boysson-Bardies et al., 1992; Stöl-Gammon and Sosa, 2007). The remainder of this discussion concerns the development of stress in English-speaking children unless otherwise stated.

Research has shown that although children aged two and three years are able to produce stressed syllables, they are slower to acquire the shorter, less salient weak syllables present in adult language. Allen and Hawkins (1980) found individual differences in rates of weak syllable production among two- and three-year-olds, ranging from 35% to 65%. The children tended either to delete weak syllables or realise them as stressed. Allen and Hawkins (1980) suggest that children do not produce weak syllables, either because they cannot perceive them, or because of articulatory constraints. Two-year-olds are more likely to produce word-final weak syllables than those occurring initially or medially (Kehoe and Stöl-Gammon, 1997). They are also more likely to delete weak syllables with sonorant onsets (for example, in *animal* or *telephone*) than those with obstruent onsets (for example, in *crocodile* or *octopus*) (Kehoe and Stöl-Gammon, 1997).

Kehoe (1998) identified three stages of stress acquisition, which were typical of children aged 22, 28 and 34 months respectively. The trochaic stage is characterised by the predominance of monosyllables, trochaic patterns and truncations (reduced word length in number of syllables). The experimental stage is characterised by stress errors and equal stress placement on all syllables within a word. The consistent stress pattern stage is characterised by more mature stress patterns and fewer errors.

There exists evidence that children initially learn the stress patterns of multi-word utterances holistically as unanalysed strings, then experiment analytically with the stress patterns of individual words, before reverting back to more mature stress patterns. Behrens and Gut (2005) studied the emergence patterns of two-word utterances in German-speaking children aged from 2;0 to 2;3. In utterances consisting of a noun plus a particle, the children aged 2;0 produced apparently mature stress patterns: stressed noun plus unstressed particle. At age 2;1, the children placed equal stress on both words, also showing a wider range of pitch movements. This appears to be akin with Kehoe's (1998) experimental stage. By age 2;2, the children had reverted to the original mature stress pattern.

A similar pattern was observed for utterances consisting of a determiner plus a noun. At age 2;0, the children produced mature stress patterns: unstressed determiner plus stressed noun. However, there were also characteristics which differed from the mature pattern, including segmental reduction of determiners and level intonation on nouns. At age 2;1, the children placed stress on the determiner and the noun was not always stressed. By age 2;3, the children produced fully mature stress patterns, with no segmental reduction of the determiner and falling intonation on the noun. The authors conclude that these patterns show evidence of gradual reorganisation from a holistic, unanalysed structure containing unstressed filler syllables, to prosodic integration of the individual grammatical constituents.

3.4. Phonological Versus Phonetic Approaches to Speech Development

3.4.1. Phonological Approaches

Phonological approaches to the acquisition of speech focus on the child's ability to produce the phones of the adult language within the domain of the single word, as well as the ability to produce contrastivity between phones (Howard, 2010). The child's realisation is compared with the mature form (known as the target). Errors are viewed as mismatches between the

child's realisation and the target, resulting from sound substitutions, distortions, omissions and additions (Howard, 2010). Analysis of speech data aims to identify the patterns occurring in the child's speech and the results may arguably become highly abstracted and removed from the actual speech data (Howard, 2010). The mental, phonological representation of a phone is viewed as an abstract entity, quite separate from the articulatory organisation of the phone in speech production (Kent, 1997). Phonological approaches to speech development include *Distinctive feature theories* (which later evolved into *Generative Phonology*) (Chomsky and Halle, 1968), *Natural Phonology* (Stampe, 1979) and *Nonlinear Phonology* (Bernhardt and Stöl-Gammon, 1994). Detailed discussion of these theories is beyond the scope of this thesis, but *Natural Phonology* is outlined below as a popular example

Natural Phonology was proposed by Stampe (1979) and has been developed for usage in developmental research and clinical contexts (Grunwell, 1987; Ingram, 1976). According to this theory, the acquisition of adult pronunciation is governed by innate phonological processes; these processes produce sound substitution errors which are less challenging for the child's developing phonological system and speech capacity than the target phones (Stampe, 1979). Over the course of development, these processes are gradually suppressed and the child's phonological system is thus revised to become more complex until it resembles that of adults who speak the same language. Grunwell (1987) distinguishes between structural simplification processes, which affect syllable and word structure, and systemic simplification processes which, affect the system of contrastive phones. Structural simplifications include weak syllable deletion, final consonant deletion, reduplication, consonant harmony and cluster reduction. Systemic simplifications emerge when structural simplifications have been suppressed and include stopping, gliding, fronting and context-sensitive voicing (Grunwell, 1987). This approach remains popular in current developmental research (for example, Dodd et al., 2003).

3.4.2. Articulatory Phonology

Articulatory (or *Gestural*) *Phonology* does not draw a distinction between phonetics and phonology. Underlying phonological structures are viewed as inextricably linked to the temporal and spatial properties of the phones produced in speech (Howard, 2010). According to this approach, the phonological representation of a phone comprises the specifications for the articulatory movements needed to produce the phone. This abstract representation of articulatory movement is known as a gesture (Browman and Goldstein, 1987; Kent, 1997; Van Lieshout and Goldstein, 2008). Examples of movements which may be specified in a gesture include those of the jaw, lips, tongue and velum (Browman and Goldstein, 1987). The timing and phasing of these movements is also specified in the gesture.

The temporal and spatial coordination of different gestures for the sounds which comprise a word or utterance results in a gestural score for the word or sentence (Browman and Goldstein, 1987; Kent, 1997). Temporal overlap of gestures within the gestural score is possible; this accounts for phonological phenomena not explained by other theories, such as assimilation (both within and between words), gemination, consonant elision and vowel reduction (Browman and Goldstein, 1987; Kent, 1997). This theory therefore moves away from the concepts of linearity between phones in speech production.

The organisation of a person's phonological system is partly determined by articulatory constraints on speech production and perception (Kent, 1997). This may explain some of the patterns specific to young children's phonology. The phonological errors encountered in children's speech are explained in terms of the temporal mis-alignment of gestures within the gestural score; this sometimes appears to give rise to phones which are not present in the adult production of the target word (Studdert-Kennedy and Goodell, 1992).

3.5. Usage-Based Phonology

Usage-based models link phonological development with development at other linguistic levels. They emphasise the roles of both input and usage in shaping and modifying the child's sound system (Bybee, 2006; Stöl-Gammon and Sosa, 2007). According to this approach, phonology is not learned independently of other linguistic levels, but is acquired in conjunction with grammar and lexis in the context of usage. Children's initial patterns of lexical selection and avoidance indicate that their choice of words is restricted by their limited phonological abilities, including speech sound production and preferences for specific syllable and word structures, known as templates. The child's vocabulary increases in conjunction with reduced phonological constraints, leading to wider inventories of speech sounds and syllable structures (Stöl-Gammon and Sosa, 2007). When the phonological system has become more mature, words and utterances which are used frequently are stored over time as neuromotor routines and retrieved as Gestalts, leading to the phonological phenomena characteristic of adult speech, including CSPs (Bybee, 2002). Bybee's account of formulaic language is discussed in more detail in Chapter One, in the discussion of research on the analytic versus holistic language learning styles.

Usage-based approaches provide popular accounts of language acquisition in general, incorporating multiple linguistic levels, including phonology, grammar and lexis (see Chapter One). Great emphasis is placed on the importance of input and usage frequency in determining developmental patterns. Bybee's (2002) argument that high frequency words and utterances are stored and retrieved holistically provides an account of phonological phenomena specific to connected speech, as well as those occurring within the individual word.

3.6. The Development of Multi-Word Speech

Multi-word utterances emerge in the second year of life, with single word utterances becoming relatively uncommon in the child's output (Howard et al., 2008; Newton and Wells, 1999; Stemberger, 1988). Matthei (1989) argues that there must be additional phonological constraints on multi-word utterances, otherwise they would be produced earlier in life.

The occurrence of additional phonological simplifications in multi-word utterances has long been recognised in the field of speech impairment, owing to the reduced intelligibility of speech impaired children in connected speech compared with single word production (Faircloth and Faircloth, 1970). However, such interactions were not investigated in typically developing (TD) children until the late 1980s, when three individual case studies were reported (Donahue, 1986; Matthei, 1989; Stemberger, 1988).

Donahue (1986) conducted a diary study of her son's linguistic development, commencing from the emergence of single words at eleven months until 1;10 (years;months), when two-word utterances were well established. She noted several developmental stages from her observations. The first single words were produced accurately. At 1;3, two-word utterances emerged briefly before being eliminated. From this point until 1;6, only single words were produced, but there was evidence of consonant harmony, which co-occurred with a sudden, rapid increase in vocabulary. When two-word utterances emerged again from 1;6 to 1;10, consonant harmony was evident across whole utterances. Donahue (1986) views this developmental pattern as evidence of attempts to work through constraints on articulatory retrieval, motor planning or execution of an utterance. She argues that phonological simplification through consonant harmony served to reduce her son's language processing load and enabled him to extend his vocabulary. She concludes that her son adopted an analytic approach to language learning, showing slow growth and periods of plateau. Donahue (1986)

acknowledges that her study focused on only one phonological phenomenon and predicted the existence of further between-word simplifications and individual differences.

Stemberger (1988) similarly predicted that different phonological phenomena would occur in multi-word utterances compared with those in single word utterances, owing to the increased processing demand on motor planning. He confirmed this prediction and Donahue's (1986) prediction of individual differences in a diary study of his daughter in the early stages of language acquisition. He found evidence of further phenomena including resyllabification of word-final consonants across word boundaries, manner assimilation of word-final plosives prior to word-initial nasals and adult-like elision of word-final cluster elements. Stemberger concludes from these findings that further research is needed in order to fully categorise phonological simplifications occurring between words, as has been done for within-word phonological simplifications.

Similarly to the study of Donahue (1986), Matthei (1989) studied the linguistic development of a boy from the age of eleven months until 1;10, and identified several developmental stages. Between eleven months and 1;3 years, only single words were produced, involving phonological simplifications including consonant harmony, limited plosive production and word shapes reduced to a consonant-vowel (CV) structure. The child's vocabulary increased gradually during this period and was restricted by phonological simplifications. From 1;5 to 1;7, the first two-word utterances emerged and the child's vocabulary increased rapidly, as reported at a similar age by Donahue (1986). Although the above-mentioned phonological simplifications were no longer applied to single-word utterances at this time, they were nevertheless evident in the first two-word utterances. Similar to the conclusion of Donahue (1986), Matthei (1989) interprets his findings as evidence for an analytic language learning strategy, whereby the child mastered skills sequentially in order to avoid processing overload.

More recent studies have been motivated by Stemberger's (1988) recommendation that further research be conducted into the nature and development of phonological simplifications occurring in connected speech. Motivations for this research include the following:-

- The need to advance linguistic theory in order to describe and account for between-word simplifications in both adult speech and phonological development;
- The possibility that individual differences may exist in the development of connected speech owing to individual differences in language learning strategies;
- The need to obtain normative data from TD children with which to compare children with speech impairments;
- The need to better understand the reduced intelligibility of speech-impaired children in connected speech;
- Recent evidence that phonological development may interact with other linguistic and communicative factors, such as syntax and formulaicity.

This research has focused on the behaviours which occur at word boundaries as words are joined in connected speech.

3.7. Word Juncture Behaviours and Connected Speech Processes

Spontaneous connected speech is not formed simply from the linear, sequential production of each composite phone of every word within an utterance (Howard et al., 2008). It is so rapidly produced that it is impossible for the articulators to adopt the ideal configuration for each composite phone of every word (Brown, 1990; Shockey, 2003). Equally, many phonetic cues in the speech signal are redundant for intelligibility owing to the listener's reliance on other linguistic and contextual cues, meaning that effective communication does not necessitate perfect speech sound production (Cruttenden, 2001). The speaker therefore maximises articulatory efficiency and fluency by unconsciously producing those gestures necessary for comprehension with reduced articulatory precision (Brown, 1990; Eastwood, 1981). As words

combine, adjacent phones influence each other, affecting the articulation of individual segments and the prosodic structure of whole utterances, thus producing qualitatively different forms of words in connected speech from the citation forms produced in isolation (Brown, 1990; Cruttenden, 2001; Howard et al., 2008; Shockey, 2003). Most of these influences are anticipatory in English, meaning that features of a phone are realised in the production of the preceding phone (Cruttenden, 2001).

Connected speech also differs from single word production, owing to the increased cognitive and speech processing required. This is evident in adult learners of a second language (Eastwood, 1981) and in children acquiring their first language (Donahue, 1986; Howard et al., 2008; Matthei, 1989).

The phonetic phenomena occurring at word boundaries are known as word juncture behaviours; Juncture (or junction) is a concept which is a major focus of Firthian prosodic analysis (Wells, 1994; Kelly and Local, 1989). Open juncture results from phonetic behaviours which render two words or syllables disjunct. In English, open juncture behaviours include the audible articulation of word-final consonants, the insertion of glottal stops at word boundaries and the occurrence of pauses at word boundaries. Close juncture results from word juncture behaviours which smooth articulatory transitions and increase cohesion between words or syllables (Howard et al., 2008; Wells, 1994). The presence of open or close juncture at word boundaries is affected by many linguistic and communicative factors known as alternations (Farnetani and Recasens, 2010). Open juncture is associated with stressed syllables, content words, emphatic speech, scripted speech, low frequency words, communication of new information and repair of miscommunication. In contrast, close juncture occurs during spontaneous speech and is particularly associated with high-frequency words, unstressed syllables and communication of given information (Bybee, 2002; Farnetani and Recasens, 2010; Howard et al., 2008; Wells, 1994). Shockey's (2003) distinction between scripted and spontaneous speech is used here in preference to Wells' (1994) distinction between formal and casual, colloquial speech conditions. This is because Shockey found no

difference in the extent of phonetic reduction (including word juncture behaviours) occurring in formal and casual conversational speech. She therefore argues that different degrees of phonetic reduction cannot reliably be explained in terms of speaking style and offers the distinction between scripted and spontaneous speech as an alternative explanation. An increased speaking rate may also produce an increase in word juncture behaviours, although this is not necessarily the case (Shockey, 2003). Equally, speaking rate does not necessarily determine speaking style, as rapid speech may be used in more formal, scripted conditions and slow speech may also occur casually and spontaneously (Cruttenden, 2001).

The automatic occurrence of context-dependent gestural overlap as phones influence each other is known as coarticulation. The resultant phonetic changes are considered to be continuous rather than categorical, and they may or may not be auditorily perceptible (Catford, 1977; Farnetani and Recasens, 2010). Coarticulation has traditionally been viewed as an innate, automatic, language-universal property of speech, not rooted within linguistic rules, but resulting purely from the physical properties of the articulators as gestures overlap in time (Chomsky and Halle, 1968; Farnetani and Recasens, 2010; Harris, 2003).

Shockey (2003) argues that some aspects of coarticulation are phonological, because they occur regularly in predictable phonetic environments according to language-specific conventions. She also notes that this varies across different regional accents. CSPs are a group of specific word juncture behaviours and word form reductions which are traditionally believed to be of phonological origin and learned alongside other linguistic aspects according to language-specific rules (Brown, 1990; Chomsky and Halle, 1968; Cruttenden, 2001; Farnetani and Recasens, 2010; Harris, 2003; Shockey, 2003). They occur optionally alongside other close juncture alternatives and open juncture. However, adult speakers tend to produce CSPs in preference to other forms at CSP sites and their absence is considered to be unusual and artificial (Cruttenden, 2001). The three CSPs which have been most frequently studied in recent developmental research will now be defined.

Assimilation is the contextual variability which results when the articulatory features of a phone (voicing, place or manner) are modified to resemble more closely those of an adjacent phone, resulting in a categorical change in contrastive phone (Cruttenden, 2001; Farnetani and Recasens, 2010; Catford, 1977). Regressive (or anticipatory) assimilation occurs when the phonetic properties of a word-final consonant are influenced by those of the following word-initial consonant, whereas progressive (perseverative) assimilation occurs when the phonetic properties of a word-initial consonant are influenced by those of the preceding word-final consonant. Regressive velar assimilation of alveolar consonants is the most commonly occurring assimilation type in English, closely followed by bilabial assimilation of alveolars (Brown, 1990; Cruttenden, 2001). Although *Gimson's Pronunciation of English* provides examples of word-final [t] assimilation (Cruttenden, 2001), Shockey (2003) points out that word-final [t] usually takes the form of a glottal stop in casual speech and is therefore not assimilatory. An example of bilabial assimilation would be in the utterance *can be* in which the final /n/ in *can* may adopt a bilabial place of articulation prior to the initial /b/ of *be* to become /kæm bi/. An example of velar assimilation would be the utterance *can go*, in which the final /n/ in *can* may adopt a velar place of articulation prior to the initial /g/ in *go* to become /kæŋ gəʊ/. In casual speech, word-final bilabial plosives and nasals may assimilate to an alveolar or velar place of articulation, although this may be considered by some to constitute substandard pronunciation (Cruttenden, 2001). Examples of this would be the pronunciation of *I'm not* as [aɪn nɒt] and *I'm going* as [aɪŋ gəʊɪŋ].

Consonant elision involves the deletion of one or more word-final consonants in a cluster of three or more consonants formed by the abutting coda and onset of two adjacent syllables or words. Such elision would not be expected in careful speech. An example of elision would be the pronunciation of *mashed potato* as /mæʃ pəteɪtəʊ/. Shockey (2003) reports that elision is more likely to occur in monomorphemic words such as *past* than in multimorphemic words such as *passed*, where the most likely sound to be elided represents a morpheme. Alveolar plosives are particularly vulnerable to both assimilation and this elision type (Shockey, 2003). Elision and assimilation may co-occur (Cruttenden, 2001). An example of this would be the

pronunciation of *kind man* as /kaɪm mæn/. A further elision type is that of word-initial /h/ following a word-final consonant (Cruttenden, 2001). An example would be the pronunciation of *find him* as /faɪnd ɪm/.

Liaison is the smoothing of an articulatory transition between two heterosyllabic vowels by the insertion of an approximant (Cruttenden, 2001; Gick, 1999). Liaison of [j] occurs following the vowels [i], [ɪ], [aɪ], [eɪ] and [ɔɪ]. An example would be the pronunciation of *tidy up* as /taɪdij ʌp/. Liaison of [w] occurs following the vowels [u], [aʊ] and [əʊ]. An example would be the pronunciation of the utterance *do it* as /duw ɪt/. Gimson argues that these liaison types occur at a phonetic, coarticulatory level rather than at a phonological level (Cruttenden, 2001). He exemplifies this by comparing the utterance *my ears* which may be produced with [j] liaison with the utterance *my years*, which contains the contrastive phone /j/.

Liaison of /r/ occurs following word-final schwa or low back vowels. In contrast with liaison of [j] and [w], /r/ liaison is believed to be an aspect of English phonology, rather than resulting purely from coarticulation (Cruttenden, 2001). One reason for this difference may be that /r/ does not share phonetic properties with preceding vowels at liaison sites in the same way as do the glides [j] and [w] (Newton and Wells, 2002). Brown (1990) describes /r/ liaison as usually involving the resyllabification of a final /r/ in a stressed syllable to the initial position of a following unstressed syllable. However, she also acknowledges that in instances when grammatical words contribute to /r/ liaison sites, the final /r/ in an unstressed syllable (grammatical word) is resyllabified to the initial position of a stressed syllable. Shockey (2003) views /r/ liaison as a manifestation of the tendency towards consonant-vowel alternation in spoken English. Liaison of /r/ is subdivided into linking /r/ and intrusive /r/. The surface manifestation of these subtypes is identical, but the underlying origin is believed to be different. Linking /r/ refers to liaison occurring when a word-final r is present orthographically and was historically pronounced (Cruttenden, 2001; Gick, 1999). An example would be the pronunciation of *far away* as /fɑr əweɪ/. In contrast, intrusive /r/ is the introduction of /r/ liaison in the absence of word-final r either in historic pronunciation or

spelling. It occurs less frequently than linking /r/ and may be considered to be substandard English (Cruttenden, 2001); (Gick, 1999). Examples would be the pronunciation of *law and order* as /lɔr ən ɔdə/ and *idea of* as /aɪˈdɪər ɒv/. In the absence of liaison, the second part of a diphthong may be absorbed at [j] and [w] liaison sites, vowel glides may occur at /r/ liaison sites, or glottal stops may be inserted to create open juncture at all liaison sites (Cruttenden, 2001).

Besides these three CSPs, others exist including vowel elision, phonetic reduction of vowels and words (often found in grammatical words) and coalescence (fusion of the places of articulation of two adjacent consonants (Barry and Andreeva, 2001; Cruttenden, 2001; Shockey, 2003). The remainder of this section focuses mainly on assimilation, which is the primary focus of the current study.

The existence of the traditional distinction between coarticulation and CSPs is widely debated, especially with regard to assimilation. Farnetani and Recasens (2010) suggest that this distinction may have arisen solely through different methods of investigation. They propose that the concept of assimilation has evolved from auditory-perceptual phonetic analysis, whereas the concept of coarticulation has been reached through instrumental phonetic analysis. Because context-dependent variability in articulatory movements can lead to phonetic variation both in auditory percept and acoustic cues, they argue that it is impossible to determine which coarticulatory behaviours are universal and which are language-specific, and whether these behaviours are governed by the same or different underlying phonetic and phonological abilities.

Similarly, Harris (2003) argues that the theoretical distinction between assimilation and coarticulation is difficult to apply in practice to individual instances. In agreement with Shockey (2003), he also notes that assimilation necessarily involves coarticulation. As a solution to the debate, he proposes that the term assimilation be used only to refer to instances when the overlap of phones is grammar-internal, in other words, when there is evidence of the

coarticulation contributing to the linguistic content of an utterance. Examples of grammar-internal assimilation would be those which indicate word boundaries or the start of a foot (Harris, 2003).

Evidence for the phonological nature of CSPs comes from the results of cross-linguistic comparisons of comparable phonetic phenomena. Barry and Andreeva (2001) compared spontaneous dialogues of adult speakers of six different European languages using auditory-perceptual and spectrographic analysis. Two of the languages were stress timed, two were syllable timed and two were considered to be intermediate. The CSPs investigated were consonant elision, residual phonetic properties of elided consonants, syllable elision, weakened consonant reduction and vowel reduction. They found that similar reduction phenomena occurred across the six languages, but with evidence of cross-linguistic variability. They concluded that this variability occurred owing to the different durations and articulatory effort invested at different points within utterances across the different languages. Similar findings were obtained in a perceptual cross-linguistic comparison of assimilation, where it was noted that the same two adjacent phones across a word boundary led to assimilation in Dutch, but not in Czech (Rechziegel, 2001).

Evidence for the more continuous phonetic overlap described in traditional definitions of coarticulation comes from instrumental studies which have identified continuity across phonetic boundaries, with individual differences between speakers of the same language. Wright and Kerswill (1989) used a perceptual experiment to investigate whether phonetically trained listeners could detect different degrees of assimilation in two-word utterances which were otherwise phonetically identical. A trained phonetician recorded the utterances which involved complete, partial or no assimilation of word-final alveolar nasals. Utterances containing either underlying bilabial or velar nasals were also produced as controls. Electropalatography (EPG) was used to ensure that the speaker was producing the desired degree of assimilation in each utterance. The listeners' judgements were made in the form of phonetic transcription, word identification and categorisation of each utterance according to

the four assimilatory conditions. They found that increased assimilation produced fewer correct identifications, with non-assimilated utterances most correctly identified and fully assimilated utterances least correctly identified. Fully assimilated utterances and utterances containing underlying bilabial or velar nasals were most often confused. The authors concluded from these results that there is a perceptual continuum of assimilation. The authors acknowledge that their study could be criticised for employing both phonetically trained speakers and listeners, who may process language differently from naive participants as a result of their training. However, they argue that there is no reason to suggest that this would have reduced the ecological validity of their results (Wright and Kerswill, 1989). They state that the utterances were recorded multiple times and that the tokens which most closely resembled the spontaneous speech of a naïve speaker in a parallel observation were selected. They also argue that their phonetically trained listeners performed similarly to a group of naïve listeners in a parallel observation.

Following these findings, Ellis and Hardcastle (2002) investigated individual differences in the assimilation produced by ten English speakers with a range of British accents. The participants read aloud a set of sentences which enabled the comparison of underlying alveolar nasals at assimilation sites with underlying velar nasals. The sentences were otherwise phonetically identical. The speakers were requested firstly to read the sentences slowly and carefully, and then to read them in a faster, more casual manner. EPG and Electromagnetography (EMA) were used to monitor the speaker's articulatory movements. Complete velar assimilation was extremely rare in the careful speech condition and more frequent in the casual speech condition, indicating that the two elicited speaking conditions produced different articulatory behaviours. Partial assimilation was relatively rare in both conditions. Non-assimilation, partial assimilation and complete velar assimilation were all found in the casual speech condition, with evident inter-speaker differences. Two participants never assimilated, while four consistently produced apparent complete assimilation. The remaining four produced a variety of realisations. Two appeared to produce either apparent complete assimilation or non-assimilation as two discrete categories, while the other two

appeared to produce the three realisation types along an articulatory continuum. A general tendency towards assimilation was identified in some participants who produced apparent complete assimilation in both speaking conditions. This study therefore shows that assimilation may have gradient, coarticulatory properties and that inter-speaker and intra-speaker differences exist.

It would appear from the studies described above that the methodological distinction between CSPs and coarticulation is not as clear-cut as Farnetani and Recasens (2010) suggest, at least in terms of perceptual versus instrumental analytical methods. The research by Barry and Andreeva (2001) promoting the phonological and language-specific status of CSPs was not based purely on auditory-perceptual evidence, but also involved spectrographic analysis. Equally, the EPG study of Wright and Kerswill (1989) promoting the coarticulatory nature of assimilation involved a fundamental auditory-perceptual element. A further issue which complicates interpretation of these findings is the different types of speech data used in different studies. The cross-linguistic studies of Barry and Andreeva (2001) and Rechziegel (2001) were both based on spontaneous speech occurring either in natural conversation or semi-structured interviews. In contrast, the coarticulatory studies of Wright and Kerswill (1989) and Ellis and Hardcastle (2002) were based on speech elicited in highly controlled experimental conditions and which in the latter case, was read from a script. However, despite these methodological differences, the results indicate that both coarticulatory and phonological aspects are involved in the production of CSPs, especially in assimilation.

Local (2003) argues that there are also lexical effects on the extent to which assimilation occurs. He reports that assimilation occurs more often in unstressed grammatical forms such as *I'm*, than in stressed, lexical forms such as *time* and *lime*. He suggests that this is because the final /m/ in *lime* needs to be clearly articulated in order to distinguish it from phonetically similar words such as *line*. The same level of phonetic contrast is not necessary for grammatical words (Local, 2003).

It is clear from the evidence above that the nature of CSPs in adult speech is extremely complex and that there is wide inter-speaker and intra-speaker variability. It is therefore predicted that such variability also exists in the connected speech of children acquiring language. In the following section, several exploratory studies of CSP development are discussed, in which this prediction is borne out.

3.8. The Development of Connected Speech Processes

The first investigation into the typical development of CSPs was conducted by Newton and Wells (1999). They proposed that CSPs are articulatory in origin and result from developmental immaturity in multi-word utterances, similarly to the phonological simplifications proposed at the single word level (Stampe, 1979). They therefore predicted that no change in the realisation of CSP sites would be observable with increasing age. They compared the realisation of assimilation, elision and liaison sites in groups of children aged three, four, five, six and seven years across sentence repetition, story re-telling and spontaneous speech tasks. This allowed for a methodological comparison across elicited, partially elicited and fully spontaneous conditions respectively. Each CSP site produced by each child in each condition was phonetically transcribed and categorically coded for quantitative analysis. The categories were close juncture (CSP present), partial CSP or open juncture (CSP absent). Group means were then calculated to compare age differences.

The spontaneous speech condition contained the most CSP sites, followed by sentence repetition, followed by story re-telling. These findings were statistically significant, although actual values varied little. There was no significant difference in the occurrence of CSPs as a function of age. Additional findings were that [t] was more often assimilated or elided than either [d] or [n]. Liaison of [j] occurred most frequently, while /r/ liaison occurred least frequently. Newton and Wells also compared the acquisition of allomorphs of the definite and indefinite articles across the age groups, as a syntactic comparison against their phonetic

measures. In contrast with their findings for CSPs, they found that usage of the correct allomorphs of articles increased with age, and that similarly to the CSP sites, articles were produced most frequently in spontaneous speech and least frequently in story re-telling. Newton and Wells tentatively concluded from their findings that CSPs are phonetic in origin and therefore emerge alongside the onset of multi-word utterances, in contrast with syntactic phenomena, which are learned over time.

Following the lack of evidence for developmental trends in the acquisition of CSPs, Newton and Wells (2002) conducted a further study, in order to investigate whether developmental trends were evident at an earlier age and to determine whether a combination of quantitative and qualitative methods could better capture these trends than the purely quantitative approach previously used. In contrast with the previous cross-sectional study, this was an individual case study, aimed to monitor one child's developmental progress from the onset of multi-word utterances at 2;4 until 3;4. The child was observed for an hour fortnightly during free play and his speech was recorded. Whereas the previous study identified CSP sites based on expected adult forms, this study investigated CSP sites as they occurred in the child's output forms, meaning that ungrammatical forms which do not occur in adult speech could be investigated. CSP sites were phonetically transcribed and were categorised according to complete presence of CSPs, absence of CSPs and idiosyncratic forms (those not encountered in adult speech), although there was no category of partial CSPs. Qualitative trends in the development of individual CSPs were also investigated.

The number of CSP sites in the child's output increased over time alongside increases in vocabulary and syntax. Adult-like bilabial and velar assimilations and idiosyncratic forms (such as glottal stop insertion) were evident from the outset. Open junctures emerged at 2;9, during an apparent developmental transition at which all three realisation types co-occurred. From 2;10 onwards, idiosyncratic forms decreased and open junctures increased, although adult-like assimilations always predominated. The occurrence of assimilation appeared to interact with grammatical development. The prepositions *in* and *on* and the article *one* in noun

phrases accounted for most early occurrences, and the auxiliary verb *can* accounted for nearly half of the occurrences in later months.

Adult-like elision also occurred from the outset, with idiosyncratic forms featuring from 2;5 to 2;10. These forms involved elision accompanied by additional phonetic phenomena including cluster elision, glottal stop insertion and vowel nasalization. At 2;4, open junctures emerged briefly before being eliminated. Open junctures re-emerged at 2;10, a month after idiosyncratic forms had disappeared, providing further evidence for a developmental transition period from 2;9 to 2;10 years. From 2;10 until 3;4, elision sites were realised predominantly with elision and less frequently with open juncture.

The authors argue that the prerequisite ability to articulate complex clusters explains the late emergence of open juncture at elision sites. Similarly to the findings for assimilation, elision always predominated over open juncture. One possibility not suggested by the authors is that some instances of elision may not have occurred purely at the phonetic and phonological levels, but rather at a morphological level as the result of morphemic omission. For instance, they give the example "drop sausage" for the target *dropped sausage*, which may have resulted either purely from [t] elision or from underlying grammatical omission of the past tense inflection *ed*.

Liaison of [j] was evident from the outset with close juncture realisations of all potential sites. However, from 2;7 to 2;9, the number of potential sites increased and were mostly realised with open juncture through the insertion of a glottal stop. From 2;10 onwards, close juncture predominated again alongside the developmental transitions in assimilation and elision which occurred at this age.

Interactions with grammatical development were also evident. Pronouns followed by forms of the verb *be* (such as "there they are") often produced liaison, but liaison did not occur in grammatically incorrect forms resulting from the omission of *be* (such as "he upside down"

and "I upstairs"). A possibility not suggested by the authors is that close juncture was more associated with formulaic utterances, while open juncture reflected some underlying knowledge of grammatical incorrectness.

Liaison of [w] was the most prevalent liaison type, although some sites were realised with open juncture through glottal stop insertion. Many of these occurred in the utterance *do it*, which (as suggested by the authors) may have been a formulaic utterance stored as a gestalt. From 2;11 onwards, close junctures increased and open junctures declined, although [w] liaison was never produced with the same consistency as either assimilation or elision.

Unlike the other liaison types, /r/ liaison, realised immaturely as [ʊ] emerged suddenly at 2;11. The authors therefore concluded, in agreement with their previous findings that assimilation, elision and liaison of [j] and [w] are phonetic in origin and therefore, emerge automatically with the onset of multi-word utterances. The exception is /r/ liaison, which they concluded to be of phonological origin. They suggest that /r/ liaison may differ from the other liaison types because unlike [j] and [w], /r/ does not share phonetic properties with the vowels which precede it at liaison sites. They also concluded that the combined quantitative and qualitative approach in this study highlights early developmental trends which were not evident from their previous (1999) study.

Thompson and Howard (2007) combined the mixed qualitative and quantitative approach of Newton and Wells (2002) with the cross-sectional design of Newton and Wells (1999). They compared CSP site realisations in the spontaneous connected speech of two- and three-year-olds. Six children were studied in total, with three children in each age group. Each child was observed and recorded individually in a free play session of 40 minutes. Potential CSP sites were identified and analysed using orthographic and phonetic transcription and were categorised as having either open or close juncture. The proportions of open and close junctures were calculated for each participant, and mean percentages were calculated for each

age group in order to make between-group comparisons. The mean length of utterance (MLU) of each child was also calculated as a comparative syntactic measure.

Two of the two-year-olds produced similar numbers of open and close junctures, whereas the third produced slightly more open than close junctures. In contrast, all of the three-year-olds produced more close than open junctures. However, there were considerable individual differences. The authors acknowledge that the statistical significance of the between-group differences was greatly influenced by a syntactically more advanced child in the older group (as measured by MLU) who produced the most close junctures, and one syntactically less advanced child in the younger group who produced the most open junctures. They also noted that the children with lower MLU produced fewer CSP sites, indicating a general effect of linguistic advancement on performance in both phonological and syntactic domains.

Adult-like word-final elision was the most frequently occurring and well established CSP across both groups and was observed even in the children with low MLU. Bilabial assimilation was better established than velar assimilation in both groups. The older children produced adult-like bilabial and velar assimilation with greater consistency than the younger children, although some unusual assimilations of voicing, place and manner were produced by the older children. /r/ liaison was observed only in the most syntactically advanced older child, although the phonetic forms used are not stated. The other children realised /r/ liaison sites with open juncture.

Thompson and Howard (2007) and Newton and Wells (2002) similarly concluded that assimilation and elision emerge early in the development of multi-word speech, whereas /r/ liaison emerges later and is therefore acquired differently from the other CSPs. The authors of both studies therefore suggest an inequality between the CSPs, with assimilation and elision having more phonetic than phonological origins and liaison having more language-specific, phonological origins. However, the authors draw different conclusions regarding the developmental process. Newton and Wells (2002, 1999) interpret their findings of consistent

close juncture predominance as evidence for the automatic, coarticulatory nature of CSPs. In contrast, Thompson and Howard (2007) interpret their finding of increased close juncture with age as evidence that CSPs are acquired as a result of gradual phonological refinement. Thompson and Howard (2007) highlight the importance of recognising individual differences in CSP development similar to those found in adult speakers. They suggest that such variability may reflect the differential effects of analytic versus holistic language learning strategies, where the former would lead to the initial predominance of open juncture, and the latter would lead to the predominance of CSPs from the onset of multi-word utterances. The relationship which they observed between MLU, CSP sites and proportions of close junctures indicates interactions between phonological and syntactic development. They therefore recommend further research in this area. One difficulty with this study is that the groups were not balanced for age and sex of participants, meaning that any sex differences or within-group age differences would not have been detected. Therefore, further research is also needed to compare males and females and to study the incremental development of CSPs with age.

Following Thompson and Howard's (2007) suggestion of interactions between CSP development and acquisition of other linguistic skills, Howard, Methley and Perkins (2008) studied the distribution of word juncture types and possible interactions with other linguistic phenomena in one TD child from age 2;3;2 to 2;10;8 (years;months;days). The child was observed in his own home during four hour-length sessions at ages 2;3;3, 2;5;7, 2;7;9 and 2;10;8. Half an hour of each session was recorded and analysed. Open junctures were in the minority throughout the study, with under a third of CSP sites realised with open juncture at any point. Typical assimilation and elision were found at CSP sites along with some unusual behaviours. At age 2;7;9, a significant increase in close juncture co-occurred with a reduction in MLU growth, suggesting competition between different linguistic processing demands. There was evidence for a trade-off between paradigmatic accuracy and syntagmatic fluency, in that open juncture was associated with accurate articulation of segments and disrupted prosody, while close juncture was associated with reduced segmental accuracy alongside more typical prosodic features. The authors concluded from this finding that the child adopted a

holistic language learning strategy by storing familiar utterances as gestalts, while syntactically productive or complex sentences relied more on segmental phonological processing and were therefore produced with open juncture. Formulaic utterances were produced more frequently, more fluently and with more instances of close juncture than novel, productive utterances. Close juncture was more often associated with high frequency verbs which serve a grammatical function in English such as *do*, *be* and *have*, whereas open juncture was more often associated with lower frequency lexical verbs such as *hide*. At a pragmatic level, open juncture was associated with the expression of emotional state, the use of novel vocatives at age 2;7;9 and giving new information (as found in adults). The authors conclude that their findings support a holistic and emergentist approach to language processing, in which processing load is spread across linguistic domains.

These findings confirm the suggestion of Thompson and Howard (2007) that interactions exist between the development of CSPs and other linguistic abilities. They also confirm the conclusions of Thompson and Howard (2007) and Newton and Wells (2002) that assimilation and elision occur in early multi-word speech and are therefore more likely to be of greater articulatory than phonological origin.

Following the findings of these studies, Bryan, Howard and Perkins (2010) conducted a pilot study which has led directly to the current investigation. The development of assimilation, elision and liaison was investigated in a TD boy, Brian, from age 2;0;12 until 4;0;4. Previously collected audio recordings were selected for analysis at approximate monthly intervals from a dense database of child language data (Lieven et al., 2009). Each recording was approximately an hour long and consisted of Thomas's spontaneous interactions with his mother while engaging in everyday activities in his home. Thomas's developmental progress was thus analysed at much more regular intervals and over a longer period than any other study to date, and was based entirely on spontaneous connected speech. Ten potential sites for assimilation, elision and each of the three liaison types were identified in each recording session (subject to availability). Each focal word juncture was phonetically transcribed and

analysed according to the juncture types present and their relative proportions over time. General trends in the occurrence of idiosyncratic forms were also analysed.

Assimilation was observed throughout the sampling period, although assimilation site realisation varied greatly with alternating periods of predominance and non-predominance over other forms. The duration of predominance periods increased over time relative to periods of non-predominance, indicating that assimilation was produced with increasing consistency over time. This finding was also reported by Thompson and Howard (2007). However, adult-like consistency was not evident in Brian's data by the end of the sampling period. Bilabial assimilation was more established than velar assimilation, also in agreement with the findings of Thompson and Howard (2007). The most frequently occurring idiosyncratic behaviours at assimilation sites were elision of word-final consonants and non-assimilation owing to the alveolar articulation of word-initial velar plosives. Bryan (et al.) suggested that the latter may have resulted from progressive alveolar assimilation.
These findings differ from those of Newton and Wells (2002), who report that assimilation was consistently predominant throughout their study, apart from during a developmental transition period from 2;9 until 2;10. Differences between these findings may partly result from different categorisation of close juncture behaviours. Following the suggestion of Cruttenden (2001), Bryan et al. (2010A) included word-final [t] in their analysis, but found that, as suggested by Shockey (2003), this was mostly produced with a glottal stop. Bryan et al. therefore analysed these glottalisations as separate close juncture behaviours from assimilation. When glottalisation and assimilation were initially analysed together as close juncture behaviours, proportions of these close juncture behaviours predominated overall from age 2;6;2, in agreement with Newton and Wells (2002).

Word-final alveolar consonant elision occurred with the emergence of potential sites at 2;5;3 and was produced with high consistency throughout the study, with a minority of open junctures and idiosyncratic forms emerging in Brian's fourth year. This elision type was unaffected by a period of high phonological variability at 3;4;3 compared with other CSPs.

The most frequently occurring idiosyncratic forms were cluster elision of abutting word-final and word-initial consonants and instances in which it was unclear which of two abutting consonants had been elided. The former behaviour was also reported by Newton and Wells (2002).

The emergence of open junctures and idiosyncratic forms at elision sites during the fourth year contrasts with the findings of Thompson and Howard (2007) and Newton and Wells (2002), who both reported more occurrences of these behaviours in the third year of life. This contrast provides further evidence for individual differences in the acquisition of CSPs. A further consideration is that Bryan's study may have captured relatively late phonological transitions which could not be captured in previous studies.

Bryan also studied elision of word-initial /h/ following a word-final consonant, an elision type not previously studied. Similarly to word-final consonant elision, /h/ elision emerged with the onset of potential sites at age 2;7;2. Open juncture emerged at 2;10;5, similar to the age at which Newton and Wells (2002) found open juncture emerging at final alveolar consonant elision sites. Idiosyncratic forms emerged from 3;1;3 to 3;7;3, at a similar age to the emergence of idiosyncratic forms at final alveolar consonant elision sites. The most frequently occurring idiosyncratic form was hyperelision of the entire consonant cluster, as found at final alveolar consonant elision sites.

Liaison of [j] emerged at 2;6;2, one month after the emergence of potential sites at 2;5;3. Liaison consistently predominated over idiosyncratic forms and open juncture, with a minority of non-predominance periods. The finding of overall predominance agrees with that of Newton and Wells (2002), with the main difference being that Newton and Wells observed the emergence of liaison alongside potential sites, whereas Bryan et al. observed an initial period of open juncture predominance. Similarly to the findings for final alveolar consonant elision sites, open junctures re-emerged at 3;5;3 to 3;9;3, contrasting with the younger age of 2;9 at which Newton and Wells (2002) observed similar phenomena. Idiosyncratic forms were in the

minority, but occurred more frequently than open junctures. These included smooth vowel transitions across word boundaries resembling diphthongisation, and elision of word-initial vowels which removed liaison sites. The former is also reported in adult speakers (Cruttenden, 2001). A further phenomenon occurring between two and three years was substitution of a different consonant for [j] at the liaison site.

Liaison of [w] emerged with potential sites at 2;4;3 and either predominated over or occurred in equal proportions to idiosyncratic forms and open junctures throughout the study, except for one period of non-predominance. The first occurrences were in utterances such as *do it*, which Newton and Wells (2002) also observed and explained in terms of gestalt storage for frequent, formulaic utterances. Open junctures were rare and emerged during the fourth year alongside open junctures at other CSP sites. Similar idiosyncratic forms were observed to those occurring at [j] liaison sites, including smooth vowel transitions, word-initial vowel elision and consonant substitution.

Liaison of /r/ emerged at 2;8;3, five months after the emergence of potential sites. The phonetic realisation of /r/ was initially the labiodental approximant, although the labial-velar approximant was introduced as an alternative at 2;11;3. This corresponded with the first liaison predominance, which occurred at the same age as liaison emergence in the participant of Newton and Wells (2002). Liaison consistently predominated from 3;3;3 onwards, except at age 3;4;3, when the participant showed high phonological variability. Mature liaison first occurred at 3;8;2, a year the emergence of immature forms. By the end of the study, both immature and mature forms were produced with changing predominance from month to month. Open junctures emerged with the emergence of liaison sites at 2;3;2, which occurred more frequently than both idiosyncratic forms at /r/ liaison sites and open junctures at [j] and [w] liaison sites. Idiosyncratic forms often involved consonant substitution and word-initial vowel elision, as found at [j] and [w] liaison sites. The late emergence of /r/ liaison compared with [j] and [w] liaison corresponds with the findings of Newton and Wells (2002), and supports their suggestion that /r/ liaison is phonological in origin and is learned over time along with other phonological phenomena. However, the similarity of idiosyncratic forms

which Thomas produced across different liaison sites suggests that all liaison types are to an extent phonetically related.

The current study follows directly on from this previous research on the development of connected speech. However, this study focuses only on assimilation, in order to investigate in detail the interactions occurring between syntax and one specific phenomenon. Bilabial and velar assimilation are also investigated individually, as reported by Bryan et al. (2010). The current study also explored the effect of maternal input on the development of assimilation, which is a new line of investigation. Before moving on to the aims and method of the current study, it is therefore necessary to discuss research on the phonology of child-directed speech.

3.9. Phonological Characteristics of Child-Directed Speech

Research has established that child-directed speech (CDS) differs from adult-directed speech (ADS) in several ways. Linguistic characteristics of CDS include simplifications in syntax and vocabulary, shorter utterances and multiple repetitions of the same utterance. Differences in suprasegmental phonological characteristics have also been found, including adaptation of words to a CVCV syllable structure, slower speech rate, longer duration of pauses and wider pitch range (Foulkes, Docherty and Watt, 2005; Khattab, 2006). However, relatively few studies have focused on the segmental phonological characteristics of CDS (Foulkes et al., 2005).

CDS is characterised by greater acoustic contrasts between vowels than those found in ADS (Andruski, Kuhl and Akiko, 1999; Bernstein Ratner, 1984; Kuhl et al., 1997). This finding is consistent across the different languages which have been investigated: American English, Swedish, Russian and Japanese (Andruski et al., 1999; Kuhl et al., 1997). In addition, there exists evidence of reduced contrast and increasing formant overlap between vowels as a function of the child's increasing MLU (Bernstein Ratner, 1984).

The research of Patricia Kuhl and colleagues focused on CDS addressed to young infants, no older than eight and a half months (Andruski et al., 1999). Bernstein Ratner's earlier study examined vowel quality in speech addressed to slightly older children, although she does not specify the children's exact ages. The only information which she gives is that all children were above nine months of age at the onset of the study, some were 17 months of age at the start of the study, and that recordings took place over a period of six months. She argues that it is more important to analyse the characteristics of CDS in relation to the children's language ability, than simply in relation to their ages. She points out that children of a similar age may be at very different stages in terms of language acquisition. She therefore characterises the vowel qualities in CDS in relation to the children's MLU. She mentions a maximum MLU of four, which falls within Brown's stage V (Bowen, 1998; Brown, 1973). In a critique of Bernstein Ratner's study, Kuhl et al. (1997) conclude that based on the information given, the older children were aged between two and four years on these grounds. However, this seems questionable, considering that all nine children recruited for the study were born over a two-year period.

It can be concluded from this research that there exist more distinct acoustic contrasts between vowels in CDS than in ADS. The contrast between vowels reduces, leading to increased overlap, as a function of advances in the child's linguistic ability. However, there exists no direct information on the possible changes as a function of children's increasing age, or the point at which vowels in CDS become equivalent to those in adult-directed speech.

In contrast with the well-established pattern of vowel production in CDS, research findings concerning consonant patterns in CDS present a much less clear picture. Some comparisons of CDS and ADS indicate that primary caregivers produce simplified speech, with more consistent phonological patterns, greater distinction between contrastive phones, fewer phonological reductions and more standard variants in CDS. However, other studies have found more evidence of phonological variability, reduction and vernacular variants in CDS.

Each study has focused on a different phonological pattern. It is therefore possible that CDS involves a combination of increased consistency and increased variability, producing different findings, depending on the parameters investigated.

The first study to compare patterns of consonant production in CDS and ADS involved eight mother-child dyads in south-east England, with children aged between two and four years (Shockey and Bond, 1980). The researchers recorded samples of spontaneous mother-child interaction and conversation between the mother and a research assistant. They hypothesised that mothers would simplify speech by maximising phonological distinctions, in order to enable the child to learn correct phonological representations. They focused on four phenomena characteristic of British English:-

- Substitution of [ʔ] for /t/;
- Elision of word-initial [ð];
- Coalescence of word-final /t/ and /d/ prior to word-initial /j/, For example, *bet you* [bɛttʃu] and *did you* [dɪd͡ʒʊ] and
- Reduction of the cluster /ts/ to /s/, for example, *it's a* [ɪs ə].

Contrary to their prediction, they found more instances of these phonological reductions in CDS than in ADS. The exception was coalescence, which was infrequently sampled in the data and could not therefore be analysed further. They found considerable individual variability in the extent to which these phenomena occurred, both in the CDS and ADS conditions. They were struck by the lack of phonological simplification in CDS, which contrasted sharply with the highly evident syntactic and lexical simplifications which they observed. They interpret this finding as evidence that phonological reductions are less conscious and more habitual than other linguistic behaviours.

Shockey and Bond propose that the phonological reduction found in CDS serves the social function of establishing a tone of intimacy between the mother and the child. They further propose that despite multiple exposures to both reduced and non-reduced forms, children

acquire the correct phonological representations for lexical items through their prior knowledge of the distribution of contrastive phones within a language. For instance, they may learn that [t] is a more standard pronunciation of /t/ than [ʔ], because [t] occurs more frequently in English. However, there exists a fundamental problem with this explanation. It is possible that in some dialects of English, such as those occurring in the Southeast of England, the glottal stop occurs more frequently in medial and final positions than [t]. In this case, children would receive no evidence in the input that [t] is more standard. One methodological confound of this study is that the ADS condition involved conversations between the mothers and the research assistant, an unfamiliar adult. This raises the question of whether the extent of phonological simplification would have been more similar across the conditions, or even more evident in the ADS condition if the ADS condition had involved conversation with a familiar adult, such as a friend or family member.

This issue was addressed in a more recent, larger-scale study of CDS and ADS (Foulkes et al., 2005). This study focused on the phonetic variations of /t/ occurring in Tyneside English. These included:-

- The standard English variant [t];
- Voiced variants such as [d] and [ɾ], which are characteristic of Tyneside English;
- The vernacular variant [ɹ] common to dialects used in the north and midlands of England;
- The form [tˀ…] in syllable coda position, which is considered especially characteristic of the Tyneside dialect.

The CDS sample was collected from 39 mother-child dyads during a play session, with toys and a book provided to elicit specific speech sounds. The children's ages ranged from two to four years. The mothers' data were then compared with that of 32 adults from a previously collected sample of ADS. This comparison focused particularly on the young working class women from the adult corpus, who were the closest demographic match to the mothers in the CDS sample.

For words containing /t/ in medial position, the mothers in the ADS condition produced glottal stops in 90% of instances and [t] in only 10% of instances. In contrast, the occurrence of [t] in the CDS condition increased to 59%, while production of glottal stops was reduced to 36%. Variations in the CDS condition were also observed as a function of the children's age and sex. More productions of [t] were observed in the mothers of two-year-olds, compared with more glottal stops in the mothers of four-year-olds. This was statistically significant for mothers of girls, but only a non-significant trend for mothers of boys. Mothers of girls showed higher proportions of [t], whereas mothers of boys produced more glottal stops.

For words containing final /t/ in prevocalic contexts (for instance, in the utterance *get in*), higher proportions of [t] and voiced variants occurred in CDS, whereas higher proportions of Glottal stops and [ɹ] were observed in ADS. Again, there were additional effects as a function of the children's age and sex. Mothers used higher proportions of [t] in their speech to younger children and girls.

In contrast with Shockey and Bond (1980), Foulkes et al. (2005) conclude that CDS provides children with more information on standard phonological contrasts. They further suggest that children use the ADS which they hear in order to learn vernacular variants. They argue that the sex differences which they observed resulted from the mothers' global behavioural adaptations, in line with their children's gender identity. There are several possible reasons for the sharply contrasting results of Shockey and Bond (1980) and Foulkes et al. (2005). Firstly, the studies looked at different phonological variables occurring in different English dialects. Secondly, Foulkes et al. had a much larger data sample of 39 mother-child dyads, compared with the eight dyads tested by Shockey and Bond. Thirdly, the ADS data corpus of Foulkes et al. consisted of conversations with self-selected conversation partners. This data was therefore based on conversation with a familiar adult, whereas the ADS data collected by Shockey and Bond was based on conversations with unfamiliar adults. However, the methodological disadvantage of the study by Foulkes et al. is that the mothers' CDS was

compared with the ADS of different adults. Although the adults in the two conditions were closely matched in terms of demographic characteristics, it is possible that inter-speaker variability was a confounding variable in this design. It would be valuable to conduct a similar study, in which mothers' CDS was compared with the mothers' speech when in conversation with a familiar adult.

A further study aimed to investigate whether mothers maximised distinctions between contrastive phones which are potentially confusable (Khattab, 2006). This investigation involved the comparison of contrastive singleton and geminate consonants in Lebanese Arabic. Spontaneous interactions of five mother-child dyads were recorded at the start and end of the children's single-word phase. The children were aged between 13 and 18 months. Disyllabic words containing singleton and geminate consonants were extracted from the data for the CDS sample. These words were then elicited from the mothers through picture description and question and answer tasks, in order to produce the ADS sample. Acoustic analysis revealed high inter- and intra-speaker variability in the durational differences between singleton and geminate consonants. There was no significant difference in durations as a function of speech condition.

Khattab argues that these results do not support the notion that CDS is simplified in terms of increased consistency and phonological contrast. This conclusion agrees with that drawn by Shockey and Bond (1980), but contrasts sharply with the results of Foulkes et al. (2005). However, in contrast with the studies of Foulkes et al. and Shockey and Bond, Khattab further concludes that the phonological characteristics of CDS and ADS are similar. It is noteworthy that the children in Khattab's study were much younger than the children in the other two studies. It may be that mothers do not emphasise standard phonological contrasts in their CDS until a point when children are ready to learn such contrasts. The main methodological problem with this study is that the ADS condition was engineered by engaging the mothers in picture description and question and answer tasks. These are formal experimental tasks,

compared with the spontaneous speech elicited in natural conversation, which constituted the ADS condition in previous studies.

In an earlier study with a different design, the possible influence of maternal input on inter-speaker phonological variability among children was investigated (Vihman, Kay, De Boysson Bardies, Durand and Sundberg, 1994). They conducted a cross-linguistic study of five mother-child dyads in each of three language groups: English, French and Swedish. As part of a long-term project, six half-hour sessions of spontaneous interaction were recorded for each mother-child dyad. Data collection began at the prelinguistic phase and ended when each child was estimated to have a vocabulary of 50 words. For the purposes of this study, one half-hour session was selected for phonological and statistical analysis, when the children were twelve or 13 months of age. The proportions and distributions of four parameters were analysed in both the mothers' and children's speech: place of consonant articulation, manner of consonant articulation, length of words (number of syllables) and frequency of word-final consonants.

In their study of CDS, Vihman et al. (1994) found that for most measures, the children's speech showed significantly higher variability than the mothers' speech. The only exception was for word length in French. Further statistical comparisons of individual mother-child dyads showed little closeness of fit, meaning that mother-child dyads did not correspond well in terms of the extent of their phonological variability. They concluded that high individual variability exists between children learning the same language, and that variability in the children's word shapes is greater than the variability in the phonetics of their mothers' speech. The finding that the mothers speech was relatively consistent in terms of phonological characteristics, contrasts with Khattab's (2006) observation of high variability in mothers' speech. However, this apparent contrast may be a reflection of the different phenomena investigated. Vihman et al. studied the distribution of isolated phones and word length, whereas the other studies described above focused on language-specific context-dependent phonological phenomena.

Vihman et al. propose that the high individual variability in children's speech, despite consistent input across mothers' speech, results from a perceptual filtering process. They argue that children filter in patterns from the input which correspond with the motor schemes which they have developed through babbling, while filtering out sounds which do not match these schemes.

In conclusion, there have been very few studies on the phonological characteristics of CDS. The studies which exist have employed different methodologies and yielded contrasting results. The phonetic and phonological patterns of consonants in CDS are therefore much less clearly established than vowel patterns. Although these issues make it impossible to draw definite conclusions, these findings taken together indicate the possibility of a pattern which has not previously been considered. It may be that when children are prelinguistic or at the earliest stages of language acquisition, their caregivers produce speech which has the same phonetic and phonological variability as ADS. This would explain Khattab's finding of high variability across the CDS and ADS conditions in their study of CDS to younger children (aged 13 to 18 months). However, a change may then occur at a later stage of language acquisition, when caregivers are unconsciously maximising phonological contrasts and using more standard forms, in order to facilitate learning. This would explain the higher proportions of standard variants than vernacular variants observed by Foulkes et al. (2005) in mothers of two-year-olds. Their further observation that mothers' usage of vernacular variants increases as a function of age then appears to indicate that mothers revert back to speech with more phonological reduction and overlap between contrasts, because four-year-old children no longer require the same level of simplification. In summary, it is tentatively concluded that caregivers provide phonological simplification only at the stage when the child is receptive to it and reduce the simplification of their input when it is no longer necessary.

3.10. Summary and Conclusions

This chapter has explored typical phonological development, recent research on CSP development and the phonological characteristics of child-directed speech. There exist many diverse theories which seek to account for patterns of phonological development. Theories with a more phonological focus, such as *Natural Phonology,* have made valuable contributions to current understanding of phonological phenomena occurring at the level of the single word. However, they do not account for phenomena specific to multi-word speech. In contrast, phonetic and linguistic approaches, such as *Articulatory Phonology* and usage-based approaches, are better able to account for these phenomena.

The concepts of word junctures and CSPs which are central to the current study have been discussed. Evidence for the coarticulatory versus phonological nature of CSPs has been explored, with an especial focus on assimilation, which is the focus of the current study. It appears that assimilation results from a combination of coarticulatory and phonological phenomena. However, the exact role of each is not clearly understood and there exists evidence of high inter-speaker and intra-speaker variability.

Developmental patterns of CSPs have been discussed in detail. Assimilation, elision and liaison of [j] and [w] are acquired much earlier than liaison of /r/. It has therefore been concluded that /r/ liaison is a learned, phonological phenomenon. However, the extent to which the other CSPs under investigation are phonetic, coarticulatory phenomena or learned phonological behaviours continues to be debated. There is also evidence that patterns of CSP development are complex and there exist interactions with other linguistic levels, such as syntax. This discussion has set the scene for the current study, in which these interactions are explored in further detail.

The final section focused on the phonological characteristics of child-directed speech, in preparation for the current investigation of the role of maternal input in assimilation

development. There exist very few studies of consonant patterns in CDS, and the results of these studies appear to be contradictory. Some research has shown that parents facilitate their children's phonological development by maximising phonological contrasts, whereas other studies have shown quite the opposite pattern of increased phonological reduction in mothers' speech to younger children. It is noted here that the methodologies used in these studies have been diverse and sometimes confounded, rendering comparison between the studies difficult. However, the comparison which has been possible has yielded a pattern not previously suggested. It is proposed here that mothers adapt to their children's level of phonological ability, by maximising phonological contrasts at the stage when the child is acquiring these contrasts. They otherwise continue to produce the same variability and reduction as found in ADS.

Chapter Four

Impressionistic Phonetic Transcription as a Methodological Tool

4.1. Introduction

Perceptual phonetic transcription is the main method of analysis employed in the current study, in order to investigate the development of assimilation. This chapter examines the strengths and weaknesses of transcription in detail, in order to gain a clear understanding of the potential biases inherent in this method. Some of these biases may be favourable and improve the reliability and validity of research. However, other biases are potential confounds of which the researcher should remain aware. This chapter begins with a summary of the arguments in favour of and against transcription. The potential biases which can occur in the processes of recording data, listening to data and transcribing data are then discussed. The final section is a discussion of methods used to ensure the reliability and validity of transcription data as far as possible, especially methods of establishing transcription agreement.

4.2. What is Perceptual Phonetic Transcription?

Impressionistic phonetic transcription is a methodological tool which aims to analyse speech in as much phonetic detail as is perceptible to the hearer. Data derived from transcription form the basis of much phonetic research as well as the clinical decisions made in speech and language therapy. It is considered by many to be a valuable tool, because it enables analysis of speech with reference to the natural context of communication. However, some argue that

its value is limited by human perceptual subjectivity and that objective speech measures are therefore more reliable and valid. The following discussion explores the debates concerning the value of transcription. Attempts to develop methods for increasing the reliability and validity of transcription are summarised. Implications of this research are discussed in relation to the method for the current investigation of assimilation development, in which phonetic transcription has been extensively employed.

4.3. Arguments For and Against Transcription

Some experimental phoneticians do not consider phonetic transcription to be a valid method of speech analysis because of its subjectivity. They argue that human perception is not able to detect all of the physical characteristics of the speech signal (such as exact durations and frequencies) and that perceptual phonetic transcription therefore provides an incomplete record of the physical event. They therefore advocate objective instrumental (acoustic and articulatory) measures as the only valid means of obtaining an accurate record of a speech event (Heselwood, 2009; Shriberg and Lof, 1991).

In response to this objection, researchers in favour of perceptual phonetic transcription do not dispute the inherently subjective nature of perceptual analysis, which results from the experience of hearing. They acknowledge that there are unavoidable sources of error arising from the biological and cognitive limitations of perception which may affect transcription accuracy. Heselwood and Howard (2008) acknowledge that to produce a perfect transcription would be extremely labour-intensive and is probably impossible, but a transcription need not be final and can be changed. Kent (1996) summarises the types of perceptual errors which occur in everyday communication including illusions, lapses in perception and uncertainty as to what has been heard. As a result, listeners inevitably perceive aspects which are in fact absent from the signal and fail to hear aspects which are present. Researchers who employ impressionistic transcription advocate this method on the grounds that it is influenced by the

same subjectivity and human perceptual limitations as is natural communication (Heselwood, 2009; Heselwood and Howard, 2008; Howard and Heselwood, 2002; Kent, 1996). It is only by listening to a speaker that phonetic phenomena associated with communicative efficacy (such as intelligibility) can be determined. A transcription can lead to research hypotheses and further investigation of phonetic phenomena (Howard and Heselwood, 2002). Kent (1996) also points out the practical advantages of perceptual analysis above instrumentation in field work, in that it is convenient and economical.

Proponents of transcription argue that spending sufficient time in completing a clear, detailed transcription increases overall efficiency by enabling the researcher or clinician to identify phonologically relevant phonetic phenomena. The highest possible level of detail is needed from the outset because the most relevant phonetic phenomena may not be clear until attempts are made to interpret the completed transcription (Crystal, 1984; Heselwood and Howard, 2008; Kelly and Local, 1989; Local and Walker, 2005; Perkins and Howard, 1995).

Researchers in favour of phonetic transcription acknowledge that instrumental measures can be used to complement and validate perceptual analysis and to increase phonetic understanding by enabling in-depth quantitative investigation of specific phenomena (Heselwood, 2009; Heselwood and Howard, 2008; Howard and Heselwood, 2002; Perkins and Howard, 1995). While Howard and Heselwood (2002) acknowledge that imperceptible acoustic features may be visible on a spectrogram, they argue that such aspects cannot possibly serve any communicative function. Heselwood (2009) also discusses the opposite phenomenon occurring when perceptible aspects are not visible from a spectrogram. He therefore advocates using perceptual analysis in conjunction with instrumental measures to explore possible cause and effect relationships between measurable parameters and audible phenomena.

The arguments above demonstrate how impressionistic phonetic transcription can be a valuable analytical tool in both research and clinical work, providing insights into the phonetic

phenomena occurring in typical and atypical speech. However, there are inherent biases and pitfalls associated with this method owing to human biological and cognitive limitations. Transcribers should therefore have a high level of phonetic knowledge and should keep an open mind, trying not to impose phonological categories on the speech they hear, but instead paying attention to phonetic subtleties which may be of research or clinical interest. They should be aware of the potential confounds affecting reliability and validity and take measures to avoid these wherever possible (Heselwood and Howard, 2008; Kent, 1996). Some potential sources of bias and attempts to improve reliability and validity are now discussed.

4.4. Audio and Video Recording

Research has shown that transcriptions are more accurate and reliable when produced from a recording rather than in a live context (Amorosa, Vonbenda, Wagner, and Keck, 1985). When phoneticians completed live transcriptions of a child with a phonological impairment, they found the task challenging owing to time pressure, and were more likely to transcribe expected adult forms than the atypical forms which actually occurred. In contrast, when they were allowed to listen to recorded speech multiple times and to confer, their transcription accuracy and inter-rater agreement increased. An expert transcription and spectrographic data were used to validate transcription accuracy.

These findings have greatly impacted on phonetic research. Many phoneticians today acknowledge that live transcription is unreliable, and that it is important to obtain a high quality recording in order to increase accuracy and reliability and to minimise subjectivity (Heselwood and Howard, 2008; Howard and Heselwood, 2002; Ladefoged, 2003). Obtaining a high quality recording posed a great challenge until recent years, because analogue cassettes were the best available method of recording and they could not capture the highest frequency sounds, such as those occurring in fricatives. The transcripts resulting from such recordings therefore lacked accuracy and reliability (Stephens and Daniloff, 1977). However,

technological advances over the last decade mean that it is now feasible to make high quality digital recordings which can capture frequencies of up to 11000 Hz, which are beyond the bounds of human hearing (Ladefoged, 2003). It is also possible to attach portable microphones to the speaker's clothing or to station radio microphones at various points in the room or house where the data are being collected. This means that the researcher can ensure that the speaker is close to a microphone at all times in accordance with the recommendations of Ladefoged (2003).

An advantage of using video recording in addition to audio recording is that it is possible to detect visual cues to articulatory activities which cannot be detected from the audio data alone. These may include silent articulation or the distinction between bidental and interdental fricatives. Such cues bring the transcription process closer to that of live transcription, in which the observer has the added advantage of observing the speakers lip and facial movements (Heselwood and Howard, 2008; Kelly and Local, 1989; Stephens and Daniloff, 1977). However, McGurk and MacDonald (1976) found that visual information can equally lead to unhelpful biases as well as helpful cues. They demonstrated that listeners can assimilate information from conflicting auditory and visual stimuli to create a percept which differs from both stimuli. When participants watched a film containing repeated lip movements for the syllable *ga* superimposed onto repetitions of the spoken syllable *ba*, they heard *da*. When the lip movements and auditory presentations were reversed, participants reported hearing either *bagba* or *gaba*. The participants perceived the auditory presentations accurately when the visual stimuli were either congruent with the auditory stimuli or when they were absent (McGurk and McDonald, 1976). These findings show that visual cues play a role in overall speech perception and may lead to perceptual inaccuracy. It is concluded from these mixed findings that visual cues to speech perception may have a positive or negative effect on perceptual accuracy and may therefore be equally advantageous or disadvantageous to the transcriber.

4.5. Listening to the Data

There exist aspects of different listening conditions which may also impact on transcription accuracy. Hewlett (1985) distinguishes between speaker-oriented transcriptions, which focus on the speaker's articulatory activity and listener-oriented transcriptions which focus on the listener's perceptions. Heselwood and Howard (2008) argue that the appropriate technological methods used to listen depend on whether the researcher takes a speaker- or listener-oriented approach. They suggest that if a listener-oriented approach is to be taken, then it is appropriate to listen to the speech at normal speed without any technological modification to the playback. In addition, the transcriber can employ their auditory feedback loop by mimicking the sounds heard in order to gain insight into the possible articulatory gestures occurring. On the other hand, they argue that if a speaker-oriented approach is to be taken, then all available technological means should be used to gain maximum insight into articulatory activity, including reducing the playback speed and playing certain portions in reverse as suggested by Ladefoged (2003).

There is also debate on how many times a phonetician should listen to an utterance when transcribing. Shriberg, Kwiatkowski and Hoffman (1984) argue that an utterance should only be heard three times, in order to avoid the auditory illusions which can occur after listening multiple times. One example of such an illusion is the verbal transformation effect. This occurs when repeated playback of the same stimulus results in the listener hearing a change in phonetic pattern, which may result in activation of a different phonological representation and consequent perception of a different word (MacKay, Wulf, Ying, and Abrams, 1993). In contrast, Ashby, Maidment and Abberton (1996) advocate an analytic listening approach which involves listening multiple times, but focusing attention on different phonetic aspects each time. Similarly, Amorosa et al. (1985) found that transcription accuracy and inter-rater reliability increased when participants were allowed to listen to a recording multiple times and to confer, compared with when they heard live speech only once (see above for a more detailed discussion). In reviews of this literature, Heselwood and Howard (2008) caution

against listening too many times while focusing on the same aspect in order to avoid unnecessarily biasing the transcription.

Munson and Brinkman (2004) disagree with the cautions against listening too many times, concluding from their experiment that multiple presentations of speech stimuli do not affect transcription reliability. Transcribers completed two transcription sessions a week apart, in which they were asked to transcribe the child's forms of /s/ in a number of words. Some of the words were presented only once, while the others were presented seven times. The researchers compared the similarity of transcriptions across participants and conditions and found that there was no significant effect of multiple presentations either on inter-rater reliability or intra-rater reliability. In a second experiment, they manipulated the variable of accuracy by artificially modifying the acoustic characteristics of /s/ in words spoken by a typical speaker. The result of this modification was some correct tokens, some partially correct tokens and some incorrect tokens. Participants were not required to transcribe, but instead to make binary judgements of whether the /s/ was correct or incorrect. There was no significant effect of presentation condition on either judgement accuracy or intra-rater reliability. However, inter-rater reliability improved with multiple presentations, contrary to the predictions of Kent (1996) and Shriberg et al. (1984). They also found that accurate productions were identified more consistently than either partially correct or incorrect judgements, indicating a perceptual bias towards accurate forms over inaccurate productions. They point out that this reduced ability to accurately detect incorrect forms may have clinical applications (Munson and Brinkman, 2004). This finding corresponds with that of Amorosa et al. (1985), who also found a tendency to normalise transcriptions to those of the expected adult form when transcribing atypical speech.

Several aspects of Munson and Brinkman's study make it difficult to interpret their findings. Firstly, the experiments involved judging only /s/ production, which is an artificial, isolated task unlike the multiple tasks involved in real life phonetic transcription. Although it may be that the researcher or clinician is concerned with only specific phonetic phenomena, they

would still be likely to transcribe whole words or even multi-word utterances, rather than an isolated sound. Also, the attempt to simulate impaired speech by artificially manipulating the acoustic characteristics of typical speech may lack ecological validity. Therefore, owing to the artificial nature of some of the stimuli and transcription methods used, the researchers' claim that multiple presentations do not generally influence transcription reliability should be viewed with caution.

As mentioned above, there are a number of listener confounds resulting from limited perceptual and processing capacities which may affect the accuracy of transcription. These are effects of the speaker's top-down processing from higher level cognition on bottom-up perception. One factor is whether or not the listener knows the intended utterance in advance of completing the transcription. Heselwood and Howard (2008) point out that having an orthographic gloss of the utterances to be transcribed is disadvantageous for completing phonetic transcription, but conversely essential for phonological analysis. When conducting phonological analysis, a gloss serves as a template on which to judge the speaker's intelligibility and accuracy (Grunwell, 1987).

In an investigation of the role of prior expectation on transcription accuracy, Oller and Eilers (1975) compared transcriptions of a phonologically delayed child when participants knew the meaning of the intended utterance and when they had to guess. They found that when the participants knew the intended meaning of the utterance, their transcriptions conformed more closely to expected adult forms than when they did not. These finding suggest that knowledge of the intended utterances may activate the listener's own phonological representations of the word. This may interfere with bottom-up perception, thus biasing the transcription. However, in a second experiment they found that knowledge of utterance meaning could also improve transcription accuracy. They concluded that expectation of utterance meaning (whether by knowing or guessing the meaning) could equally give rise to phonetic expectations which affect transcription accuracy and validity. In some cases, this expectation bias appeared to lead to perceptual errors and omissions, but in other cases, it appeared to facilitate the

avoidance of these errors, resulting in more accurate transcriptions. Highly skilled and less skilled transcribers were also compared in this study. The authors concluded that phonetic expectation biases even those transcriptions produced by highly skilled phoneticians. They argue that expectation may have a positive effect by reducing the listener's cognitive processing load and consequently the magnitude of the task, but also adversely affects transcription accuracy by limiting bottom-up perception.

A similar issue which is also linked to incorrect activation of the listener's phonological representations is *phonemic false evaluation* (Buckingham and Yule, 1987). This occurs when phonetic deviation from the expected adult form of a segment causes the listener to perceive it as a contrasting phonological unit. The sound which the speaker produces is not necessarily so different from the intended sound as to create a phonological contrast, but deviant phonetic features of the incorrect sound cause the listener to interpret it as a phonologically contrasting one. This in turn leads to word level misinterpretation, a phenomenon termed *lexical identification shift*, which can lead to miscommunication (Kent, 1996; Pitt, 1995). The opposite problem may also occur when a transcriber perceives speech errors which are phonologically contrastive, but fails to perceive equally relevant, but more subtle, non-contrastive phonetic differences which may be attempts at phonological contrasts on the speaker's part (Cucchiarini, 1996; Hewlett, 1985).

A further top-down perceptual bias is the ability to perceive segments which are not present in the signal, a phenomenon known as the *phonemic restoration effect*. Warren and Obusek (1971) found that when participants were presented with sentences in which the first segment of a word had been removed to produce *eel*, the participants claimed that they had heard *meal*, *wheel*, *heel* or *peel* according to the semantics of the sentence. This finding shows that humans can perceive elements which are absent and that this may be a further top-down bias affecting transcription accuracy (Warren and Obusek, 1971).

Shriberg and Lof (1991) discuss anecdotal evidence of *observer drift*, a phenomenon occurring throughout the course of a transcription exercise in which the transcriber develops a more or less stringent perceptual standard for identifying a specific phenomenon. For example, if a child always produces /s/ with a dentalised alveolar place of articulation, the transcriber may come to perceive this as purely alveolar and cease to notice the dentalisation over time.

Kent (1996) points out that top-down processing effects on speech perception are not always disadvantageous, but can be desirable, for instance when communicating in adverse listening conditions. He argues that these positive effects may be as relevant to transcription as to communicative situations. This argument combined with the potentially advantageous effects of phonetic expectation reported by Oller and Eilers (1975) and the potentially disadvantageous effect of visual cues reported by McGurk and MacDonald (1976) demonstrates that top-down processing may have either positive or negative effects on transcription accuracy.

4.6. Transcribing the Data

A phonetic transcription is not in itself the raw speech data, but an abstracted record of the listener's subjective experience of hearing it (Cucchiarini, 1996; Heselwood and Howard, 2008). The transcription which is produced is underpinned by phonetic theory, which enables the transcriber to associate the sounds heard with specific articulatory activities. Although transcriptions are derived from the auditory percepts of sounds, they are classified according to articulatory categories in transcription systems such as the IPA (International Phonetic Association, 2008). However, the association of sounds with articulatory gestures can be problematic, because there is not a simple one to one correspondence between articulatory gestures and the sounds perceived. A similar sound may result from a number of different articulatory activities, while conversely, similar articulatory gestures may produce very different sounds. A further problem is that the process of transcription using the IPA requires

the transcriber to analyse the speech which they hear into discrete segments, each of which is represented by an IPA symbol. Cucchiarini (1996) argues that this process perpetuates the notion that speech is made up of a linear sequence of segments, when it is in fact a much more continuous stream of sound. Kelly and Local (1989) argue that transcribing sounds according to the categories provided in the IPA places constraints on listening. As a result, the transcriber may either fail to notice or ignore phenomena for which there are no symbols. Furthermore, the IPA does not provide a systematic means of transcribing variability in the degree of a specific aspect (Kelly and Local, 1989). For instance, all tokens of /t/ may be dentalised and would therefore be marked with the relevant diacritic, but some realisations may be more dentalised than others and the IPA provides no guidance on how to indicate this in transcription. Kelly and Local (1989) argue that this information may be crucial in understanding a speaker's phonological system. One aspect of more recent IPA notation conventions which has led to a greater degree of transcription accuracy is the introduction of indeterminacy symbols (Crystal, 1984). These can be used when a transcriber cannot clearly identify a sound, but can perceive some of its characteristics. For instance, in the extensions to the IPA (ExtIPA) (Duckworth et al., 1990), (C,Vls), indicates that a voiceless consonant is perceived, but that the manner of articulation is indeterminate.

4.7. Transcribing Spontaneous Speech

The current study involved the transcription of spontaneous connected speech. This method enables the researcher to observe all of the CSPs and interactional phenomena occurring in natural speech, which are not necessarily observable from the elicited speech resulting from more traditional phonological assessments (Howard et al., 2008; Local and Walker, 2005). Local and Walker (2005) set out specific guidelines for the phonetic and phonological analysis of spontaneous speech. They argue that only data obtained from talk-in-interaction should be used, in other words, speech which occurs in spontaneous conversation, without performing a specific function or adhering to any script.

Local and Walker (2005) argue that the interactional context of an utterance should be analysed alongside phonetic analysis, because the two are inextricably linked. Interpretations of the communicative functions of specific phonetic phenomena should be evidence-based and data-driven, rather than based on speculation which is far removed from the data. It is important to compare different instances of the same phonetic and phonological phenomena. Each instance should also be considered in the context of syntactic structure and turn sequences (Local and Walker, 2005).

The data should be listened to sufficiently to ensure that any claims concerning phonetic phenomena are based on what can be heard, rather than on expectation. This issue is addressed more fully above, in the discussion on listening. Local and Walker (2005) also emphasise the importance of conducting both quantitative and qualitative analyses.

4.8. Reliability and Validity in Phonetic Transcription

It is clear from the discussions above that transcription is a valuable research tool, but that it can potentially be confounded by a number of factors including recording quality, listening conditions, speaker characteristics and listener biases arising from higher level cognitive processing. It is therefore important to use scientific methods to ensure as far as possible that the data obtained from transcription are both reliable and valid. The subjective nature of perceptual analysis means that this task presents many challenges. Research in clinical phonetics has been criticised on the grounds that researchers have either done little to ensure reliability and validity or have not explicitly reported the validation measures which they have taken (Cordes, 1994; Cucchiarini, 1996; Pye, Wilcox, and Siren, 1988; Shriberg and Lof, 1991). The following discussion outlines some methodological difficulties in ensuring the reliability and validity of phonetic transcription and some approaches taken to resolve these difficulties.

4.8.1. Defining Reliability and Validity

In a discussion of the challenges in establishing reliability and validity in clinical phonetic analysis, Cordes (1994) acknowledges that it can be difficult to separate the two in practice. Reliability is a statistical measure, which concerns consistency and replicability of data using the same measuring instrument (Cordes, 1994; Cucchiarini, 1996). Data are reliable if the values obtained are not confounded by aspects other than the phenomena being measured, such as poor quality recording in the case of speech data (Cordes, 1994). Cordes argues that it is difficult to establish reliability of phonetic transcription in the true statistical sense, because there is no statistical score against which to measure the observed behaviour and no way of correlating exact scores for comparison.

In contrast, validity concerns whether the measuring instrument used measures the phenomena it purports to measure (Cordes, 1994). In terms of transcription, this involves determining whether the phonetic phenomena perceived can be related to acoustic measures and articulatory activity (Heselwood, 2009; Heselwood and Howard, 2008; Shriberg and Lof, 1991). Cordes (1994) illustrates the difference between reliability and validity in transcription with the hypothetical scenario that two transcribers may make the same phonetic judgements, meaning that they agree and that there exists inter-rater reliability, but that these may be inaccurate and therefore not valid. Similarly, Cucchiarini (1996) argues that although researchers in many phonetic studies routinely report reliability indices of over 80%, this reflects only reliability and does not necessarily reflect accuracy and validity.

4.8.2. Transcription Agreement

A method commonly used to assess reliability in phonetic research is to compare the extent of agreement across several transcriptions of the same data. The transcriptions to be compared may be produced by different transcribers analysing the same data (inter-rater agreement) or by the same transcriber analysing the data on two different occasions (intra-rater agreement)

(Cordes, 1994; Shriberg et al., 1984). Cucchiarini (1996) argues that transcription agreement cannot be equated with reliability in the true statistical sense, which involves repeatedly taking the same measurements with the same instrument in the collection of interval data. Instead, she suggests that the term agreement should be used in relation to transcription, since this involves equal ratings of the same measurement and can be applied to nominal data, such as the values given to phonetic symbols. Several factors have been found to influence transcription agreement. Nasals, glides and plosives are more often associated with high agreement, whereas fricatives, affricates and liquids are associated with lower agreement in both inter-rater and intra-rater comparisons (Shriberg and Lof, 1991). Shriberg and Lof (1991) argue that when assessing transcription agreement, it is insufficient simply to assess a small proportion of tokens from the analysis data for inter-rater or intra-rater reliability. Instead, they encourage researchers to provide multiple sources of evidence for any claims of reliability and validity.

Percentage agreement indices are commonly used to calculate transcription agreement, which involve point-by-point comparison of several transcriptions and a final calculation of their overall correspondence (Cordes, 1994; Cucchiarini, 1996). String alignment is a computer-based method for calculating this, which involves a process of matching between transcriptions. However, this method would only be accurate if the transcriptions being compared contained exactly the same number of equivalent symbols (Cucchiarini, 1996).

A further problem which Cucchiarini (1996) identifies with such measures is that disagreements are equally weighted, giving rise to discrete classifications of total agreement or total disagreement. She argues that measures of agreement should be continuous, to account for the fact that some disagreements are great, while others are relatively small. For instance, /p/ disagrees with /z/ in terms of voicing, place and manner, whereas the disagreement between /p/ and /b/ is only in voicing and is thus smaller. Cucchiarini (1996) suggests that a more accurate solution is to assign numerical values to distinctive phonetic features and then to calculate the dissimilarity between equivalent segments across transcriptions. In this way,

the extent of disagreement can be ranked as ordinal data. In an attempt to overcome the issues of equal weighting and the non-equivalence involved in string alignment, Cucchiarini (1996) devised computerised feature matrices. These aimed to provide criteria for aligning transcription pairs, and to calculate dissimilarity indices between segments in terms of distinctive features.

Similarly, Oller and Ramsdell (2006) offered an alternative to percentage agreement indices in order to avoid the problem of giving equal weighting to disagreements. Transcriptions of infant vocalisations were aligned using a computer program and the differences between them were calculated using numerical values. If two equivalent segments shared all distinctive features, then a value of 1 was given, whereas the value 0 was given when no features were shared. The individual features were assigned numerical values, which were subtracted from 1 as specific features were found to differ. The values were such that incremental differences in agreement could be expressed. For instance, disagreement between obstruents and sonorants was expressed as numerically greater than that occurring between different obstruents. Oller and Ramsdell (2006) conclude similarly to Cucchiarini (1996) that computer-based methods of transcription alignments and numerical ranking of disagreement types may lead to more accurate measures of transcription agreement.

4.8.3. Inter-Rater Reliability

As discussed above, the best practical method for investigating inter-rater reliability of transcription data has involved assessing the extent of agreement between transcriptions completed by different phoneticians listening to the same data (Cordes, 1994). This process can be extremely challenging, because different transcribers may interpret the same data very differently (Cordes, 1994). There may not be a simple distinction between right and wrong when comparing transcriptions, as different interpretations of the data may equally be valuable and insightful (Ashby, Maidment and Abberton, 1996). Because transcription is qualitative

and subjective, it is challenging to define the dimensions and units to be agreed upon (Kent, 1996; Shriberg and Lof, 1991). Two judges may perceive the same phenomenon, but may define it differently (Kent, 1996). For instance, one person may transcribe a devoiced [b], while another transcribes a [p] with voicing, in which case, the symbols used are different, but the two phoneticians' percepts may be very similar.

A solution to the difficulty of achieving inter-rater agreement is for each transcriber to complete an independent transcription and then for all of the transcribers together to produce a consensus transcription which reflects the transcribers' combined judgements. In this way, the exact qualitative and quantitative aspects of any disagreements can be made explicit (Cordes, 1994; Shriberg et al., 1984). Shriberg et al. (1984) recommend that at least two and preferably three transcribers be involved in the consensus transcription process. They argue that consensus transcription serves to minimise the biases introduced by an individual transcriber which may result from misperception or inattentiveness. They advocate the use of rigorous procedures in consensus transcription in order to avoid confounds introduced by individual transcribers. For instance, they suggest that the transcribers involved in consensus may not otherwise have equal influence on the final transcription, owing to personal factors such as competence, professional status and personality. They also suggest that having specific criteria and rules for reaching consensus helps to avoid discarding data on which consensus cannot be reached. They argue that it is precisely this more challenging data which may hold key information regarding a person's speech processes and any speech impairment. They propose that if there is disagreement regarding a segment, then the segment should be played a maximum of three times and each transcriber should attempt both to hear the other person's interpretation and to confirm their own. In a test of their proposed transcription procedure, they found that this process often resulted in immediate consensus as one transcriber was able to hear the other's interpretation, which was usually a salient feature to which they had not initially attended. They also suggest that a segment should be judged as incorrect unless proven to match the intended segment.

Pye et al. (1988) also advocate consensus transcription, although they found that the presence of a third transcriber led to greater disagreement in search of increased accuracy compared with only two transcribers. They summarise variations on the consensus method used in different clinical studies. These include:-

a. Two independent transcribers with agreement resolved by consensus;

b. One independent transcriber with a proportion of utterances transcribed by a second transcriber;

c. Two transcriptions completed live and two from recordings, with disagreements either resolved by consensus or excluded from the data when agreement could not be reached;

d. Two independent transcribers with a third transcriber deciding on aspects of disagreement.

Pye et al. (1988) argue that a difficulty with the composite transcriptions which arise from consensus methods is that they capture only those aspects which have been agreed upon; individual insights which are not agreed upon, but which may be equally informative are lost. They argue that if individual transcriptions differ greatly from the composite transcription, then there is no reason to believe that the composite transcription is anymore objective or accurate than the original transcriptions from which it was derived.

They demonstrated the potential pitfalls of consensus procedures in a study in which three phoneticians transcribed the highly unintelligible speech of a hearing impaired child. A composite transcription was then created from the three independent transcriptions. When two or more transcribers agreed on a segment, it was included in the composite transcription. When all transcribers disagreed on a segment, a compromise segment was chosen which shared the most phonetic properties with those occurring in the independent transcriptions. When this was not possible, all of the possibilities were entered into the composite transcription in brackets. They found that two transcribers seldom used the same diacritics in transcribing a specific segment, and it was therefore necessary to use broad transcription in order to reach agreement. However, they also found that the diacritics supplied in the original transcriptions were insightful and helped to resolve some disagreements. They argue that

consensus can be used to eliminate inter-transcriber differences, but that compromises in the amount of detail may also result. They conclude from these findings that achieving transcription agreement by consensus can be a worthwhile activity, but that transcribers should remain aware of these pitfalls and the possible influence of other factors, such as the complexity of transcribing connected speech and the reduced agreement added by a third transcriber.

Cordes (1994) similarly concludes that the challenge of establishing inter-rater reliability will always exist in this field owing to human error. However, he suggests that errors can be minimised by ensuring sufficient training of transcribers, clearly defining and operationalising the target behaviours to be observed, and by paying attention to methodological detail. Perkins and Howard (1995) agree that consensus techniques are valuable for obtaining a final transcription on which to base interpretations and theoretical claims, whether this be achieved using formal or informal procedures.

A challenge when assessing inter-rater agreement which was brought to light by Pye et al. (1988) is deciding how narrow and detailed the transcriptions should be. There is often high inter-transcriber agreement between broad, phonological transcriptions, whereas detailed narrow transcriptions which aim to achieve optimum accuracy can produce greater disagreement (Cucchiarini, 1996; Pye et al., 1988; Shriberg and Lof, 1991). If this inevitable reduction in agreement equates with reduced reliability, then this is a strong argument against the value of transcription (Heselwood, 2009). However, Cucchiarini (1996) argues in favour of narrow transcription that it is less likely to be influenced by chance agreement than broad transcription, because there exist more parameters and categories for the classification of sounds. This suggests that the potential accuracy achieved through narrow transcription is worth pursuing, despite its potential for inter-transcriber disagreement.

4.9. Summary and Conclusions

This chapter has discussed the efficacy of impressionistic phonetic transcription as a methodological tool. Firstly, the arguments for and against phonetic transcription were reviewed. Some phoneticians who prefer instrumental methods argue that transcription is invalid, owing to its subjectivity and the imprecision which may result from the limitations of human perception. However, those in favour of transcription argue that it is by nature a suitable technique for investigating those aspects of speech which are relevant to human communication.

However, it is acknowledged that transcribers can unwittingly introduce a number of biases into the transcription process at the levels of recording, listening and transcribing. These potential biases cannot be altogether eliminated, but can be minimised if the researcher remains aware of them and ensures that specific methodological controls are in place. A further issue which is especially challenging is establishing transcription agreement as a measure of reliability. A number of methods are available, including point-by-point transcription agreement, computer-based transcription alignment, consensus transcription, percentage agreement indices and statistical calculation of agreement between independent transcriptions. Each of these methods has associated advantages and challenges.

The research reviewed in this chapter is used to inform the methodology of the current study, in order to optimise accuracy and reliability of the data obtained through transcription. This methodology is presented in full in Chapter Five.

Chapter Five

Aims and Method of the Current Study

5.1. Aims of the Current Study

The current study was inspired by the pilot study of Bryan et al. (2010) described in Chapter Three and was based on the same language corpus of Brian. This study aimed to further investigate the possibility that there may exist interactions between the development of connected speech processes (CSPs) and syntax, as suggested in previous research (Howard et al., 2008; Newton and Wells, 2002; Thompson and Howard, 2007). The developmental trajectory of Thomas's assimilation was investigated in detail and was compared with his pattern of syntactic acquisition.

This research focused on utterances containing the auxiliary verbs *can* and *can't*. There are two reasons why these utterances were selected for analysis. Firstly, their developmental patterns have been established in previous research (Richards, 1990). Secondly, *can* has been found to produce potential assimilation sites prior to main verbs such as *go* and *be*, from age 3;4 onwards; *can't* has been found to contribute to potential assimilation sites from age 3;6 onwards (Bryan et al., 2010A; Bryan, Howard, and Perkins, 2010B; Newton and Wells, 2002). The documentation of *can* and *can't* acquisition in the literature on both grammar development and assimilation development provides a link for the comparison of interactions across these two linguistic domains.

A further aim of the current study was to investigate the ways in which Thomas's mother (M) realised potential assimilation sites, and to relate these findings to the developmental patterns observed for Thomas. The following research questions were posed:-

1. a. What are the relative proportions of assimilations, open junctures and other non-assimilation phenomena at potential assimilation sites with *can* and *can't* in M's speech?

   b. How do these proportions compare across potential bilabial and velar assimilation sites?

   c. How do these findings relate to Thomas's assimilation development at different points in time?

   d. How do these findings relate to the literature on adult assimilation?

2. a. Are there any specific high frequency constructions in M's language which lead to potential assimilation sites?

   b. If so, do these correspond with high frequency constructions in Thomas's language? For instance, does M produce a high proportion of potential bilabial assimilation sites in the constructions *can be*, *can/can't put* and *can play* and a high proportion of potential velar assimilation sites in the constructions *can/can't get*, *can/can't go* and *can keep*?

3. a. Are there observable phonetic matches or mismatches (either segmental or prosodic) across Thomas' and M's productions of the same potential assimilation sites with *can* and *can't*?

   b. Are there observable changes over time?

It was difficult to reach a hypothesis, owing to the contrasting methodologies and results of studies which have investigated the phonological characteristics of child-directed speech (CDS). However, it was considered that the phenomena investigated in the current study were most similar to those explored by Foulkes et al. and by Shockey and Bond. The common feature of the current study and these two previous studies is the focus on language-specific, context-dependent phenomena, which have two phonetic variants in complementary distribution. Based on the findings of the larger-scale study of Foulkes et al. (2005), it was predicted that M would produce more non-assimilated forms at potential assimilation sites when Thomas was two years of age, in order to facilitate his learning of standard phonetic and phonological forms. It was also predicted that M's proportions of assimilations would

increase over time as a function of Thomas's age, until they predominated over other forms. Unfortunately, it was not possible to analyse M's CDS during Thomas's second year, because there exists no data for this period. If mothers' CDS to children younger than age two is not simplified (as suggested by Khattab, 2006), then it would be expected that M would produce more assimilations during his second year than his third year. It was also not possible to compare assimilations in M's CDS with those in her (ADS), because the data corpus consisted almost entirely of conversations between M and Thomas, with only an extremely small proportion of ADS utterances.

5.2. Method

5.2.1. Participant and Data

The participant in this study was one typically developing child, Thomas, from the Dense Database (DDB) compiled by researchers at the University of Manchester and the Max Planck Institute for Evolutionary Anthropology, Leipzig (Lieven et al., 2009). Thomas was the only child of a monolingual English-speaking family in Manchester. His father spoke with a strong regional accent, while his mother displayed some features of a Manchester accent.

The data used in this study comprised naturalistic audio and video recordings of Thomas' spontaneous interactions, primarily with his mother, but also with his father, the research assistants and occasionally with other people, such as neighbours. Thomas's data were collected from the age of 2;0;12 until 4;11;20 (ages given in years;months;days). The first 14 months of this study period (from 2;0;12 to 3;2;12) constituted a period of very intensive data collection, in which Thomas was recorded for approximately one hour for five days in every week. This is referred to in the data corpus as the *very intensive period*. From age 3;3;2 to 4;11;20, Thomas was recorded less intensively for approximately one hour five times per week, but for only one week in every month. This is referred to in the data corpus as the

intensive period. Four out of the five weekly recordings were purely audio, while the fifth was also a video. For the purposes of the current study, Thomas's data were sampled over a two-year period from age 2;0 until 4;0.

The recordings were orthographically transcribed and morphosyntactically analysed by Lieven et al. (2009), using the Codes for the Human Analysis of Transcription (Chat) transcription system (MacWhinney, 2000). Each line of transcription was linked to its corresponding utterance in the audio and video files using Computerized Language Analysis software (CLAN) (MacWhinney, 1996, 2000). This preparation was carried out by Lieven et al. for the purposes of their own research and with the intention of incorporating the Dense Database into the Child Language Data Exchange System (CHILDES) (MacWhinney, 1996, 2000). CHILDES is a large, international database comprising both typical and atypical child and adult language data from many research studies, which allows data sharing and provides a means of comprehensive linguistic analysis using the CLAN programme.

5.2.2. Ethical Considerations

Thomas's data have been made publicly available through CHILDES and Thomas is therefore referred to by his real name in this study. Thomas has also been referred to as Brian in previous research carried out prior to the public release of his data (Lieven et al., 2003; Lieven et al., 2009). Elena Lieven has granted permission for the DDB to be used for the purposes of the current study, subject to the researchers following specific ethical guidelines in relation to the data (see Appendix One). This project has also been ethically approved by the research ethics committee in the Department of Human Communication Sciences, University of Sheffield (see Appendix One).

5.2.3. Data Sampling

All of the recordings for Thomas aged two were sampled and all of the different utterances containing *can* and *can't* were analysed. This high sampling rate was used in order to capture all of the novel contexts in which Thomas produced these verbs, including the first potential assimilation sites and instances of newly acquired syntactic constructions. A further reason for sampling all of the data at this age was because *can* and *can't* initially emerged as low frequency items, meaning that individual recordings maximally contained only small numbers of utterances for analysis. When Thomas began to produce multiple tokens of specific utterances without evidence of phonological or syntactic change (such as "I can't"), these multiple tokens were no longer included in the qualitative analysis. The points at which specific utterances became highly frequent and were eliminated are noted in Appendix Three below the corresponding data tables. However, these multiple tokens were included in the quantitative analyses.

A different approach was adopted from age three until age four, owing to the transition from the very intensive to the intensive period from 3;3 onwards, and the increasing frequency of *can* and *can't* in Thomas's vocabulary. The sampling rate of 5 hours for one week in every month was therefore used from 3;0 onwards, three months earlier than this approach was adopted by Lieven et al. (2009). This is a creative sampling technique, which aims to compromise between the realistic limits on research resources and the need to produce accurate estimates of developmental phenomena (Tomasello and Stahl, 2004). Such a compromise is achieved by reducing the time required for data analysis to that typical of a less dense sample, while also ensuring that individual sampling periods are sufficiently dense to capture linguistic structures of both high and low frequency (Tomasello and Stahl, 2004). Only those tokens of *can* and *can't* occurring at potential assimilation sites were transcribed for four out of the five monthly sessions, in order to reduce transcription time as tokens became increasingly frequent. However, all tokens of *can* and *can't* were transcribed in one

monthly session (usually the third of five recordings), as a control measure of Thomas's productions of *can* and *can't* in non-assimilation contexts.

5.2.4. Phonetic Analysis of Potential Assimilation Sites

The data were analysed using the CLAN program on a Personal Computer. A key word search was conducted for each recording sampled using the *kwal* command in CLAN to search for the string *can**. In response to the *kwal* commands, CLAN automatically generated a Chat file for each recording, containing all of the utterances in which the string *can* was found, along with the relevant Chat transcript line numbers and audio links to the sound files. The symbol * is a wild card in CLAN, which can be placed after a string, in order to detect any words which begin with the string. The search for *can** was therefore not limited to the verb *can*; it also detected instances of *can't* and a minority of other words which were not relevant to the current study, for example, *candle* and the noun *can*.

The utterances to be analysed were played using the audio links in CLAN. Listening took place through high quality headphones in a quiet room, as advised by Ladefoged (2003). Listening was carried out a maximum of three times with the transcriber's attention on any one segment, in accordance with advice given in the literature on listening techniques (Ashby et al., 1996; Shriberg et al., 1984). Each utterance was initially transcribed orthographically. The orthographic transcriptions produced in the current study mostly matched those given in the Chat transcripts, except in instances in which the current author perceived a different utterance from that given in the Chat transcript. If there were two or three different possible interpretations of a word, utterance or part of an utterance, all were entered into the orthographic transcription with a forward slash (/) symbol in-between the different interpretations, in order to indicate that the researcher was in doubt as to the intended utterance, and could not draw a definite conclusion. In instances when the current author identified errors in the Chat transcripts, but was in no doubt as to the intended utterance, then

the revised version was entered into the orthographic transcription. Each utterance involving a discrepancy between the current interpretation and that given in the Chat transcript was marked with an asterisk next to the orthographic transcription. Such discrepancies were detected in twelve utterances sampled at age two and 87 utterances sampled from age three to four. Specific details of each discrepancy involving a completely new interpretation are given in notes at the end of each set of data tables in Appendices Three and Four, including the new interpretation, original interpretation and relevant Chat transcript line number. A minority of utterances were eliminated from the analysis, which were found to contain neither *can* nor *can't*, contrary to information given in the Chat transcripts. Conversely, it is possible that a minority of utterances containing *can* and *can't* were omitted from the analysis, because they were transcribed differently in the Chat transcripts. There was unfortunately no time-efficient means of avoiding this potential source of error.

The focal word *can* or *can't* and the following word in each utterance (where applicable) were impressionistically transcribed according to the conventions of the IPA (International Phonetic Association, 2008) and the indeterminacy conventions of extIPA (Duckworth et al., 1990). It was considered neither practical nor feasible to transcribe each utterance in its entirety, since this would have presented the time-consuming and labour-intensive challenges involved in transcribing long utterances, babble and unintelligible speech (Perkins and Howard, 1995). Transcriptions were therefore mostly confined to two words in order to analyse phonetic behaviours occurring in *can* and *can't*, both at potential assimilation sites and in non-assimilation contexts. There were a minority of exceptions, in which three words were transcribed, either in order to indicate placement of stress, or in instances when *can* or *can't* appeared to merge with another word on a segmental level. The words which were phonetically transcribed in each utterance are underlined in the orthographic transcriptions (see Appendices Three and Four). In addition to the narrow phonetic transcription conducted for each utterance, a short description of the apparent phonetic phenomena was also included. These descriptions enabled categorical classification of the phonetic behaviours occurring at potential assimilation sites. This classification facilitated the calculation of frequency counts

to determine the predominant behaviours occurring at potential bilabial and velar assimilation sites at specific points in time.

The three broad categories which emerged from the analysis were assimilation, other non-assimilation phenomena and open juncture. For the purposes of this study, assimilation specifically refers to regressive place assimilation of the final /n/ in *can* and the final /nt/ cluster in *can't*. Examples of bilabial assimilation include *can be* [kʰæ̃m bi] and *can't put* [kʰɑ̃mʔ pʰʊt]; examples of velar assimilation include *can go* [kʰæ̃ŋ gəʊ] and *can't get* [kʰɑ̃ŋʔ gɛt]. Although there was evidence of other types of assimilation in Thomas's speech, such as progressive alveolar assimilation, these were not included in the assimilation category. However, they are discussed in Chapter Six, in the section on other phonetic phenomena occurring at potential assimilation sites. Other non-assimilation phenomena refers to instances when the final coda /n/ in *can* and the final /t/ or /nt/ cluster in *can't* were not realised. This often gave rise to close juncture, with a smooth articulatory transition between the two abutting words. These specific examples are referred to in Chapter Six as non-assimilation close junctures. Examples include *can be* [kʰæ̃ bi], *can't put* [kʰɑ̃nʔ pʰʊt], [kʰɑ̃n pʰʊt], [kʰɑ̃ pʰʊt], *can go* [kʰæ̃ gəʊ] and *can't get* [kʰɑ̃nʔ gɛt], [kʰɑ̃n gɛt], [kʰɑ̃ gɛt]. However, open juncture also sometimes accompanied final coda consonant elision, for example, when pauses or glottal stops occurred at the word boundary (Wells, 1994). However, for the purposes of this study, these junctures are categorised as other non-assimilation phenomena. In contrast, the open juncture category is preserved exclusively for those junctures at which final /n/ in *can* or /nt/ in *can't* were realised alongside the following abutting word-initial consonant. Examples include *can be* [kʰæ̃n bi], *can't put* [kʰɑ̃nt pʰʊt], *can go* [kʰæ̃n gəʊ] and *can't get* [kʰɑ̃nt gɛt]. This distinction is made in order to emphasise the differences between those word junctures involving final coda consonant or cluster elision (which could occur at both open and close junctures) and those instances in which all consonants in the citation forms were realised.

It was often necessary to refer back to the original audio recordings and Chat transcripts during transcription, in order to hear focal utterances within the context of the interaction in which they occurred. This helped to resolve the problem of unintelligibility which occurred when different speakers' utterances overlapped in time. It was also possible to investigate the long-domain interactional contexts of specific phonetic and linguistic behaviours, as recommended in the literature (Local, 2003; Local and Walker, 2005). A minority of utterances were eliminated from the analysis because they could not be transcribed. There were two main reasons for this: firstly, some utterances could not be accessed owing to corrupted, disrupted or missing sound files; these were marked as untranscribable. Secondly, some utterances were inaudible owing either to conversational overlap, extraneous background noise or because Thomas was too far away from the microphone; these were marked as unclassifiable.

5.2.5. Segmental Phonological Analysis

Thomas's realisations of target alveolar and velar plosives were investigated. This is because a minority of Thomas's potential assimilation sites in early months showed evidence of velar fronting, progressive alveolar assimilation and consonant harmony. For example, "can get" at age 2;8 was realised as [ˈt̪ʰæ̃n d̪ɛt̪] in the utterance "can get it out". Thomas also realised the words *can* and *can't* with variable initial alveolar and velar consonants at age two. This control segmental phonological analysis was therefore necessary, in order to understand the more global developmental context for these phenomena.

A total of eight recordings were selected at three-monthly intervals, starting at age 2;3, when *can* first emerged in Thomas's speech. The remaining recordings were sampled at ages 2;6, 2;9, 3;0, 3;3, 3;6, 3;9 and 4;0. For each recording, up to ten words were identified, which contained target alveolar plosives and target velar plosives in each of three word positions: word-initial, medial and final. This produced a maximum sample of 60 words per recording.

However, earlier recordings yielded a much smaller sample than this, owing to the relatively small size of Thomas's vocabulary in earlier months of the study. Words containing target voiceless plosives were transcribed where possible, to make this analysis most comparable with findings for *can* and *can't*, which both have initial target /k/. However, in instances when there was an insufficient number of words containing voiceless plosives, then words containing target voiced alveolar or velar plosives were used to produce a sample as close as possible to ten words for each word position. Some words were sampled twice, if they contained a target plosive in more than one word position. For example, *cake* has both an initial and final target voiceless velar plosive. Multiple tokens of the same word were not sampled, unless they showed variable realisation of plosives.

5.2.6. Syntactic Analysis

The utterances containing can and can't which were transcribed phonetically were also analysed syntactically, in order to determine qualitative changes in Thomas's syntactic development over time. Examples of qualitative changes observed include increased syntactic complexity and the emergence of novel main verbs and constructions. Quantitative syntactic measures of Thomas's syntactic development were also investigated, including mean length of utterance (MLU) (measured in morphemes) and maximum length of utterance (measured in words). These analyses were carried out using the MLU and Max Word commands in CLAN. Frequency counts for *can* and *can't* were also calculated for each recording using the *freq* command in CLAN, and total frequency counts were calculated for each month of the study.

Reference is made in the results chapters to the relative productivity versus formulaicity of specific utterances. It was beyond the scope of the current study to adopt rigorous methods for quantifying the relative formulaicity of utterances, such as the morpheme matching and operations procedure employed by Lieven et al. (2003) or the computerised *Traceback* method employed by Lieven et al. (2008). Instead, judgements of the relative productivity or

formulaicity of utterances were based on the formulaicity criteria specified by Peters (1983) as evident from the interactional context. Utterances were considered to be formulaic if:-

- The complete utterance form was used repeatedly without alteration;
- The utterance was linked to specific contexts and events;
- The utterance was sometimes used in a contextually incongruous way;
- The utterance was produced without pauses and with a smooth intonation contour;
- There was evidence that the utterance was a formula used by other members of the child's community (Peters, 1983).

It was not necessary for an utterance to meet all of these criteria in order to be considered formulaic. For instance, it was not always possible to confirm whether or not a formula was also used by others in the child's community; this extended to situations such as nursery, where recording did not take place. Peters' second formulaicity criterion was not applied in this study; this states that an utterance is formulaic if it contains grammatical elements not found in the child's productive language. The current study was limited to only those utterances containing *can* and *can't* and a detailed grammatical analysis of other utterance types was not included. It is recognised that Peters' criteria are not sufficiently rigorous to provide quantitative information on the degree of formulaicity or to detect partially analysed utterances. They are used as a general guide here to detect utterances which may be processed as a single unit or lexical item, alongside the more detailed phonetic, phonological and syntactic analyses, which are more focal in the current study.

5.2.7. Investigating the Role of Maternal Input and Interactional Context in Thomas' Assimilation Development

A further study was conducted, in order to examine the possible influence of Thomas's mother's (M's) usage of assimilation on Thomas's assimilation development. Comparisons between Thomas and M were made on two levels. Firstly, Thomas's and M's realisations of

potential assimilation sites were analysed on a global level, independently of local interactional context. Secondly, a more fine-grained analysis was carried out, to compare both the segmental and prosodic aspects of Thomas's and M's potential assimilation sites, in adjacent or near-adjacent pairs of similar utterances.

In order to gain a global picture of M's realisations of potential assimilation sites, five weekly recordings were sampled at each of three different points, which corresponded with each of the three stages identified in Thomas's assimilation development. The first sample of M's potential assimilation sites was collected from recordings of Thomas aged 2;6, following the emergence of both *can* and *can't* in Thomas's language, but prior to the emergence of potential assimilation sites. The second sample was collected from recordings of Thomas aged 3;3, the first point at which establishment of bilabial assimilation was evident and velar assimilation was emerging in constructions with *can* and *can't*. The final sample was collected from recordings of Thomas aged 4;0, a point at which he was no longer producing high numbers of potential assimilation sites with *can* and *can't*. This gave rise to three points in time at which M's speech was sampled: T1, T2 and T3, with intervals of nine months between them. Whereas weekly samples of five recordings were available at both T1 and T2, only four recordings were available at T3. Although this produced a smaller sample of potential assimilation sites at T3, overall proportions of assimilations remained similar to those observed at T2.

All of M's productions of *can* and *can't* occurring at potential assimilation sites within this sample were analysed using impressionistic phonetic transcription and classification of phonetic phenomena, similarly to the method applied in the main study of Thomas's assimilation development. Each phonetic realisation was classified according to one of the three broad categories which emerged during the analysis of Thomas's data: assimilation, open juncture and other non-assimilation phenomena. Relative proportions of these behaviours were calculated at each point in time. The relationship between Thomas's pattern of assimilation development and M's usage of assimilation was then qualitatively analysed.

One aim of this study was to examine M's usage of assimilation in relation to the literature on adult assimilation, in order to investigate whether M's assimilation patterns are typical of British, English-speaking adults. However, although there exists literature which describes the nature of assimilation in adult speech (Cruttenden, 2001; Shockey, 2003), there does not appear to exist any normative adult literature showing proportions of assimilations relative to other behaviours. M's findings were therefore compared with the norms obtained by Newton and Wells (1999) for typically developing children aged three to seven years. This is the oldest age group in which proportions of assimilations over other forms have been investigated. In order to make this comparison, it was necessary to compare M's overall proportions of assimilations relative to other behaviours, inclusive of both bilabial and velar assimilation in the same analysis. Bilabial and velar assimilation were then analysed separately, as previously carried out for Thomas. The types of constructions leading to potential assimilation sites in M's speech were also compared with those used by Thomas.

The role of interactional context in Thomas's assimilation development was investigated by comparing Thomas's and M's productions of the same potential assimilation sites. Portions of interaction were identified in which both Thomas and M produced the same potential assimilation sites, in adjacent or near-adjacent pairs of utterances, which were either identical or extremely similar. There were only three instances in which Thomas repeated M's potential assimilation sites, compared with 22 instances in which M repeated Thomas's potential assimilation sites. This gave rise to 25 pairs of utterances in which Thomas's and M's realisations of potential assimilation sites could be directly compared.

The prosodic characteristics of each utterance within these pairs were analysed. In addition, M's potential assimilation sites with *can* and *can't* were analysed using segmental impressionistic phonetic transcription, as previously carried out for Thomas (see Chapter Four). The phonetic behaviours occurring were then classified either as assimilation, open juncture or other non-assimilation phenomena. The two utterances within each adjacent or near-adjacent pair were then compared, to determine whether they matched on four levels:

segmental realisation of the potential assimilation sites, overall stress pattern, locus of the tonic syllable and nuclear tone.

5.2.8. A Test of Inter-Rater Transcription Agreement

The results of the current study were obtained using impressionistic phonetic transcription, which is subject to a number of biases (see Chapter Three). It was therefore necessary to assess the reliability of the current author's transcription, using a test of inter-rater agreement. The term *agreement* is used here in preference to *inter-rater reliability*, because Cucchiarini (1996) promotes it as more appropriate in relation to the categorical data obtained from phonetic transcription.

Point-by-point comparison of the two transcriptions was not considered to be a suitable method, owing to the variation likely to occur between the two transcriptions; this results from different interpretations of the data and different usage of symbols in the transcription of similarly perceived phenomena (Ashby et al., 1996; Kent, 1996). Moreover, the narrow phonetic transcription method used to capture the maximum possible detail in the current study is likely to produce more disagreement in inter-rater agreement testing than broad transcription, although the latter captures considerably less detail (Pye et al., 1988). There also exist methodological difficulties with calculating percentage agreement indices from point-by-point comparisons; firstly, independent transcriptions do not necessarily have equal numbers of equivalent symbols, making the matching process difficult. Secondly, this methodology assumes equal weighting of all points of disagreement, without taking account of the extent of individual discrepancies in terms of the phonetic parameters in common and those which differ across the independent transcriptions.

An alternative to point-by-point transcription agreement is consensus transcription (Shriberg et al., 1984). This method has several disadvantages; firstly, striving for agreement may result in

loss of detail from the insights of individual transcribers (Pye et al., 1988). Secondly, more weight may be given to some transcribers' interpretations than others owing to transcriber variables including age, gender, status and experience (Shriberg et al. 1984). Nevertheless, consensus procedures continue to be viewed as valuable in enabling researchers to interpret their findings (Perkins and Howard, 1995). An alternative consensus procedure was therefore adopted in the current study (after Newton and Wells, 2002); this aimed to avoid the disagreements likely to result from comparison of independent narrow phonetic transcriptions, while at the same time attempting to establish overall agreement concerning the phonetic phenomena under investigation.

A total of 313 potential assimilation sites were narrowly transcribed by the current author. A selection of 30 (9.4%) of the sites was sampled for the consensus procedure. This sample comprised ten utterances categorised by the current author into each of the three broad categories emerging from the analysis: assimilation, open juncture and other non-assimilation phenomena. The sample was also carefully selected to include examples of each category at both potential bilabial and velar assimilation sites and in constructions with both *can* and *can't*. The sample included utterances which spanned the entire period studied following the emergence of potential assimilation sites, from 2;8;21 to 4;0;7.

The utterances were judged independently by the current authors' two project supervisors, both of whom are expert phoneticians. This was in accordance with the recommendation that consensus should be established between three raters (Shriberg et al., 1984). A discussion then followed, in which the three raters described the phenomena which they perceived and stated their categorical preferences. Although the current author was able to rate each of the potential assimilation sites selected, the second and third raters were unable to rate a minority of instances. The categories assigned by the three raters are given in table 5.1 below.

Table 5.1. Categorical Ratings of Utterances Selected for Consensus Testing

Legend

Open: open juncture

Other: other non-assimilation phenomena

Gradient: perceived as gradient between assimilation and non-assimilation

Unclassifiable: could not confidently be classified as belonging to one of the three categories

Age	Utterance	Rater One	Rater Two	Rater Three
2;8;21	"Can get it out"	Open	Open	Open
2;9;21	"I can't get through now"	Other	Other	Other
2;10;8	"Ah! I can't (re)member"	Open	Open	Open
2;10;24	* "Can't (re)member it"	Other	Other	Other
2;11;00	"I can't blow it"	Assimilation	Assimilation	Assimilation
2;11;25	"You can go home now"	Open	Open	Open
3;1;3	"I can (.) make some room now"	Assimilation	Assimilation	Assimilation
3;2;3	"You can count number one"	Open	Open	Open
3;2;5	"I get a, I, I can pl, Teletubbies can play with my fishing-rod"	Assimilation	Assimilation	Assimilation
3;2;5	"I get a, I, I can pl, Teletubbies can play with my fishing-rod"	Open	Open	Assimilation
3;2;5	"Look this, you can put in"	Other	Other	Other
3;3;2	"This can be a train spotter"	Other	Assimilation	Assimilation
3;3;2	"But I can't get it in"	Other	Other	Gradient
3;3;6	"Erm you can be a girl (.) fast asleep"	Assimilation	Assimilation	Assimilation
3;3;6	"Ah you can, I can be, you can be a bear"	Assimilation	Assimilation	Assimilation
3;3;6	"Ah you can, I can be, you can be a bear"	Assimilation	Assimilation	Assimilation
3;3;6	"You can be, I can be a girl"	Assimilation	Assimilation	Assimilation

3;3;6	"You can be, I can be a girl"	Open	Assimilation	Assimilation
3;3;7	"You can come"	Open	Open	Open
3;4;3	"Because over here checking the wagon can go fast"	Open	Open	Open
3;4;3	"I can collect it now"	Assimilation	Assimilation	Assimilation
3;5;2	"No I can't (be)cause I'm a busy man"	Other	Other	Unclassifiable
3;6;0	"Anybody can get on"	Open	Open	Open
3;6;0	"You can keep (th)em"	Open	Assimilation	Assimilation
3;7;2	"Can't catch me (be)cause I'm the pooey-man"	Assimilation	Assimilation	Assimilation
3;7;2	"Can't catch me (be)cause) I'm a gingerbread candle"	Other	Other	Other
3;7;4	"You can, I can build/built your house couldn't I?"	Assimilation	Other	Unclassifiable
3;8;3	"This can go dud-dud"	Other	Other	Unclassifiable
3;8;5	"I can blow on this"	Other	Other	Other
4;0;7	"I can go schhhooo"	Other	Other	Other

These results show high agreement between the three raters. All three raters agreed on the classification of 22 out of 30 potential assimilation sites (73.3%). At least one other rater agreed with the current author's classification of a further four sites (13.3%). This means that agreement occurred between the current author and at least one other rater on the classification of 26 out of 30 sites (86.6%). Both of the second and third raters either disagreed with the current author's rating, or were unable to classify the phonetic phenomena occurring at only four potential assimilation sites (13.3%). Tests of intra-rater transcription were not carried out, owing to the current author's extreme familiarity with and memory of the data following such in-depth analysis.

Chapter Six

Thomas' Assimilation Development in Constructions Containing *Can* and *Can't*

6.1. Introduction

This chapter begins with a section on aspects of Thomas' segmental phonological development, which may have influenced the phenomena occurring at potential assimilation sites with *can* and *can't*. Firstly, his realisation of target velar plosives is explored, which is relevant to the later discussion of specific phonetic phenomena occurring at potential assimilation sites. Thomas' realisation of word-final target bilabial and velar nasals is also noted, in order to provide the necessary background information for understanding his pattern of assimilation development. Following this introductory section, the remainder of the chapter focuses on the quantitative behavioural trends and qualitative details of individual phonetic phenomena occurring at Thomas' potential assimilation sites with *can* and *can't*.

6.2. A Control Analysis of Thomas' Segmental Phonological Development

Prior to investigating patterns in Thomas' assimilation development, it was necessary to explore those aspects of his segmental phonological development which could potentially have influenced the phenomena observed at potential assimilation sites. The main phenomenon which was found to be important was Thomas' velar fronting of target alveolar plosives. This affected his realisations of the onsets of both *can* and *can't*, as well as leading to progressive alveolar assimilation and consonant harmony at potential velar assimilation sites. It was also considered important to note whether Thomas was able to produce word-final target bilabial,

alveolar and velar nasals prior to the emergence of bilabial and velar assimilation in connected speech.

6.2.1. Thomas' Realisations of Target Alveolar and Velar Plosives

Appendix Two shows the phonetic transcriptions obtained from a detailed phonetic analysis of Thomas' target alveolar and velar plosives. Quantitative summaries of realisations of these results are shown in Tables 6.1 and 6.2 below.

Table 6.1. Thomas' Realisations of Target Alveolar Plosives

Age	Number of Data Points	Alveolar Realisations	Velar Realisations	Other Realisations
2;3	13	12	1	0
2;6	30	30	0	0
2;9	30	28	2	0
3;0	30	29	1	0
3;3	30	30	0	0
3;6	30	30	0	0
3;9	30	30	0	0
4;0	30	30	0	0

Table 6.2. Thomas' Realisations of Target Velar Plosives

Age	Number of Data Points	Alveolar Realisations	Velar Realisations	Other Realisations
2;3	15	7	5	3
2;6	28	10	13	5
2;9	30	3	26	1
3;0	30	1	27	2
3;3	30	2	28	0
3;6	30	0	30	0
3;9	28	0	28	0
4;0	30	1	29	0

These results show that the vast majority of Thomas' target alveolar plosives were realised with an alveolar place of articulation throughout the study. Other phonetic features were also evident in individual productions; for instance, many alveolar plosives were dentalised and some were retracted to a more postalveolar place of articulation. The number of aspirated initial alveolar plosives also increased with age. Only four instances out of a total of 223 words were realised with a velar place of articulation; all of these were word-initial. Three of these occurred at age two and the final one occurred at age 3;0. No places of articulation other than alveolar or velar were sampled. From 3;3 onwards, all target alveolar plosives were realised with alveolar articulation, indicating that Thomas had achieved consistent and stable production of alveolar plosives at this age.

Thomas' realisations of target velar plosives were much more variable, with less than half of the sample realised with velar articulation at ages 2;3 and 2;6. A clear pattern of velar predominance emerged from 2;9 onwards, which persisted until the end of the sampling period at 4;0. The most commonly occurring non-velar realisation was alveolar, although a minority of other places of articulation also occurred, including palatal, uvular, glottal and bilabial. A total of 24 alveolar articulations occurred out of a sample of 221 words, whereas only eleven other articulations were found in the sample. Out of the 24 alveolar realisations, 20 occurred at age two and four occurred at ages three and four, indicating an increase in phonological stability and consistency over time. The distribution of alveolar realisations of target velar plosives was relatively similar across word positions; nine were word-initial, eight were word-medial and seven were word-final. Out of the eight word-medial instances, five were intervocalic and three occurred in consonant clusters. Age 3;3 was the last point at which alveolar realisations were seen with any regularity; one instance occurred in each word position. The only alveolar realisation sampled after this was at age 4;0 in *excavator*, which is relatively complex, containing four syllables and a three-consonant cluster. Places of articulation other than alveolar and velar were not sampled from age 3;3 onwards, which also reflects Thomas' increasing consistency over time.

The combined results for this control analysis indicate greater variability in the production of target velar plosives than target alveolar plosives, especially at age two. Age 3;3 appears to have been a milestone in terms of phonological stability, which marked the elimination of all variant forms apart from a minority of alveolar realisations of target velar consonants. These findings are consistent with previous research, which reports that velar fronting typically occurs at ages two and three and is usually eliminated by age four (Grunwell, 1987; Hewlett, 1988; Ingram, 1976).

6.2.2. Thomas' Realisations of Word-Final Target Nasals

In order to confirm that word-final bilabial, alveolar and velar nasals were present in Thomas' speech prior to the emergence of assimilation, examples of relevant words were obtained by scanning the utterances transcribed for the purpose of the segmental analysis of target velar plosives (see Appendix Two). It was evident that Thomas was able to produce word-final target bilabial, alveolar and velar nasals at age 2;3, seven months prior to the emergence of bilabial assimilation and eleven months prior to the emergence of velar assimilation. Word-final /m/ was evident in the words *time* [ˈt̪aˑĩm] and *ice-cream* [ə̃ˑʔnˈd̪ĩm]. Word-final /n/ was evident in the words *doing* [ˈd̪uwĩˑn], *can* [ˈt̪ɐ̃̀n] and *coming* [ˈtˈʌ̃mĩn]. Word-final [ŋ] was evident in the word *bang* [ˈbæːŋˑ].

6.3. Quantitative Trends in Thomas' Assimilation Development

This section discusses the quantitative trends in Thomas' assimilation development relative to other phonetic behaviours occurring at potential assimilation sites. Some qualitative information is given here, but a more detailed qualitative discussion of specific phonetic phenomena is provided in the next section. The impressionistic phonetic transcriptions on which these results are based are given in Appendices Three and Four.

6.3.1. The Development of Bilabial Assimilation

Figure 6.1 and Table 6.3 show the proportions of different phonetic behaviours occurring at potential bilabial assimilation sites. The table shows a complete breakdown of the proportions of individual phenomena. In contrast, the graph simply shows proportions of assimilations relative to all other phenomena, including open junctures and other non-assimilation phenomena grouped together.

Ten potential assimilation sites were sampled at age two, all of which occurred in constructions with *can't*. Four occurred at age 2;10 and six occurred at 2;11. The first potential site occurred at 2;10;8 and was realised with open juncture. However, all subsequent forms sampled at age two were realised with close juncture, showing clear predominance of close juncture forms as soon as potential sites emerged. Bilabial assimilation emerged at age 2;10;13, only five days after the emergence of potential sites at 2;10;8. Assimilation occurrence increased from one out of four potential sites at 2;10 to four out of six potential sites at 2;11. This increase is not considered to be evidence of clear predominance of assimilation at this age, owing to the small numbers of potential assimilation sites sampled.

Figure 6.1. Proportions of Bilabial Assimilations and other Phenomena at Potential Sites

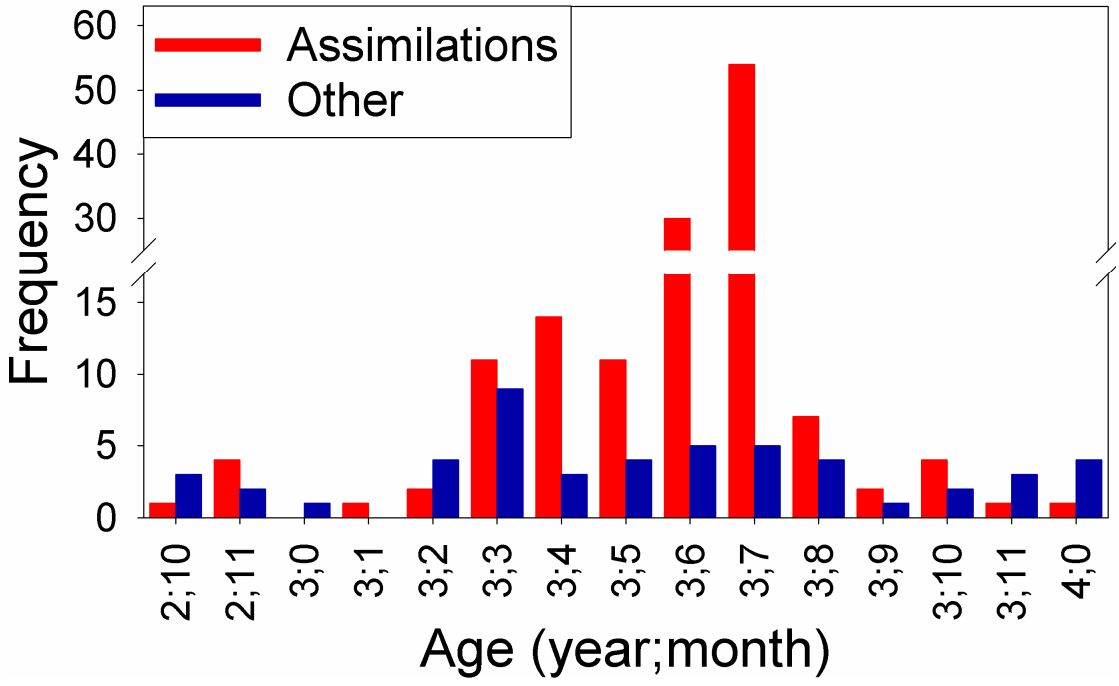


Table 6.3. Summary of Potential Bilabial Assimilation Site Realisations

Age	Total Sites	Assimilations	Other Non-Assimilation Phenomena with *Can*	Other Non-Assimilation Phenomena with *Can't*	Open Junctures with *Can*	Open Junctures with *Can't*	Indeterminate/ Unclassifiable	Untranscribable
2;10	4	1	0	2	0	1	0	0
2;11	6	4	0	2	0	0	0	0
3;0	1	0	0	1	0	0	0	0
3;1	1	1	0	0	0	0	0	0
3;2	6	2	0	2	1	0	1	0
3;3	20	11	0	1	5	0	1	2
3;4	17	14	0	0	2	0	0	1
3;5	15	11	0	0	3	0	0	1
3;6	35	30	1	1	3	0	0	0
3;7	59	54	1	0	3	1	0	0
3;8	11	7	2	0	2	0	0	0
3;9	3	2	0	0	1	0	0	0
3;10	6	4	0	0	2	0	0	0
3;11	4	1	1	0	0	0	1	1
4;0	5	1	3	1	0	0	0	0

All bilabial assimilations at age two were realised with elision of final /t/ and continuous voicing across the word boundary from the coda of *can't* to the voiced onset consonant of the following word. For example, "can't remember" was realised as [kʰɑ̃m ˈmɛmbəʊ] and "can't blow" was realised as [kʰɑ̃m ˈbləːʊːw].

Only one potential bilabial assimilation site was identified each at 3;0 and 3;1. The latter was the first potential assimilation site with *can*: "can make", which was realised with assimilation, despite a perceptible pause at the word boundary: "I can (.) make some room now" [ˈaɪ kʰæ̃m (.) meɪ]. Age 3;2 was characterised by a substantial increase in the frequencies of *can* and *can't* and the range of main verbs with which they occurred. Bilabial assimilation continued to be a minority behaviour, occurring at two out of five transcribable sites.

Age 3;3 marked the start of a developmental trend towards higher numbers of potential assimilation sites and the predominance of assimilation over other phonetic forms. This assimilation predominance continued to increase up to age 3;7. This change corresponds with the emergence of the construction "can be" at age 3;3 and its subsequent high frequency throughout this period. Figures and percentages of assimilation predominance are as follows:
Age 3;3: eleven out of 18 transcribable sites (61.1 %);
Age 3;4: 14 out of 16 transcribable sites (87.5%);
Age 3;5: eleven out of 14 transcribable sites (78.6%);
Age 3;6: 30 out of 35 sites (85.7%);
Age 3;7: 54 out of 59 sites (91.5%).

It is noteworthy that the sample of potential assimilation sites at age 3;5 is similar in size to those found at 3;3 and 3;4, despite the availability of only four recordings. This therefore indicates that a further increase in the number of potential assimilation sites may have been evident at 3;5, if a full quota of five monthly recordings had been available.

The period from age 3;8 until 4;0 was characterised by substantial reductions both in the numbers of potential assimilation sites sampled and the proportions of sites realised with assimilation. There were reductions both in the numbers of recordings in which sites were sampled and in the numbers of sites occurring within individual recordings. Assimilation continued to show a trend towards predominance at 3;8. However, this predominance had reduced to a level comparable with that observed at age 3;3. This predominance cannot be viewed as statistically meaningful, owing to the low numbers of potential sites sampled. From 3;9 onwards, there was no clear evidence of assimilation predominance over other forms. Numbers of sites were reduced to single figures. Although assimilation occurred in over half of potential sites at ages 3;9 and 3;10, these numbers are again too small to draw conclusions of statistical predominance. Assimilation appeared to be a minority behaviour in the last two months of the study. Despite the reduction in assimilation predominance throughout this period, close juncture forms continued to predominate over open juncture forms. Proportions and percentages of assimilations during this period are as follows:

Age 3;8: seven out of eleven sites (63.6%);

Age 3;9: two out of three sites (66.7%);

Age 3;10: four out of six sites (66.7%);

Age 3;11: one out of four sites (25%);

Age 4;0: one out of five sites (20%).

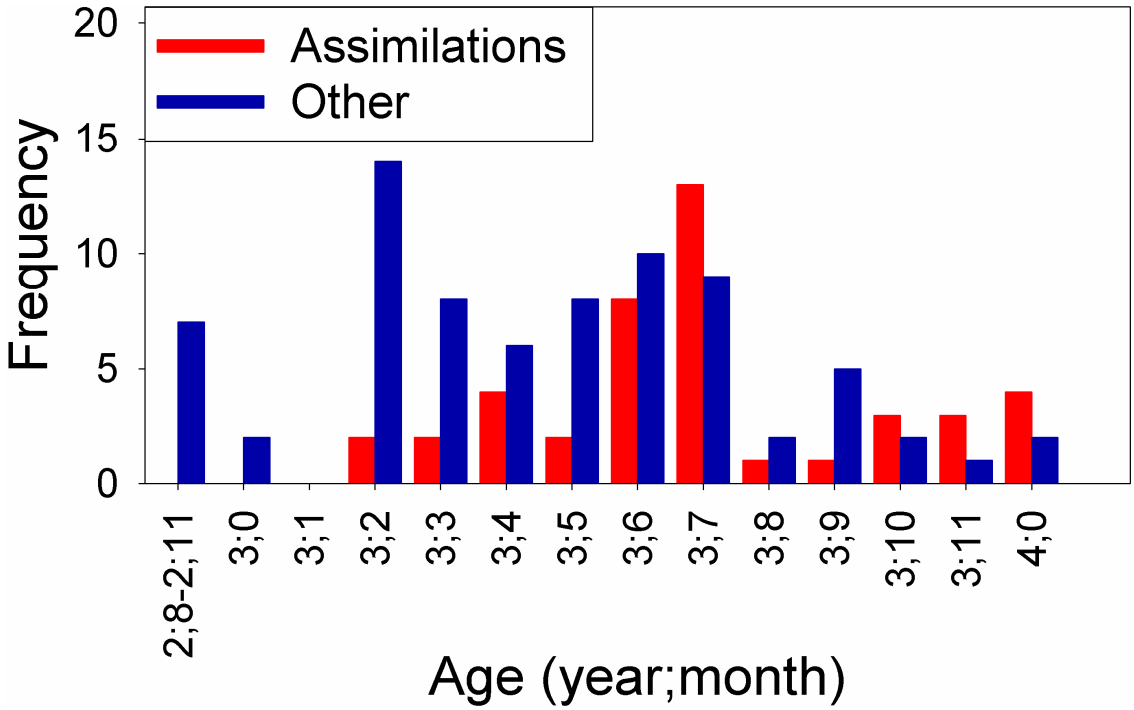
6.3.2. The Development of Velar Assimilation

Table 6.4 and Figure 6.2 show the proportions of different phonetic behaviours occurring at potential velar assimilation sites. The table shows a complete breakdown of the proportions of individual phenomena. In contrast, the graph simply shows proportions of assimilations relative to all other phenomena, with open junctures and other non-assimilation phenomena grouped together.

Table 6.4. Summary of Potential Velar Assimilation Site Realisations

Age	Total Sites	Assimilations	Other Non-Assimilation Phenomena with *Can*	Other Non-Assimilation Phenomena with *Can't*	Open Junctures with *Can*	Open Junctures with *Can't*	Indeterminate/ Unclassifiable	Untranscribable
2;8-2;11	7	0	0	4	2	1	0	0
3;0	2	0	0	2	0	0	0	1
3;1	0	0	0	0	0	0	0	0
3;2	16	2	0	9	4	0	0	1
3;3	10	2	0	4	2	0	0	2
3;4	10	4	0	1	4	0	0	1
3;5	10	2	1	5	1	0	1	0
3;6	18	8	0	2	7	0	0	1
3;7	22	13	0	2	5	0	1	1
3;8	3	1	1	0	1	0	0	0
3;9	6	1	0	0	4	0	0	1
3;10	5	3	0	1	1	0	0	0
3;11	4	3	0	0	1	0	0	0
4;0	6	4	0	0	1	0	0	1

Figure 6.2. Proportions of Velar Assimilations and Other Phenomena at Potential Sites



Although Thomas began to produce potential velar assimilation sites from age 2;8, no assimilations were observed during the following six months. However, there was a trend towards the predominance of non-assimilation close juncture forms over open junctures. Close juncture forms accounted for four of the seven sites sampled at age two (54.1%) and both of the two sites occurring at age 3;0. No potential velar assimilation sites were sampled at 3;1.

At age 3;2, the increased frequency and range of constructions with *can* and *can't* led to a sharp rise in the number of potential velar assimilation sites, from two at age 3;0 to 16 at age 3;2. This was also the point at which velar assimilation emerged as a minority behaviour. Numbers then dropped to a sample of ten sites monthly from 3;3 to 3;5. The sample at age 3;5 was taken from only four available recordings. An increase in potential velar assimilation sites may therefore have been evident if a full quota of five monthly recordings had been available. Velar assimilation remained a minority behaviour throughout this period. Close juncture forms continued to predominate over open junctures, except for age 3;4, when

proportions were more equal (five close junctures, four open junctures and one untranscribable site).

A period of assimilation establishment was evident at ages 3;6 and 3;7, when numbers of potential sites and proportions of assimilations reached their highest levels. At age 3;6, the number of sites increased from ten to 18 and assimilation occurred at eight of the 17 transcribable sites (47.1%). At age 3;7, these numbers rose further to 22 sites; assimilation occurred at 13 out of 21 transcribable sites (61.9%). Thus, assimilation occurred at just under half of potential sites at age 3;6 and showed a trend towards predominance only at age 3;7. Close juncture continued to predominate over open juncture.

Similarly to the findings reported for bilabial assimilation development, there was a sharp decline in numbers of potential velar assimilation sites sampled from age 3;8 to 4;0. There was a reduction in both the numbers of monthly recordings in which potential velar assimilation sites were sampled and numbers of sites within individual recordings. Numbers of potential sites were reduced to single figures throughout this period. Although relative proportions of assimilations were relatively high from age 3;10 to 4;0, numbers are too small to draw conclusions of clear predominance. Close juncture forms continued to predominate over open junctures, except at age 3;9, when open junctures occurred at four out of six sites. Numbers and proportions of assimilations during this period are as follows:

3;8: one out of three sites (33.3%);

3;9: one out of six sites (16.7%);

3;10: three out of five sites (60%);

3;11: three out of four sites (75%);

4;0: four out of six sites (66.7%).

6.4. Specific Phonetic Phenomena Occurring at Potential Assimilation Sites

6.4.1. Emerging Phonetic Forms of *Can*

The verb *can* first emerged as the final element in the sentences "I can" and in multiple occurrences of "yes I/we/it can". The latter was a formula learned from the song *Bob the Builder*, from a popular British children's television cartoon of the same title. These strong forms of *can* were most frequently realised either as [ˈkʲʰæ̃n] or [ˈt̪ʰæ̃n]. However, other phonetic variations were evident in individual forms, such as dentalisation and aspiration of initial /t/ and variable vowel features (see Appendix Three).

From age 2;7 onwards, Thomas continued to produce these strong forms of *can* in similar contexts. However, *can* also emerged as an auxiliary followed by a main verb. The majority of auxiliary forms of *can* occurring from age 2;7 to 2;11 were realised as weak forms, and primary stress was placed either on the main verb or on another syntactic element. However, there existed much phonetic variation between these weak forms. Examples include "I can see you" [a m̩ ˈs̪i] at 2;7;9, "yes can hear (and a)" [xə̃n hɪɐʊ] at 2;7;15, "can see flowers" [(C,Vls)n̩ si] at 2;9;3, "I can see a little carrot" [gə̆̃ sɨj] at 2;9;7, "I can see Purdie" [ɣə̃ s̪i] at 2;9;21 and "you can go home now" [kʰn̩ gəʊ] at 2;11;25. The last of these was the only instance in which *can* occurred at a potential velar assimilation site at age two.

Only a minority of auxiliary forms of *can* occurring from age 2;7 to 2;11 were realised as strong forms; these were "can see please?" [ˈt̪ʰæ̃n si] at 2;7;26, "Bob the Builder, can we fix it?" [ˈt̪ʰæ̃n ˈwɨ] at 2;8;21 (*Bob the* Builder formula) and "can get it out" [ˈt̪ʰæ̃n d̪ɛt̪] at 2;8;21 (possible alveolar/dental assimilation throughout the whole utterance. A single unstressed form of *can* with no vowel reduction occurred, which was realised differently from the stressed forms; this occurred in "I can (.) off people" [kʰĩn: (.) ˈʔɒf] at 2;11;10. It can be seen that all except two weak forms of *can* were realised with initial velar consonants, although

these included a range of both voiced and voiceless plosives and fricatives. In contrast, all stressed forms of the auxiliary *can* were realised with an initial /t/, with phonetic variation across individual forms such as dentalisation or retraction. The segmental analysis of Thomas' development of velar plosives showed variability between alveolar and velar production of target velar plosives during this period (see section 6.2.1).

During the period from age 3;1 to 3;4, these weak forms of the auxiliary *can* were eliminated at potential assimilation sites. Only two weak forms were sampled in non-assimilation contexts: "can eat" [kn̩ ʔit] and "can hold" [kʰäd həʊl] at 3;3;4. The second of these showed only partial vowel reduction and could therefore be considered a partial weak form. Instead, most productions of *can* were realised with full non-reduced vowels, including those which were unstressed. Examples include "can stroke" [kʰæ̃n stʊəʊk], "I can make" [ˈaɪ kʰæ̤̃m (.) meɪ], "can count" [ˈkʰæ̃n (.) ˈt̤aɪ] and "can go" [kʰæ̃ŋ gɛ]. However, phonetic variations occurred in a minority of instances. One occurrence was evidently an immature form: "can count" [t̤ʰæ̃n ˈtaʊ]. Other forms showed the types of phonetic variation also reported in adult connected speech (Cruttenden, 2001; Shockey, 2003). These forms had a variety of initial velar consonants, including "can see" [ˈkʰæ̃ sij], "can eat (th)em" [g̊æ̃n ˈiṭ ə̃m], "can do" [xæ̃n duw], "can see" [kxæ̃ ˈsi], "can" [ˈkʲą̃], "can be" [ɣæ̃m bij], and "we can" [ˈcæ̃n].

Age 3;5 marked the re-emergence of weak forms of *can* at potential bilabial assimilation sites. All instances at this age were realised with close juncture. Bilabial assimilation was observed in a minority of weak forms at this age. Examples include "can be" [t̤ə̃m bi] and [kə̃m bi]. However, weak forms at potential velar assimilation sites were realised with non-assimilation close juncture, for example "can go" [kə̃ gəʊ]. Age 3;6 marked further developments in the emergence of weak forms of "can". Open juncture forms emerged at potential bilabial assimilation sites, for example "can put" [kʰn̩ pʰʊ(Pl,Vls)] at 3;6;0. Weak forms with velar assimilation emerged, for example "can catch" [kʰə̃ŋ ˈkʰætʃ] at 3;6;4. Weak forms also re-emerged in non-assimilation contexts, for example "can I watch" [kʰn̩ aɪ] at 3;6;2.

It can be seen from these examples that weak forms of *can* were mainly realised either as [kʰə̃n] or [kʰn̩] (with place assimilation of the final nasal at many potential assimilation sites). However, other forms also occurred in which there appeared to be only partial vowel reduction towards a weak form. Examples include "can put" [kʰə̈m pʊʔ] at 3;6;2 and "can pinch" [kʰə̈m ˈfĩt͡ʃ] at 3;7;2. Assimilated weak forms of *can* emerged in constructions which had previously been sampled with assimilated strong forms of *can* (for example "can be" and "can catch"). It is noteworthy that they were also sampled in constructions with novel verbs, for example "can cross" [kʰə̯ŋ ˈkʊɒs] at 3;7;2. However, this finding should be interpreted cautiously, because it is possible that these constructions had previously occurred with assimilated strong forms of *can* in Thomas' speech, but these were not sampled in the data. In contrast, the weak form of *can* in the novel construction "can crash" was produced with open juncture as [kʰə̃n ˈkʊæʃ] at 3;7;3. Following their re-emergence, weak forms of *can* continued to occur alongside strong forms until the end of the study at age 4;0. Assimilations, non-assimilation close junctures and open junctures were found.

Non-assimilation close junctures with *can* occurred at potential assimilation sites when there was no final nasal in *can*. This phenomenon emerged in a minority of weak forms of *can* at age 2;9 and strong forms of *can* at age 3;3, for example "can see" [gə̯̆ s̪ij] at 2;9;7 "can be" [kʰæ̃ bij] at 3;3;2. However, occurrences were much more frequent in weak forms of *can* both at potential assimilation sites and in non-assimilation contexts, following their re-emergence at 3;5. Examples include "can go" [kə̃ gəʊ] at 3;5;5, "can stay" [kʰə̯̃ ˈsteɪ] at 3;6;2, "can be" [kə bij] at 3;7;5, "can bake" [kʰə ˈbeːɪkʲʰ] and "can blow" [k͡xə̃ bləʊ] at 3;8;4, "can I" [kə̃ ˈaɪj] and "can look" [kʰə̃ ˈlʊk] at 3;9;3. In some instances, the schwa vowel in *can* was nasalised, indicating that these were clearly weak forms of *can*. However, in some instances in later months, the schwa was not nasalised, leading to difficulties in distinguishing between weak forms of *can* and *could*. A control analysis of utterances containing *could* at 3;9;5 revealed very similar phonetic forms, such as "could be" [kə bij]. This confusion led to the elimination of two utterances at age 3;9 and one utterance at 3;10 from the analysis, which were transcribed as containing *can* in the Chat transcripts.

6.4.2. Emerging Phonetic Forms of *Can't*

Productions of *can't* were first sampled at age 2;6, three months after the emergence of *can*. Similarly to the patterns observed for *can*, *can't* initially occurred regularly as the final element in a sentence, for example "I can't". In this context, *can't* was always produced with a fully released final [t]; examples include [ˈkʰɑ̃nʔtː] at 2;6;19 and [ˈt̪ɑ̃ːnt̪] at 2;6;25. Variant phonetic forms without a final [t] did not emerge in this context until 2;11;14, when the forms [kʰɑ̃n] and [kʰɑ̃nʔ] were sampled in two occurrences of "I can't".

As also found in early productions of *can* at age two, initial target /k/ in *can't* at age two was variably realised either as [t] or [k], with additional occasional occurrences of initial [d] or [g]. In contrast with the findings for *can*, there were no instances with initial palatal plosives or palatalised alveolar plosives. Interestingly, there were no occurrences of initial [t] in *can't* at age three, although initial [t] occurred in a minority of instances of *can* at this age. These differences between the phonetic variants of *can* and *can't* may have been the result of coarticulation towards the following back vowel [ɑ] in *can't*, whereas the front [æ] vowel in *can* may produce a phonetic environment more conducive to alveolar fronting of velar plosives.

Although Thomas consistently produced final [t] in all productions of sentence-final *can't* until age 2;11;14, forms of *can't* prior to a main verb were almost always realised without final [t] from their initial emergence at 2;6, creating close juncture with the initial consonant of the following main verb. This produced either assimilation or non-assimilation close juncture. The most frequently occurring close juncture forms of *can't* were realised either with a final nasal-glottal stop cluster [nʔ], [mʔ] or [ŋʔ], a final singleton nasal [m], [n] or [ŋ] or a final nasalised vowel, sometimes followed by a glottal stop, but without a nasal. Final nasal-glottal stop clusters and singleton nasals were observed at close junctures with *can't* until the end of

the study at age 4;0. Productions either with a final nasalised vowel or a nasalised vowel followed by a glottal stop were not observed with any regularity beyond age 3;5, except for an isolated instance in a reduced form of "can't see" [kə̃ʔ ˈs̪ij] at 3;10;2. These phenomena occurred both at potential assimilation sites and in non-assimilation contexts, although there existed phonetic variability across individual forms, especially at age two. Examples at potential assimilation sites include "can't get" [ˈkʰə̃n gɛ̰ʔ] at 2;9;28, "can't get" [ˈkʰə̃nʔ gʲɛ̰t] and "can't (re)member" [kʰə̃m ˈmɛmbəʊ] at 2;10;13, "can't (re)member" [kə̃nː ˈmɛmbəʊ] at 2;10;24, "can't (re)member" [gɒ ˈmɛmbeʊ] at 2;10;26, "can't get" [ˈkʰə̃ŋ gɛʔ] at 3;2;4, "can't get" [kʰə̃ŋʔ gɛ̰t] at 3;2;6 and "can't pinch" [kə̆m ˈpɪn͡tʃ] at 3;3;4. Examples in non-assimilation contexts include "can't see" [k̟ə̃ʔ sij] at 2;6;12, "can't reach" [kʰə̟̃nʔ ˈʊi̯t͡ʃ] at 2;7;1, "can't sleep" [gə̈ʔ ˈθlɪp] at 2;7;5, "can't see" [kʰə̃n ˈs̪i] at 2;10;7, "can't see" [ˈkʰə̃ si] at 2;11;25 and "can't get" [ˈkʰə̃n gɛt] at 3;3;2.

A minority of Thomas' productions of *can't* showed evidence of vowel reduction. Such forms have not been documented in adult speech. These forms were produced with either a schwa vowel or a centralised [ɑ] vowel. Examples include "can't pay" [ˈkə̃n feɪ] at 3;2;5, "can't see" [kʲə̃n̊ʔ ˈs̪i̥ʔi] and "can't" [kə̃t̪] at 3;10;2 and "can't get" [kɜ̃ŋ gɪt̪] at 3;10;3.

From age 2;9 to 3;7, a minority of instances of the auxiliary *can't* were realised with adult open juncture between the final [t] in *can't* and the abutting initial consonant of the next word; examples include "can't see" [kʰə̃nt̪ s̪i] at 2;10;0, "can't (re)mem(ber)" [ˈkʰə̃ntʷ (.) ˈmɛə̃m] at 2;10;8, "can't tell" [ˈkʰəhə̃ːn̪̊t̪ (..) tɛ] at 3;6;2 and "can't mum" [ˈkʰə̃nt (.) Mʊm] at 3;7;2. Notably, most of these were also produced with a perceptible pause at the word boundary. Others were produced with an epenthetic vowel at the word juncture; examples include "can't see" [kʰə̃ndə̆ si] at 2;9;18, "can't get" [ˈkə̃ntʷʏ dɛʔ dɪs] at 2;9;25 and "can't we" [ˈkʰə̃n̪̊t̪ə̆ wɨ̈] at 3;7;3.

6.4.3. Possible Progressive Alveolar Assimilation

This phenomenon occurred at a minority of potential velar assimilation sites with *can* and *can't*, following their emergence at age 2;8 until age 3;2. This was produced when a final /n/ in *can* or *can't* was followed by alveolar realisation of the following word-initial target velar plosive. Examples include "can get" [ˈt̪ʰæ̃n d̪ɛt̪] at 2;8;21, "can't get this" [kʰɑ̃n d̪ɛˈt̪ɪːs̪] at 3;0;0, "can't get it" [kʰɑ̃n ˈd̪ɛt̪ ɪt̪] at 3;0;1 and two instances of "can count" [t̪ʰæ̃n ˈtaʊ] and [ˈkʰæ̃n (.) ˈt̪ạɪ] at 3;2;3. These findings correspond with those of the segmental analysis of target alveolar and velar plosives; this showed variable alveolar and velar realisations of target velar plosives at age two and a minority of alveolar realisations of target velar plosives at age three. This pattern of progressive alveolar assimilation is linked with Thomas' pattern of velar fronting at this age (see sections 6.2.1 and 6.4.1).

6.4.4. Progressive Nasal Assimilation and Gemination

This occurred when the initial target plosive in a word following *can* or *can't* was realised as a nasal. The first occurrence of this was sampled at age 2;11;6 in "I can't put it on now fall down the floor again" [ˈkʰɑ̃nʔ mʊːt]. This was the first potential bilabial assimilation site occurring in an utterance which may have been an attempt at clause subordination. All other instances occurred at age three in the utterance "can be" and the apparent progressive nasal assimilation co-occurred with regressive place assimilation. Sometimes, the resultant nasal at the word boundary was short and it was therefore unclear whether the following word-initial consonant had been omitted, rather than assimilated to a nasal manner of articulation; examples include "can be" [kʰæ̃mi] and [kʰæ̃ˈmi] at age 3;3;6, [kʰæ̃mij] at 3;5;4 and [kʰə̃mij] at 3;7;5. However, in other instances, the nasal at the word boundary was longer and was more clearly the result of gemination; examples include "can be" [kʰæ̃m mi] (alternative transcription [kʰæ̃mːi]) at age 3;4;3 and [kʰə̃m mij] at 3;8;3. Notably, the last instance in each set of examples occurred in a weak form of *can* following the re-emergence of weak forms, indicating that these phenomena could equally occur in strong and weak forms. In one

instance, the progressive nasal assimilation appeared to be only partial, leading to a sound which had some perceptual characteristics of both a /b/ and /m/: "can be" [kæ̃m b̃iˑ] at 3;5;5.

6.4.5. Partial Assimilation and Gradual Assimilation

In their study of connected speech process (CSP) development, Newton and Wells (1999) employed a category for those utterances which they judged to have been realised with partial CSPs. These were utterances which they rated as having intermediate values, which did not show clear evidence of CSPs occurring, but which equally did not show clear evidence of open juncture occurring (Newton and Wells, 1999). This did not emerge as a major category from the classification of Thomas' phonetic behaviours in the current study. However, two potential assimilation sites had realisations which could be classified as partial assimilation. The first occurred when the final /n/ in *can* became palatalised in anticipation of the following /g/ in *go* to produce "can go" [kʰæ̃nʲ (..) ˈgəʊ]. Notably, there was also a pause at the word boundary. The second was realised with an audible transition from alveolar to velar articulation of the final /n/ in *can*: "can get" [kʰɑ̃nŋ ˈgɛt̪ˢ]. Such a transition is known as *articulatory slide* in the extIPA (Duckworth et al., 1990).

6.4.6. Assimilation and Pauses

Assimilation is a close juncture phenomenon, whereas the insertion of a pause at a word boundary is considered to be a characteristic of open juncture (Wells, 1994). It was therefore surprising to find two instances in which Thomas produced assimilation despite there being a pause at the word boundary. The first occurred at age 3;1;3 in the utterance "I can (.) make some room now", realised as [ˈaɪ kʰæ̃m (.) meɪ]. The second occurred at age 3;7;5 in the utterance "…mummy, you can (.) be a doctor couldn't you? A vet couldn't you?", realised as [kʰə̃m (.) ˈbɨ̈]. In a further utterance at age 3;6;0, assimilation occurred despite low syntagmatic fluency, which took the form of equal stress placed on each word and perceptibly

more open juncture between words in the utterance (Wells, 1994). Thus, the utterance "ˈI ˈcan ˈbe ˈone and you can sit next to me I can be a driver", the first potential assimilation site was realised as [ˈkʰæ̃m ˈbeɪ].

6.4.7. Assimilation in False Starts and Revisions

From age 3;2, Thomas began to produce false starts and revisions in his speech, reflecting his usage of increasingly long utterances and complex sentences. In cases when false starts and revisions occurred at a potential assimilation site, the phonetic behaviours occurring in *can* or *can't* were sometimes observed to change as the word was repeated. In all instances except one, this occurred in the construction "can be", which was used to assign roles in imaginative play. Utterances in which this phenomenon was observed can be broadly divided into two categories: those in which the whole assimilation site was repeated (for example " … can be … can be…"), and those in which the potential assimilation site was interrupted prior to the revision (for example "… can … can be …". Different patterns of occurrence were observed for each category.

In the earlier occurrences of whole assimilation site repetition, the false start was realised with assimilation, but the revision was realised with open juncture. Table 6.5 shows some examples.

Notably, the assimilated false start in "You can be, I can be a girl" was also realised with progressive nasal assimilation (see Table 6.5). The false start in "I can be a co, I can be a man can't I?" was realised with an assimilated weak form of *can*, whereas the revision was realised with an open juncture strong form of *can* (see Table 6.5). However, the reverse was true in a later instance at age 3;7;4, which was realised with a strong form in the false start and a weak form in the revision. Both of these forms were realised with assimilation; "you can be daddy xxx, you can be a mummy again" [kʰæ̃m bi), [kə̃m bij].

In utterances containing interrupted potential assimilation sites, regressive place assimilation was evident in the final /n/ of *can*, even though the utterance was interrupted prior to the production of the following word-initial consonant to which the nasal was assimilating. This first occurred at age 3;6;3 in the utterance "ah you can, I can be, you can be a bear" [ˈɣa̰m], [kʰæ̃m bi], [kʰæ̃m bij]. Although assimilation occurred in the first production of *can*, this was

Table 6.5. Summary of Assimilated False Starts and Open Juncture Revisions

Age	Utterance	Realisation
3;2;5	"I get a, I, I can pl, Teletubbies can play with my fishing-rod"	[kʰæ̃m pl̥], [kʰæ̃n ˈpleij]
3;3;6	"You can be, I can be a girl"	[kʰæˈmi], [kʰæ̃n bij]
3;6;3	"I can be a co, I can be a man can't I?"	[kʰə̃m bɪj], [kʰæ̃n bij]

not strictly a potential assimilation site, because *be* had not yet been produced. This also occurred later at age 3;7;2 in the utterance "you can, (0.5) you can bake something" [kʰæ̤̃mː], [kʰæ̃m ˈbeɪk]. In these cases, it is noteworthy that both the false starts and revisions were realised with assimilation. However, a new pattern emerged from age 3;7;4 onwards, in which false starts were realised with open juncture and revisions were realised with assimilation. This is the reverse pattern from that observed in instances of whole assimilation site repetition in the former half of Thomas' third year. These are summarised in Table 6.6 below.

These examples show a range of open junctures and assimilations in both strong and weak forms of *can*. This indicates that assimilations and weak forms of *can* occurred independently of each other.

The overall findings show a tendency towards assimilated false starts and open juncture revisions from age 3;2 to 3;6, increased consistency of assimilation across both false starts and revisions at 3;6 and 3;7 and a tendency towards open juncture false starts and assimilated

revisions from age 3;7 until 4;0. These changes correspond with the ages at which Thomas began to produce a wider range of syntactic constructions of increasing complexity (see Chapter Seven). It may not be necessary to distinguish between those false starts and revisions in which the whole potential assimilation site was repeated and those in which the potential assimilation site was interrupted. The categorisation used here emphasises the differences between them. However, because no instances of whole assimilation site repetition were sampled beyond 3;6, it is impossible to know whether these two types of false starts and revisions had similar or different emergent patterns of realisation. It is possible that whole potential assimilation site repetition was a characteristic of early false starts and revisions,

Table 6.6. Summary of Open Juncture False Starts and Assimilated Revisions

Age	Utterance	Realisation
3;7;4	"You can, I can build/built your house couldn't I?"	[k^hæ̃n], [k^hæ̃m bɪʊʔ]
3;11;3	"I got something (.) you can you can make (.) pet"	[gĩn], [k^hə̃m ˈmeɪk^x]
4;0;7	"If he can, (..) if we can go outside"	[k^hə̃n̊], [kə̃ŋ gəʊw]

while potential assimilation site interruption was characteristic of a later style of development. A larger data sample size would be necessary to draw more definite conclusions.

6.4.8. Summary of Specific Phenomena Observed at Potential Assimilation Sites

The verb *can* first emerged at age 2;3 in utterance-final position, for example in "I can". Usage of *can* as an auxiliary alongside another main verb then emerged at age 2;7. Most occurrences were realised as weak forms with initial velar consonants, although considerable phonetic variability existed between individual forms. In contrast, stressed forms of the auxiliary *can* were realised with an initial /t/. Velar fronting was a global pattern in Thomas' speech at this age, as shown by the segmental analysis of target velar plosives in initial, medial and final word positions (see section 6.2.1).

From age 3;1 to 3;4, weak forms of *can* were replaced almost exclusively by a range of stressed and unstressed forms which showed no vowel reduction; there were only two exceptions occurring in non-assimilation contexts. From 3;5 onwards, weak forms of *can* re-emerged alongside continued usage of strong forms, to include a range of assimilations, non-assimilation close juncture and open juncture forms over time. Judging from the range of assimilated and open juncture strong and weak forms occurring in utterances with false starts and revisions, it would appear that assimilation and weak forms are independent close juncture phenomena. Weak forms of *can* without final nasals created non-assimilation close juncture forms, which were sometimes indistinguishable from productions of *could* in later months. Only three forms of *can* with initial /t/ were sampled at age three, corresponding with the greatly reduced realisation of /k/ as [t] observed at age three.

Forms of *can't* were variably realised with initial [t] or [k] at age two, but no forms with initial [t] were sampled at age three. The word *can't* first emerged in sentence-final position and was realised with a final /t/. However, as it emerged as an auxiliary alongside main verbs, it took the close juncture forms [k^hãnʔ], [k^hãn] or [k^hã]. These forms were found both at potential assimilation sites and in non-assimilation contexts. The non-assimilation close juncture forms [k^hãnʔ] and [k^hãn] co-occurred alongside assimilation as it emerged. Phonetic forms of *can't* varied greatly at age two, but this variability declined at age three. Forms of *can't* showing vowel reduction were sampled at age three, but were relatively rare compared with weak forms of *can*. Open juncture forms of the auxiliary *can't* were also rare and were mostly realised either with a pause or an epenthetic vowel at the word boundary.

Progressive alveolar assimilation featured at some potential velar assimilation sites from age 2;8 to 3;2. This corresponds with findings of variable alveolar and velar production of /k/ at age two and a minority of alveolar realisations of /k/ at age three. Progressive nasal assimilation and gemination occurred most frequently in the construction "can be" from age

3;3 to 3;8. Gemination was more evident in some instances than others owing to differences in length of nasal.

In contrast with the findings of Newton and Wells (1999), only two utterances contained potential assimilation sites which could be classified as being realised with partial assimilation. Both of these occurred at potential velar assimilation sites. The first was realised with a word-final [nʲ] prior to a word-initial [g]. The second instance was realised with a gradual transition from alveolar to velar articulation of the word-final nasal. Because there were only these two occurrences, partial assimilation was not a major category adopted in this study for the classification of phonetic phenomena.

From age 3;2 onwards, Thomas produced some utterances with false starts and revisions occurring at potential assimilation sites. In the earliest occurrences, the whole assimilation site was repeated. From 3;2 to 3;6, the false start was realised with assimilation, whereas the revision was realised with open juncture. However, a single instance sampled at age 3;7 indicates that Thomas was beginning to use assimilation more consistently in both the false start and revision over time. In the second type of false start and revision, the potential assimilation site was interrupted in the false start, before being produced in full in the revision. The first two instances of this at ages 3;6 and 3;7 were realised with assimilation in both the false start and revision. However, later instances occurring from 3;7 to 4;0 were produced with assimilated false starts and open juncture revisions. This is the opposite pattern from that observed in the earliest false starts and revisions from 3;2 to 3;6. Therefore, the general emergent pattern appears to constitute assimilated false starts and open juncture revisions from 3;2 to 3;6, more consistent assimilation across false starts and revisions at 3;6 and 3;7 and open juncture false starts and assimilated revisions from 3;7 to 4;4. Because the two different types of false starts and revisions were not observed in parallel over time, it is difficult to know whether they had different emergent patterns, or whether in fact whole potential assimilation site revision and potential assimilation site interruption marked two consecutive developmental stages.

6.5. Conclusion

This chapter has focused on the phonetic phenomena occurring at potential assimilation sites in constructions containing *can* and *can't*. The trends in Thomas' assimilation development have been explored, as well as the nature of individual phonetic phenomena occurring at potential assimilation sites. Specifically, an interesting pattern in the development of strong and weak forms of *can* has been revealed, which is discussed in further detail in relation to syntactic development in Chapter Seven.

Chapter Seven

Thomas' Syntactic Development in Constructions Containing *Can* and *Can't*

7.1. Introduction

This chapter discusses the quantitative and qualitative advances in Thomas' syntactic acquisition which are relevant to his assimilation development. The quantitative analyses carried out include mean length of utterance (MLU), maximum length of utterance and frequency counts for the auxiliary verbs *can* and *can't*. A broad overview of Thomas' global syntactic development in constructions containing can and *can't* is then provided. This sets the scene for the Final discussion of those aspects of Thomas' syntactic development which appear to be directly linked with his assimilation development.

7.2. Typical Syntactic Development from Age Two to Four Years

This section summarises the typical milestones in syntactic development from age two to four years, as a framework for the interpretation of Thomas's data. Children are usually producing two-word utterances by their third birthday (Crystal, 1992; Buckley, 2003). These utterances consist of many nouns and an increasing range of other elements, including pronouns, verbs and adjectives (Buckley, 2003). Most of these utterances are declaratives, but rudimentary question forms such as "what doing" may also be evident (Crystal, 1992).

The period from age 2;0 to 2;6 is characterised by the emergence of sentences containing three clause elements. These elements include subject (S), verb (V), object (O), complement (C) and adverbial (A). Possible combinations in three-element utterances are SVO, SVC, SVA,

VCA and VOA. Question forms now also have three elements, for example, "what you doing" (Crystal, 1992). The first prepositions also emerge during this period.

From 2;6 to 3;0, children begin producing clauses of four or more elements and phrases of increasing complexity (Crystal, 1992). Possible clause structures include SVOA, SVCA, SVOC and SVAA. According to Crystal, tag questions and those with subject-verb inversion also emerge during this period, although Buckley (2003) documents these developments as occurring after age three. This discrepancy in the literature reflects individual differences in children's patterns of acquisition. Notably, the auxiliaries *can, will* and *be* emerge during this period, as well as the negative forms *can't* and *don't* (Buckley, 2003).

The period from age three to four is characterised by the emergence of complex sentences with clause and phrase coordination and subordination. According to Crystal (1992), these developments take place in the first half of the fourth year, whereas Buckley (2003) states that they are more characteristic of the end of the fourth year. The development of complex sentences corresponds with the emergence of conjunctions in the child's vocabulary, including *and*, *or*, *but*, *what*, *which*, *because*, *when* and *so*. Questions with initial auxiliaries also emerge during this period, such as *can I…* and *are you…* (Buckley, 2003). By their fourth birthday, the majority of children are producing most types of complex sentences and most grammatical elements. However, grammatical errors may be expected to continue over the next couple of years (Buckley, 2003).

7.3. Quantitative Analyses of Thomas' Syntactic Development

7.3.1. Mean Length of Utterance

This analysis aimed to provide a quantitative trajectory of Thomas' syntactic development. Table 7.1 shows Thomas' mean length of utterance in morphemes (MLUm) (for each month of the study).

These findings show the predicted general, gradual increase in Thomas' MLU throughout the study, from values typical of Brown's Stage I to those typical of Stage V. Thomas' MLU generally fell within the range of stage I from age 2;0 to 2;6, with the exception of one small rise beyond 2.0 at age 2;4. Thomas' MLU values then rose briefly to correspond with those typical of stage II for only two months at age 2;7 and 2;8. A further increase to stage III

Table 7.1. Thomas' Mean Length of Utterance

Age	MLUm	Brown's Developmental Stage
2;0	1.544	I
2;1	1.651	I
2;2	1.853	I
2;3	1.950	I
2;4	2.053	II
2;5	1.987	I
2;6	1.999	I
2;7	2.199	II
2;8	2.412	II
2;9	2.543	III
2;10	2.850	III
2;11	3.020	IV
3;0	2.819	III
3;1	3.006	IV
3;2	2.956	III
3;3	3.349	IV
3;4	3.398	IV
3;5	3.693	IV
3;6	3.941	V
3;7	4.181	V
3;8	3.891	V
3;9	3.888	V
3;10	3.919	V
3;11	3.824	V
4;0	3.451	V

values was then evident at ages 2;9 and 2;10. This was the first point at which Thomas' MLU exceeded the expected level for his age; the lower bound for stage III is 36 months (Brown, 1973).

The period from age 2;11 to 3;2 was characterised by fluctuation in Thomas' MLU across the boundary between stages III and IV. From age 3;2 to 3;5, MLU values fell clearly within the range for Stage IV, corresponding closely with the lower bound of 40 months specified for this stage (Brown, 1973). Between ages 3;6 and 3;11, MLU values fell within the stage V range, which again corresponds with the lower bound specified by Brown. The MLU value at age 4;0 dropped to 3.451, which is more typical of stage IV. One reason for this may be that one of the four available recordings at age four was cut short, owing to Thomas feeling unwell and being unwilling to speak.

These results initially indicate that Thomas was following a typical pattern of language acquisition, according to Brown (1973). However, his MLU values are in fact substantially lower than the MLU norms reported in recent research (Rice et al., 2010). Table 7.2 shows the MLUm results reported in six-month intervals for both typically developing (TD) children and children with specific language impairment (SLI) by Rice et al. (2010), compared with the equivalent mean MLUm values for Thomas. Thomas' values were obtained by averaging the six MLU values obtained on a monthly basis within each six-month period specified by Rice et al. (2010). For the purposes of this comparison, Thomas' MLU values are rounded to only two decimal places, as are the data reported by Rice et al. (2010).

Table 7.2. Comparison of Thomas' MLUm Values with Recent Normative Data.

Age	MLU Norms for TD Children	MLU Norms for Children with SLI	Thomas' MLU
2;6-2;11	3.23	2.59	2.50
3;0-3;5	3.81	3.07	3.20
3;6-3;11	4.09	3.36	3.94

These findings show that Thomas' MLU values were more similar to those of the children with SLI than those of the TD children within the first two six-month intervals, from age 2;6 to 2;11 and from 3;0 to 3;5. This supports previous reports that Thomas was initially a relatively slow language learner, although still falling within the normal range (Lieven et al., 2009). However, from age 3;6 to 3;11, Thomas' MLU corresponded more closely with that reported for the TD children by Rice et al. (2010). This was the period during which Thomas achieved stage V MLU values, indicating an acceleration in his syntactic development.

7.3.2. Maximum Length of Utterance

The aim of this analysis was to quantitatively substantiate qualitative changes observed at various points in the study, which were indicative of increasing syntactic length and complexity at specific points in time. Table 7.3 shows the results of the *Max Word* analysis; these figures give the length of the longest utterance which Thomas produced each month. Note that according to the Chat transcripts created by Lieven and colleagues, an utterance is considered to be akin to a sentence. This is demonstrated in the *Transcription Manual* section of the *Read me* document supplied alongside the transcripts; this states that each utterance must end with a punctuation mark indicating the end of a sentence: either a full stop, question mark or exclamation mark.

These results show increases in maximum sentence length throughout the study, although the pattern is not linear. From age 2;0 to 2;5, values fell between 5 and 12. Values increased from age 2;6 to 2;11, falling between 13 and 19. From age 3;0 to 4;0, values fell at 20 or above, with the exception of a value of 18 at age 3;2. The highest value of 40 was observed at age 3;8 and occurred in a string of narrative:

"You saw some snakes at the zoo and Mummy, you shouted "Daddy, I want to go look at the erm crocodiles", and then Mummy, one of the snakes come out of his home and you know what he done Mummy?".

Arguably, this utterance could be divided into a number of sentences, although it appears that it was transcribed as a single sentence, because of Thomas' linking of clauses with *and*.

Table 7.3. Thomas' Maximum Length of Utterance

Age	Maximum Length of Utterance (Words)
2;0	5
2;1	8
2;2	12
2;3	9
2;4	11
2;5	9
2;6	13
2;7	19
2;8	11
2;9	12
2;10	15
2;11	19
3;0	23
3;1	20
3;2	18
3;3	27
3;4	29
3;5	25
3;6	24
3;7	25
3;8	40
3;9	24
3;10	23
3;11	38
4;0	21

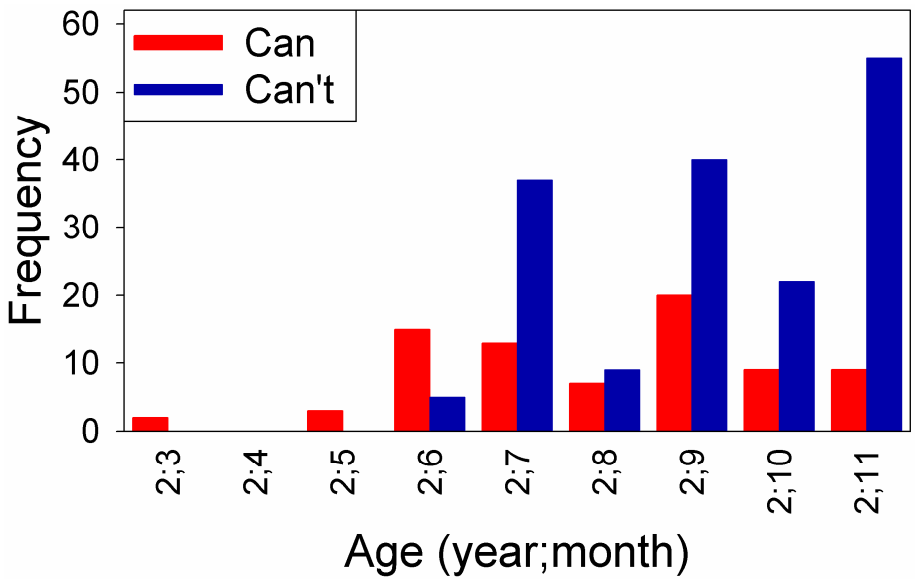
7.4. Thomas' Syntactic Development of Constructions Containing Can and Can't at Age Two

This section provides a mainly qualitative overview of Thomas' syntactic development at age two, drawing on the evidence from his constructions with the verbs *can* and *can't*. Information on constructions which do not contain potential assimilation sites has come from the complete analysis of Thomas' productions of *can* and *can't* at this age. Quantitative frequency counts for *can* and *can't* are also discussed.

7.4.1. Frequency of Occurrence of *Can* and *Can't*

Figure 7.1 summarises total frequency counts of *can* and *can't* in Thomas' language at monthly intervals during his third year. The graph begins at age 2;3, when *can* first emerged. (See Appendix Three, tables 52 to 59 for the daily frequency counts of *can* and *can't* for each month and Table 60 for a summary in tabulated form.) The graph shows increased usage of *can* and *can't* over time, although substantial fluctuation exists from one month to the next. Higher frequencies are evident for *can* than *can't* from age 2;3 to 2;6, but then frequencies of *can't* exceed those of *can* from age 2;7 until 2;11.

Figure 7.1. Frequencies of *Can* and *Can't* at Age Two



7.4.2. Emergence and Usage of *Can* and *Can't*

Thomas first used *can* in utterance-final position at age 2;3;10, in two occurrences of "I can", which were imitations of M's previous utterances. The next occurrences of *can* were sampled three months later in frequent repetitions of the formula "Yes I/we can" from age 2;5;10 onwards. This formula was evidently acquired from the *Bob the Builder* song (see Chapter Six for further details). No instances of *can't* were sampled during this period.

The verb *can't* emerged three months after the first occurrences of can at age 2;6. It firstly occurred alongside a main verb in a single imitative occurrence of "can't see it" at 2;6;12. At age 2;6;19, Thomas began to produce *can't* in utterance-final position in frequent productions of "I can't". These were apparently more productive and less formulaic than early occurrences of *can*; they were not immediate imitations of his mother, unlike the first productions of *can*. They were also contextually more congruous than the formula "yes we can". For instance, Thomas would spontaneously use "I can't", as a request for help. The verb *can* remained more frequent than *can't* at this age (see Figure 7.1).

From age 2;7 to 2;11, productions of "I can't" increased in frequency. Overall frequencies of *can't* were higher than those of *can* throughout this period (see Figure 7.1). In addition, *can* and *can't* both emerged alongside main verbs, in a wide range of constructions with the clause elements: subject (S), verb (V), object (O) and adverbial (A). Table 7.4 shows some examples, which have been selected in order to demonstrate the emergence of novel clause and phrase elements.

During this period, the auxiliary *can* was sampled in constructions with only five different main verbs: *see*, *hear*, *have*, *go* and *get*. In contrast, *can't* occurred in constructions with 17 different main verbs, including *see*, *hear*, *smell*, *reach*, *sleep*, *talk, open*, *close*, *find*, *sit*, *fix*, *get*, *do*, *put*, *blow*, *hold* and *remember*. The relative differences in both the frequencies of *can* and *can't* and the range of constructions in which they occurred indicate individual and

Table 7.4. Qualitative Advances in Thomas' Syntactic Development

Age of Occurrence	Utterance	Clause Level Analysis	Notes on Phrase Elements
2;7;1	"Can't reach"	V	First productive auxiliary *can't* + main verb
2;7;5	"I can't sleep"	SV	S restricted to one element
2;7;9	"Can't see you"	VO	O restricted to one element
2;8;24	"Not can't see now"	VA	Negator prior to auxiliary. A restricted to one element
2;9;3	"Can see flowers"	VO	
2;9;7	"I can see a little carrot"	SVO	Multiple elements in object noun phrase: indefinite article, adjective and noun
2;9;22	"I can't hear her"	SVO	First object pronoun other than *you*
2;9;28	"Bob Builder can't do Wendy's zip"	SVO	First usage of possessive noun
2;9;23	"I can't get it open"	SVOA	
2;9;25	"I can't get this postbox out"	SVOA	First definite article
2;10;21	"Can't see it under there"	VOA	First adverbial prepositional phrase
2;11;14	"Your (sic) can't see a big, huge box yet"	SVOA	Four elements in object noun phrase: indefinite article, two adjectives and noun. First *you* pronoun in construction with *can't*
2;11;18	"I can't see a rainbow up in sky"	SVOA	First adverbial with multiple prepositions
2;11;25	"You can go home now"	SVAA	First subject pronoun other than *I* in construction with *can*
2;11;6	"I can't put it on now fall down the floor again"	SVOAA VAAA?	Difficult to analyse. Possibly first attempt at coordination

differential acquisition of these verbs. The relatively high frequency and productivity of *can't* indicates that it was acquired in a more analytic way. In contrast, the relatively low frequency

and limited productivity of *can* indicates that this verb may have been learned in a more formulaic way, as part of the whole utterance.

In summary, it can be seen from these results that the latter half of Thomas' third year was characterised by major advances in the syntactic development of constructions containing *can* and *can't.* Thomas progressed from sentences of two or three words with only two clause elements (S and V) to sentences with three or four elements, consisting of S, V, O and A. . Object noun phrases began to extend beyond one element from age 2;9;7 onwards, to include articles and adjectives. Prepositions emerged in adverbials from age 2;10;21 onwards. By the end of this period, Thomas was also beginning to increase his usage of pronouns. The emergence and development of *can*, the acquisition of four-element sentences and the emergence of prepositions are all typical patterns of syntactic advancement during the latter half of the third year (Crystal, 1992; Buckley, 2003).

These observations are substantiated by the results of the quantitative analyses. The establishment of the auxiliaries *can* and *can't* from age 2;7 to 2;11 is classified by Brown as a stage III behaviour, and co-occurred with the progression in MLU from values typical of stage II to those typical of stages III and IV. Increases in both MLU and maximum length of utterance both reflect Thomas' increasing syntactic length and complexity over time, to include the range of novel clause and phrase elements described above.

7.5. Thomas' Syntactic Development of Constructions Containing Can and Can't at Ages Three and Four

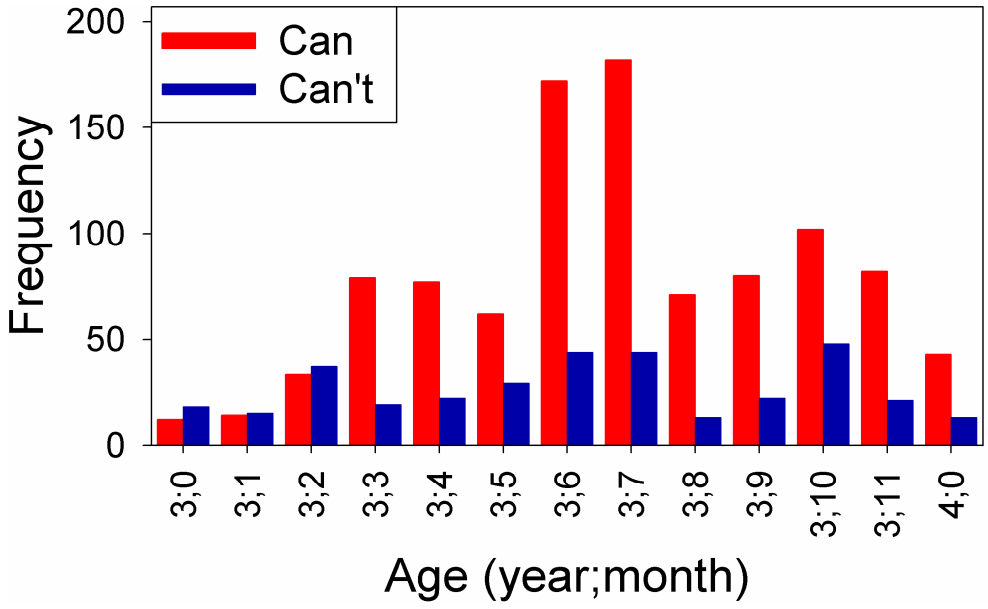
This section provides a mainly qualitative overview of Thomas' syntactic development at ages three and four, drawing on the evidence from his constructions with the verbs *can* and *can't*. Information on constructions which do not contain potential assimilation sites has come from

the control monthly sampling of all constructions containing *can* and *can't*. Quantitative frequency counts for *can* and *can't* are also discussed.

7.5.1. Frequency of Occurrence of *Can* and *Can't*

Figure 7.2 summarises total frequency counts for *can* and *can't* in Thomas' language at monthly intervals during his fourth year. (See Tables 53 to 65 in Appendix Four for the daily frequency counts of *can* and *can't* for each month and Table 66 for a summary in tabulated form.) These results show that frequencies of *can't* continued to exceed those of *can* at age 3;0 and 3;1. However, frequencies of *can* increased sharply from age 3;3 and remained higher than frequencies of *can't* throughout the remainder of the study.

Figure 7.2. Frequencies of *Can* and *Can't at Ages Three and Four*



7.5.2. Usage of *Can* and *Can't* from Age 3;0 to 3;6

Syntactic patterns at ages 3;0 and 3;1 remained similar to those observed in the latter half of Thomas' third year. The auxiliary *can't* continued to occur more frequently than *can* (see

Figure 7.2). Increases in Thomas' repertoire of main verbs, prepositions and adverbs were evident in the following utterances:-

- "I can't do it like that";
- "You can leave my toys on my shoulders like this";
- "Then, you can see him";
- "I can (.) make some room now".

Age 3;2 was characterised by substantial increases in overall frequencies of *can* and *can't* and the range of constructions in which they occurred. Frequencies of *can* and *can't* doubled, increasing from 14 and 15 respectively at age 3;1 to 33 and 37 respectively at age 3;3. (See below for a discussion of specific novel constructions, which produced potential assimilation sites.) The first instance of clause subordination was evident in the utterance "you know, he can get on there". Another development at age 3;2 was the onset of false starts and revisions. These occurrences indicate that Thomas' increasing syntactic creativity was challenging his linguistic processing capacity. The two instances which were sampled were:-

- "I get a, I, I can pl, Teletubbies can play with my fishing-rod";
- "Now I can't pl play anymore this".

Age 3;3 appeared to mark the start of a period of rapid syntactic growth and development. Frequencies of *can* doubled from 33 at age 3;2 to 79 at age 3;3. From this point until the end of the study, *can* continued to occur much more frequently than *can't* (see Figure 7.2). The quantitative findings show that MLU increased to stage IV values, where they remained until age 3;6. In addition, maximum length of utterance remained above 20 from this point until the end of the study.

7.5.3. The Development of Complex Sentences

The most noteworthy qualitative development at age 3;3 was the emergence of much clearer instances of complex sentences with clause coordination and subordination. This is typical during the first half of the third year (Crystal, 1992). There exist evident links between the emergence of specific coordinators and adverbials and the types of coordination and subordination which occurred. The emergence of coordination and subordination was most evident from age 3;3 to 3;7. Table 7.5 shows some examples, which have been selected in order to demonstrate the emergence of specific phenomena.

7.5.4. The Emergence of Tag Questions

A further development at age 3;3 was the emergence of tag questions. This corresponds with the age norms given by Buckley (2003). From age 3;3 to 3;5, "can't you?" was the only form sampled. However, a wider range of forms began to emerge at 3;6, including instances of grammatical mismatch. New developments were most noticeable during the period from 3;3 to 3;7. Table 7.6 shows some examples, which have been selected in order to demonstrate the emergence of specific phenomena in tag questions.

Table 7.5. Examples of Coordination and Subordination

Age	Utterance	Description of Phenomena
3;3;2	"Don't know why he can't get out today"	Subordinate clause qualifying the adverb *why*
3;3;4	"Then I can (.) just put on one stalk, then can sweep it/sweep in/sweeping"	First usage of *then* as a coordinator
3;3;4	"Oh no, I need that because you can be the lady looking for number one"	Subordinate clause, qualifying adverb *because*
3;3;4	"You can be a postman changing it"	Subordinate clause, qualifying object noun
3;3;6	"Erm, you think he can (.) cry and he get his hat back"	First usage of the coordinator *and*
3;4;0	"It means dustbin wagons just can't go through it, and cars, and big dustbin wagons"	Object clause. Coordination of noun phrases
3;4;3	"That's why I'm checking it it can go fast look!"	Subordinate clause qualifying the adverb *why*
3;4;3	"I can be a low wagon, driving a big man (be)cause I'm a dustbin xxx"	Two subordinate clauses.
3;5;2	"No I can't (be)cause I'm a busy man"	Emergence of "I can't, (be)cause", followed by a qualifying clause. Three instances at this age
3;5;5	"And you can't, (.) you can, (.) go (..) call for policeman can't you?".	Coordination of two verb phrases
3;6;0	"You want anymore, I can put it back where it goes"	Concurrent clause coordination and subordination. No coordinator
3;6;1	"When I'm better I can get it"	Adverbial clause at the beginning of the sentence
3;6;2	"It can fly on its own, but it doesn't fit in there"	First usage of *but* as a coordinator
3;7;1	"You can have that, you can be mummy, (.) and (0.5) you can put the light in there and I can be a man to sell something"	Coordination of three clauses
3;7;3	"When it's down there, I can't reach it can I"	First instance of concurrent subordination and tag question
3;7;4	"Erm (.) mm (.) this one could be away and you can crash couldn't you?"	First instance of concurrent coordination and tag question

Table 7.6. Examples of Tag Questions

Age	Utterance	Description of Phenomena
3;3;4	"There go flowers you can have a drink now, (…) can't you?"	First tag question sampled
3;6;0	"Then you can get on, (..) shall we?"	Nonmatching pronouns across declarative and tag.
3;6;0	"Mummy this can be a work can't it?"	First tag question with pronoun *it*
3;6;2	"I pull (th)em off can we"	Non-matching pronouns across declarative and tag
3;6;3	"I can be a co, I can be a man can't I?"	First tag with the pronoun *I*
3;6;4	"Erm you can put it through that little slot won't you?"	Non-matching verb forms across declarative and tag
3;6;4	"Then I can be, I can be a seller-man can't you?"	Non-matching pronouns across declarative and tag
3;7;1	"I can be a man a shopkeeper couldn't I?"	Non-matching verb forms across declarative and tag. First tag with the conditional *couldn't*
3;7;2	"Mummy, you can be sly fox can'tn't you?"	First of several tags with idiosyncratic form "can'tn't". Appears to indicate transition between the usage of *can't* and *couldn't*
3;7;2	"Mummy, you can cross the road with your cat can'tn't you?"	See above
3;7;3	"And now we can play outside can't we Farty-Pants?"	First instance with pronoun *we* matched across declarative and tag. First tag with vocative
3;7;3	"When it's down there, I can't reach it can I"	First instance of concurrent subordination and tag question
3;7;3	"You can be, you could be in that, couldn't you?"	Matching of conditional verb forms across declarative and tag, following a false start and revision
3;7;4	"Erm (.) mm (.) this one could be away and you can crash couldn't you?"	First instance of concurrent coordination and tag question

7.5.5. The Emergence of Interrogative Constructions

From age 3;4 onwards, Thomas began to produce interrogative constructions with *can* and *can't,* as is typical during the fourth year (Buckley, 2003). Steady increases in Thomas' usage of interrogatives were especially evident from age 3;6 until 4;0. It is interesting that the first question form to emerge, "can I roll it?" had the same syntactic structure as the formula "can we fix it?", which emerged as a formula eight months previously at age 2;8. Table 7.7 shows the emergence of different interrogative constructions during Thomas' fourth year. It is evident from close scrutiny of these data that only five out of the 22 interrogative constructions reoccurred at a later date.

Table 7.7. Developmental Trajectory of Interrogative Emergence

Age	Novel Question Forms with *Can*	Previously Encountered Question Forms with *Can*
3;4	"Can I roll it?"	"Can we fix it?"
3;5	None	None
3;6	"Can I have" (2), "can we have", "can I watch" (3)	None
3;7	"Can I see" (2), "can I keep"	"Can I have"
3;8	"Can you hold", "can I help"	"Can I have" (2)
3;9	"Can you hear", "can you look after"	"Can I have" (2)
3;10	"Can I mess", "can I play", "can I give", "can you have", "can you shut", "can I get"	"Can I have" (8), "can you hold"
3;11	"Can I weigh", "can I do", "can I eat"	"Can I have" (11), "can you have"
4;0	"Can you make", "can I go"	"Can I do", "can you look after"

7.5.6. Summary of Thomas' Syntactic Development at Ages Three and Four

It is evident that Thomas' syntax developed substantially during his fourth year. The first major changes occurred at age 3;2, when there were increases in both the overall frequencies of *can* and *can't* and the range of constructions in which they were sampled. Thomas also began to produce false starts and revisions at this age, which may indicate that his linguistic processing capacity was being challenged.

Further advances were evident at age 3;3. The frequency of *can* doubled again and *can* continued to occur much more frequently than *can't* from this point onwards. MLU increased to values typical of stage IV, where they remained for three months. Thomas' maximum length of utterance remained above 20 from this point until the end of the study. The most noteworthy qualitative developments in Thomas' syntactic structure were the emergence of complex sentences, subordinate clauses and tag questions. The emergence of interrogative constructions then followed at age 3;4.

The period from age 3;6 to 3;7 appears to have been the peak of Thomas' syntactic development of declarative constructions with *can* and *can't*. Frequencies of *can* reached their maximum values of 172 at age 3;6 and 182 at age 3;7. Frequencies of *can't* also reached the high value of 44 at both ages 3;6 and 3;7, although this was exceeded by a value of 48 at age 3;10. Age 3;6 was also the point at which MLU reached values typical of stage V, where they remained until age 4;0. Aged 3;6 also marked the start of the six-month interval during which Thomas' MLU values resembled those of TD children, rather than those of children with SLI, as reported by Rice et al. (2010). Further developments in coordination and subordination were also evident from age 3;6, as well as substantial increases in the number and range of tag questions. This period was then followed by a further change from age 3;8 to 4;0, when the most evident feature was the development of interrogative constructions with *can* and *can't*.

7.6. The Syntactic Development of Potential Assimilation Sites in Constructions Containing Can and Can't

7.6.1. The Syntactic Development of Potential Bilabial Assimilation Sites

Table 7.8 shows the emergence and usage of main verbs and a minority of other forms at potential bilabial assimilation sites with *can* and *can't* from age two to four years. If a construction occurred more than once at a specific age, this is indicated by a number in brackets following the construction.

Only three different main verbs contributed to potential bilabial assimilation sites at age two: *(re)member*, *blow* and *put*. The verbs *(re)member* and *blow* were realised with both assimilated and non-assimilated forms. All of these co-occurred in constructions with *can't*. No potential assimilation sites with *can* were sampled at age two. This reflects Thomas' higher frequency and more complex usage of *can't* than *can* during the latter half of the third year (see Appendix Three, tables 52 to 59 for individual frequency counts and Table 60 for a summary).

At least one novel construction was sampled monthly at potential bilabial assimilation sites from age 3;0 to 3;8. However, only those months in which major changes occurred are described in detail here.

Age 3;1 marked the emergence of potential assimilation sites in constructions with *can*: *can make*. As noted above, age 3;2 was characterised by substantial increases in overall frequencies of *can* and *can't*. There was also a qualitative increase in the range of constructions and main verbs with which *can* and *can't* co-occurred. These included *can put*, *can play*, *can't play* and *can't pay*. These patterns of increasing frequency and usage

Table 7.8. Constructions Producing Potential Bilabial Assimilation Sites

Age	Novel Constructions	Previously Encountered Constructions
2;10	"Can't (re)member" (4)	None
2;11	"Can't blow" (2), "can't put"	"Can't (re)member" (3)
3;0	"Can't bang"	None
3;1	"Can make"	None
3;2	"Can put", "can play" (2), "can't play", "can't pay"	"Can make"
3;3	"Can be" (18), "can't pinch" (1)	"Can play",
3;4	"Can mess", "can move", "can mix", "can broom", "can maybe"	"Can be" (9), "can make", "can put", "can't put"
3;5	"Can press"	"Can be" (14)
3;6	"Can build" (2), "can pinch" (3), "can pull", "can put/post", "can post" (2), "can pick up", "can't mum"	"Can be" (16), "can put" (8)
3;7	"Can bake" (2)	"Can be" (37), "can put" (6), "can pinch" (4), "can make" (2), "can play" (2), "can build/built" (6), "can't mum"
3;8	"Can push"	"Can be" (3), "can make", "can post", "can bake", "can put" (2), "can blow", "can't (re)member"
3;9	None	"Can be", "can pull" (2)
3;10	"Can (re)member"	"Can be", "can play" (2), "can put", "can mix",
3;11	"Can manage"	"Can be", "can make" (2),
4;0	"Can't manage"	"Can be" (3), "can't mum"

continued over the following months (see Appendix Four, Table 66 for a summary of *can* and *can't* frequencies at ages three and four, and Tables 7.5 to 7.7 above for examples of constructions which emerged). Consequently, tokens of *can* and *can't* were sampled at potential assimilation sites in all recordings during the period from age 3;2 to 3;8.

From age 3;3 until the end of the study at 4;0, *can* was sampled much more frequently than *can't* (see Table 66). This change is partly explained by the emergence of the construction *can be*, which was the most frequently occurring construction to produce potential bilabial

assimilation sites from age 3;3 to 3;7. Thomas used the construction *I/you can be (noun phrase)* in order to assign roles in imaginative play: for example, "you can be a bear and I can be a girl". Only one instance of the construction *(subject) can be (complement)* was sampled, which is more similar to adult usage of *can be*; this occurred at age 3;10;0 in the utterance "Stacking the saucers up can be great fun". Numbers and percentages of potential assimilation sites with the construction *can be* are as follows:

Age 3;3: 18 out of 20 sites (90%);

Age 3;4: nine out of 18 sites (52.9%);

Age 3;5: 14 out of 15 sites (94.3%);

Age 3;6: 16 out of 35 sites (45.7%);

Age 3;7: 37 out of 61 sites (60.7%).

There were several points at which many novel constructions emerged, showing evidence of increasing syntactic creativity. At age 3;4, these included *can move*, *can mess*, *can mix* and *can broom*, and the first construction with the adverb *maybe*. At age 3;6, six novel constructions were sampled, including *can pinch*, *can pull*, *can pick up*, *can post*, *can build* and *can't Mum*. This increased range of novel constructions at age 3;6 corresponds with a peak in Thomas' general syntactic development in constructions with *can* and *can't* at this age (see Chapter Six). The most frequent constructions which led to potential bilabial assimilation sites overall were *can be*, *can put*, and *can play*.

As the numbers of potential bilabial assimilation sites and proportions of assimilations reduced substantially from 3;8 onwards, there was less evidence of potential assimilation sites emerging from novel declarative constructions. Whereas 19 novel declarative constructions with *can* and *can't* were sampled at potential assimilation sites from 3;2 to 3;7, only four novel constructions were sampled from 3;8 to 4;0. The proportions of sites produced by the construction *can be* also decreased during this period, rising again at age 4;0. However, overall numbers of sites are too small to render these proportions statistically meaningful. Numbers and percentages are shown below:

Age 3;8: three out of eleven sites (27.3%);

Age 3;9: one out of three sites (33.3%);

Age 3;10: one out of six sites: (16.7%);

Age 3;11: one out of four sites (25%);

Age 4;0: three out of five sites (60%).

7.6.2. The Syntactic Development of Potential Velar Assimilation Sites

Table 7.9 shows the emergence of main verbs and a minority of other forms at potential velar assimilation sites with *can* and *can't* from age two to four years. If a construction occurred more than once at a specific age, this is indicated by a number in brackets following the construction.

Similarly to the findings reported for potential bilabial assimilation sites, five out of the seven potential velar sites sampled at age two occurred in constructions with *can't*, whereas only two occurred in constructions with *can*. Again, this reflects Thomas' more frequent and complex usage of *can't* than *can* at this age (see Appendix Three, Table 60 for a summary of frequency counts at age two). The constructions sampled at age two were *can go*, *can't close*, *can get* and *can't get*. The latter was most frequent, accounting for four out of seven sites at age two and also the two sites which occurred at age 3;0.

The increased frequency and usage of *can* and *can't* at age 3;2 led to a sharp rise in the number of potential velar assimilation sites sampled. The most frequently occurring constructions during this period were *can't get*, *can get*, *can't go* and *can go*. It is interesting that when assimilation emerged at 3;2, it occurred in two instances of the most frequent construction at this age: *can't get*. Assimilation emerged in constructions with *can* at 3;3, in *can go*. This corresponds with the age at which *can* became a more frequent lexical item than *can't*. The period from age 3;2 to 3;7 was characterised by rapid syntactic growth, with

Table 7.9. Constructions Producing Potential Velar Assimilation Sites

Age	Novel Constructions	Previously Encountered Constructions
2;8-2;11	"Can go", "can get", "can't get" (4), "can't close"	None
3;0	None	"Can't get" (2)
3;1	None	None
3;2	"Can count" (2), "can catch", "can't go" (2)	"Can go", "can get", "can't get" (9)
3;3	"Can cry", "can come"	"Can get", "can't get" (4), "can go", "can't go" (2)
3;4	"Can keep", "can collect"	"Can go" (6), "can't go", "can't get"
3;5	"Can't (be)cause" (3)	"Can go" (3), "can get" (2), "can't get" (3)
3;6	None	"Can go", "can't go" (3), "can get" (6), "can't get" (2), "can keep" (4), "can catch", "can't (be)cause"
3;7	"Can cut", "can cross", "can crash", "can crawl", "can't catch" (7)	"Can go", "can't go" (2), "can get", "can't get", (3), "can/can't get", "can collect", "can come" (2)
3;8	"Can (be)cause"	"Can go", "can collect"
3;9	"Can call" (2)	"Can go", "can cut" (2), "can't get"
3;10	None	"Can go", "can get", "can't get" (3)
3;11	"Can climb"	"Can go" (2), "can keep"
4;0	None	"Can go" (3), "can come", "can keep", "can't get"

relatively large monthly samples of potential assimilation sites, produced by 13 novel constructions. No novel constructions were sampled at age 3;6, when both numbers of potential sites and proportions of assimilations increased from ten at age 3;5 to 18 at 3;6. However, five novel constructions then emerged at age 3;7, the age at which both the sample of potential sites and the proportion of assimilations were highest. The most frequent constructions which led to potential velar assimilation sites overall were *can/can't get* *can/can't go* and *can keep*.

From age 3;8 to 4;0, the decline in numbers of potential velar assimilation sites and assimilations corresponded with a reduction in the number of novel constructions emerging at

potential assimilation sites. Only three novel constructions were sampled, compared with 13 in previous months. These findings are comparable with those for potential bilabial assimilation sites during this period.

There are several factors which may have contributed to the reduction in potential assimilation sites and assimilations during the period from age 3;8 to 4;0. The sample size was reduced in part by several factors concerning the audio recordings. One recording at 3;9 was missing from the corpus and one at age 3;11 was only partially transcribed. Only four recordings were available at age 4;0, one of which was cut short. These may therefore have contained focal utterances which could not be analysed.

Another factor is that from age 3;9 onwards, a degree of uncertainty arose concerning whether Thomas was actually saying *can* or *could* in some cases. This is because a number of utterances coded as containing either *can* or *could* in the Chat transcripts were realised with extremely similar phonetic forms, such as [k^{h}ə]. Contextual usage of these constructions was also indistinguishable: for instance "I can/could be …", "I can/could go…". This was a potential confound and two recordings were excluded from the analysis on these grounds: one at 3;9 and one at 3;10. However, these factors do not fully account for the reduction in potential assimilation sites. A total of six complete recordings simply did not capture any potential bilabial assimilation sites with *can* or *can't*: two each at ages 3;8, 3;9 and 3;11. Similarly, four recordings did not capture any potential velar assimilation sites: two at age 3;8 and two at 3;10.

This reduction in potential assimilation sites could not be explained by changes in overall frequencies of *can* and *can't*, which remained high throughout the study. It is therefore suggested that the reduction phase resulted from qualitative changes in Thomas' usage of *can* and *can't*. The great reduction in potential bilabial assimilation sites is linked to a decline in the use of the construction *can be* in imaginative play. There was also a decline in the range of novel declarative constructions producing potential bilabial and velar assimilation sites

during this period. A further change was Thomas' increased usage of *can* and *can't* in interrogative constructions, such as "can I" and "can you", which do not produce potential assimilation sites. (It is possible that constructions with *can we* produce potential assimilation sites, but assimilation of /n/ prior to /w/ has not been included in the current study.) It appears that the reduction in potential assimilation sites resulted directly from altered usage of *can* and *can't*, away from a repertoire consisting only of declaratives and tag questions, towards one which included higher numbers of interrogative constructions. (See section 7.5.5 for details of Thomas' interrogative development).

7.6.3. Summary of Thomas' Syntactic Development at Potential Assimilation Sites

There appear to be direct parallels between Thomas' syntactic development in constructions with *can* and *can't*, and the establishment of assimilation in these constructions. This is especially evident in Thomas' third year. Increases in the frequencies of *can* and *can't* at age 3;2 coincided with increases in the range of both potential bilabial and velar assimilation sites and the emergence of velar assimilation. The period during which bilabial assimilation became established, occurring at the majority of potential sites from age 3;3 to 3;7, corresponds exactly with the period in which syntactic advancements were most evident. The later period when velar assimilation became established, occurring at approximately half of potential sites, appeared to be especially dependent upon these advancements, as it corresponded exactly with the peak of this syntactic growth at age 3;6 and 3;7. The following reduction in potential assimilation sites and assimilations does not appear to indicate a regression in Thomas' syntactic development. Instead, it appears to reflect altered usage, towards more interrogative constructions with *can* and *can't*. It is also possible that Thomas' previous usage of *can* and *can't* in specific contexts had been replaced by other declarative constructions, which were not detected in the current analysis. For instance, the forms *can be* and *can go* appear to have been replaced to some extent by *could be* and *could go*, once

Thomas had acquired the conditional form. From age 3;8 onwards, weak forms of *can* and *could* became phonetically more similar and were therefore sometimes difficult to distinguish.

7.7. Conclusion

This chapter has investigated both the quantitative trends in Thomas' syntactic development and the specific qualitative advancements observable in his constructions containing *can* and *can't*. Thomas' increasing usage of *can* and *can't* over time in increasingly long and complex sentences and in a wider range of constructions is quantitatively substantiated by increases in his MLU, maximum length of utterance and frequency counts for *can* and *can't*. It is striking that *can't* occurred more frequently and productively than *can* in the latter half of Thomas' third year. However, frequencies of *can* became much more frequent than those of *can't* from age 3;3 until 4;0.

The interactions between Thomas' syntactic acquisition and assimilation development have also been explored. It appears that the establishment of assimilation (that is, its predominance at potential bilabial assimilation sites and its occurrence at approximately half of potential velar assimilation sites) was dependent on increased development and usage of constructions containing potential assimilation sites. When Thomas' language changed to include fewer of these constructions, assimilation in his speech declined.

Chapter Eight

The Role of Maternal Input and Interactional Context in Thomas' Assimilation Development

8.1. Introduction

The results discussed in Chapter Seven revealed that assimilation development in Thomas' speech was dependent on his usage of syntactic constructions giving rise to potential assimilation sites. The usage-based approach greatly emphasises the role of input on the child's language acquisition. The current chapter therefore aims to further explore the value of the usage-based approach in explaining these interactions between syntax and assimilation development, by investigating the occurrence of assimilation in Thomas' mother's (M's) speech. This chapter reports two analyses. Firstly, the occurrence of assimilation in Thomas' speech is compared with that occurring in M's speech. Secondly, Thomas' and M's realisations of potential assimilation sites in adjacent or near-adjacent utterances are compared, in order to investigate the immediate effect of interactional context on the occurrence of assimilation in Thomas' and M's speech. Pairs of utterances were identified in which both Thomas and M produced the same potential assimilation site in the same or an extremely similar utterance, within a couple of turns of each other. These pairs were then compared for matching of the segmental phenomena occurring at potential assimilation sites, as well as with prosodic phenomena. This analysis was carried out in order to investigate the possible links between assimilation and prosodic phenomena in interaction and their possible implications for Thomas's acquisition of assimilation.

8.2. The Occurrence of Assimilation in M's Speech

Appendix Five shows the phonetic transcriptions of M's potential assimilation sites in constructions containing *can* and *can't*, sampled at each of three points in time: T1, T2 and T3. T1 was sampled when Thomas was aged 2;6, shortly following the emergence of *can* and *can't* in his language. T2 was sampled when Thomas was aged 3;3, at the point when bilabial assimilation was becoming established and velar assimilation was emerging in his speech. T3 was sampled when Thomas was aged 4;0, at a point when there was a decline in both his usage of constructions giving rise to potential assimilation sites and proportions of assimilations relative to other phenomena.

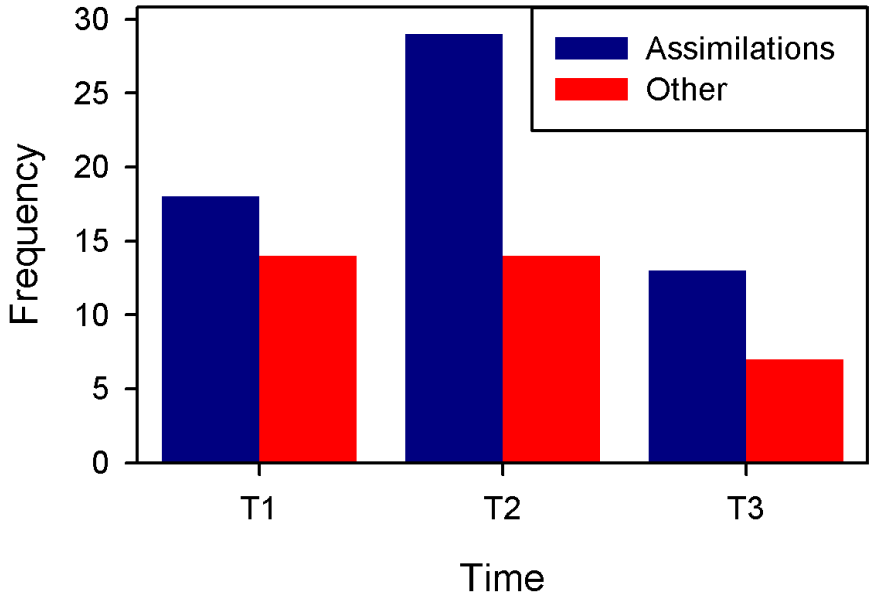
8.2.1. Overall Proportions of Assimilations

Table 8.1 shows the proportions of assimilations, other non-assimilation phenomena and open junctures sampled in M's speech at each of the three points in time. Figure 8.1 simply shows proportions of assimilations relative to all other phenomena occurring at potential sites.

Table 8.1. Proportions of Phonetic Behaviours Occurring at M's Potential Assimilation Sites in Constructions Containing *Can* and *Can't*

Time	Total Sites	Assimilations	Other Non-Assimilation Phenomena	Open Junctures	Untranscribable
T1	32	18	6	7	1
T2	43	29	3	9	2
T3	20	13	4	2	1

Figure 8.1. Proportions of M's Assimilations and Other Phenomena



Actual numbers of M's potential assimilation sites ranged from 20 to 43. The highest number of 43 was sampled at T2 and coincides with M's highest proportion of assimilations. This finding corresponds with the onset of bilabial assimilation establishment in Thomas' speech. The lowest number of potential assimilation sites sampled in M's speech was 20, occurring at T3, when the data sample was reduced to only four recordings. This corresponds with the period characterised by reductions of potential assimilation sites and assimilations in Thomas' speech. M produced assimilation at just over half of potential assimilation sites at T1 and showed a tendency towards predominance at T2 and T3. Percentages of M's assimilations relative to other phenomena at potential sites are as follows:

T1: 18 out of 32 sites (56.3%);

T2: 29 out of 43 sites (67.4%);

T3: 13 out of 20 sites (65.0%).

There exists no literature on proportions of assimilations occurring in adult speech with which to compare M's assimilation patterns. This is therefore the first normative study of assimilation patterns in adult speech, to the current author's knowledge. The nearest possible comparison is with the findings of Newton and Wells (1999), who studied older children aged

three to seven years. M's proportions of assimilations are similar to those observed by Newton and Wells, although they are slightly lower overall. Newton and Wells observed that assimilation occurred at between 73% and 76% of potential sites in children aged three to six, and at the reduced proportion of 54% of potential sites in seven-year-olds. However, these figures included assimilation of /t/, /d/, /n/ and /ʃ/. The current study is only concerned with the assimilation of /n/ and the /nt/ cluster. When Newton and Wells (1999) investigated the assimilation of /n/ in isolation, they found that assimilation occurred at between 74% and 78% of potential sites. These figures are considerably higher than those found in M's speech.

There are several possible reasons for the higher occurrence of assimilation in children aged three to seven than in the speech of an adult. Firstly, it is noteworthy that the results of Newton and Wells (1999) were obtained from elicited speech in sentence repetition and story re-telling tasks, as well as from samples of spontaneous speech. In contrast, the data for M in the current study were obtained entirely from spontaneous speech samples. These different methodologies may therefore limit the comparability of these two data sets. Secondly, it is possible that the differences between M and the participants of Newton and Wells resulted from individual differences between speakers. Thirdly, the assimilation data for M occurred in CDS and may therefore have contained lower proportions of CSPs, including assimilation, than her speech directed to adults, as similarly reported by Foulkes et al. (2005). Fourthly, it may be that further changes in the development of assimilation take place between seven years of age and adulthood, again, limiting the extent to which these findings should be compared. It would be necessary to investigate any changes in assimilation patterns occurring in late childhood and adolescence, in order to draw more definite conclusions. Additional studies on normative adult assimilation data would also be valuable.

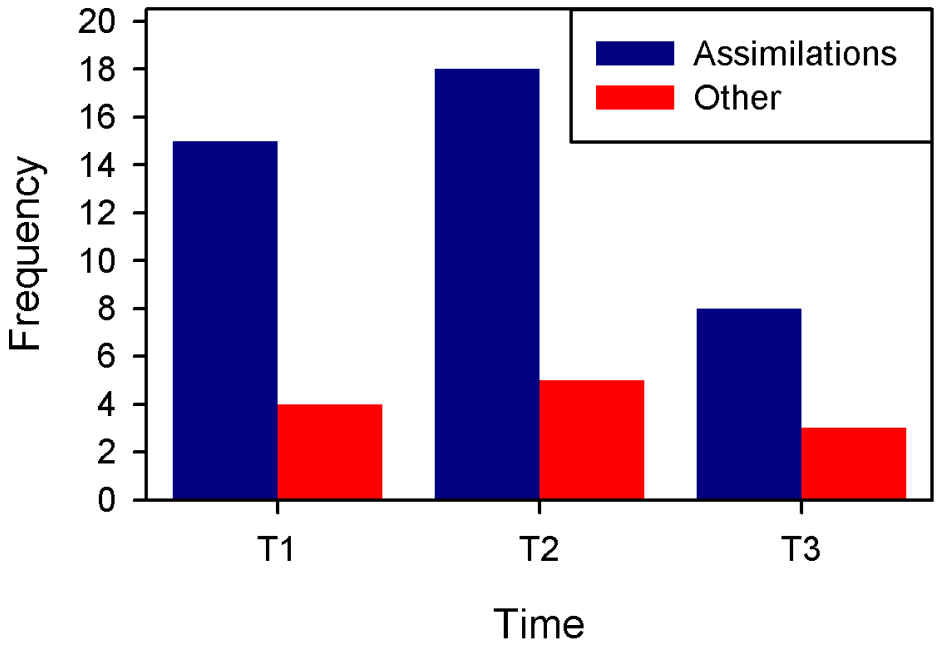
8.2.2. Bilabial Assimilation

Table 8.2 shows the proportions of assimilations and other phonetic phenomena occurring at potential bilabial assimilation sites in M's speech. Figure 8.2 simply shows proportions of bilabial assimilations relative to all other phenomena occurring at potential sites.

Table 8.2. Proportions of Phonetic Behaviours Occurring at M's Potential Bilabial Assimilation Sites in Constructions Containing *Can* and *Can't*

Time	Total sites	Assimilations	Other Non-Assimilation Phenomena	Open Junctures	Untranscribable
T1	19	15	1	2	1
T2	23	18	0	5	0
T3	11	8	1	1	1

Figure 8.2. Proportions of M's Bilabial Assimilations and Other Phenomena



Actual numbers of M's potential bilabial assimilation sites ranged from eleven to 23. The highest number of 23 sites occurred at T2, corresponding with the onset of bilabial assimilation establishment in Thomas' speech; the lowest number of eight occurred at T3, corresponding with Thomas' period of reduction. It can be seen that assimilation clearly predominated over other phenomena at all three points in time. Percentages of M's assimilations are as follows:

T1: 15 out of 19 sites (78.9%);

T2: 18 out of 23 sites (78.3%);

T3: eight out of eleven sites (72.7%).

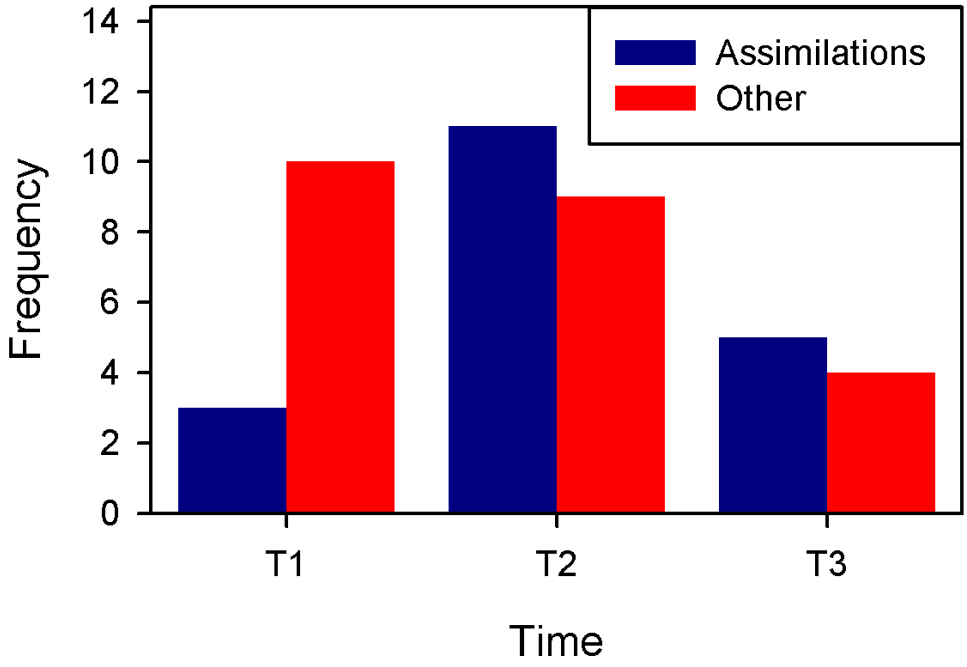
8.2.3. Velar Assimilation

Table 8.3 shows the proportions of assimilations and other phonetic phenomena occurring at potential velar assimilation sites. Figure 8.3 simply shows proportions of velar assimilations relative to all other phenomena occurring at potential sites.

Table 8.3. Proportions of Phonetic Behaviours Occurring at M's Potential Velar Assimilation Sites in Constructions Containing *Can* and *Can't*

Time	Total Sites	Assimilations	Other Non-Assimilation Phenomena	Open Junctures	Untranscribable
T1	13	3	5	5	0
T2	20	11	3	4	2
T3	9	5	3	1	0

Figure 8.3. Proportions of M's Velar Assimilations and Other Phenomena



As reported for bilabial assimilation, the highest number of potential sites was sampled at T2 and the lowest number was sampled at T3. Similarly to the findings observed for Thomas, M showed no clear pattern of velar assimilation predominance. At T1, velar assimilation was found as a minority behaviour, occurring at only three out of 13 potential sites (23.1%). It is striking that this single point at which M produced a clear minority of velar assimilations corresponds to a point prior to the emergence of velar assimilation in Thomas' speech. M produced assimilation at approximately half of the potential sites sampled at T2 and T3. Figures were eleven out of 20 sites (55.0%) and five out of nine potential sites (55.5%) respectively. This finding corresponds with the observation that Thomas produced assimilation at maximally half of potential sites during his period of velar assimilation establishment.

8.2.4. A Comparison of Assimilation in M's and Thomas' Speech

M's pattern of bilabial assimilation does not confirm the predictions stated in Chapter Five, based on previous research. It was predicted that M would produce more open junctures at T1

in order to optimise Thomas' learning of canonical (non-assimilated) forms. It was also predicted that M's proportions of assimilations would increase at T2 and T3, when simplification of child-directed speech (CDS) was no longer required. However, bilabial assimilation consistently predominated at all three points in time, with even a small decline at T3 owing to a reduced data sample size. It is therefore evident that M did not simplify her CDS by producing more non-assimilated forms at potential bilabial assimilation sites. In contrast, M's pattern of velar assimilation in CDS more closely confirms the predictions stated in Chapter Five. She produced only a minority of assimilations at T1, which then increased to approximately 50% at T2. However, there was no further increase in velar assimilation from T2 to T3.

It is noteworthy that M's proportions of bilabial and velar assimilations are similar to those found for Thomas during his establishment periods. M's consistent predominance of bilabial assimilation corresponds with Thomas' bilabial assimilation predominance from age 3;3 to 3;7. Similarly, M's production of velar assimilation at approximately half of potential sites at T2 and T3 corresponds with Thomas' proportions of velar assimilations at age 3;6 and 3;7. Moreover, T2 corresponds with a point shortly following the emergence of velar assimilation in Thomas' speech. These findings strengthen the hypothesis that M unconsciously adapted the realisations of her potential velar assimilation sites over time, in response to Thomas' increasing linguistic ability.

These results further support the conclusion drawn in Chapter Six, that bilabial and velar assimilation should be treated as separate phenomena. It has already been demonstrated that Thomas showed substantially different patterns of acquisition and proportions of occurrence for the two assimilation types. It also appears that M's proportions of assimilations in CDS reflect these differences.

The process by which Thomas learned bilabial assimilation appears to have been directly driven by the many exemplars in the input which he received from M. However, as M

initially produced only a minority of velar assimilations prior to Thomas' acquisition of velar assimilation, the question which remains is how Thomas learned to produce velar assimilation at all. It is proposed here that only a minority of exemplars was required in the input in order for Thomas to learn assimilations as acceptable phonological variants. Although M predominantly produced non-assimilated forms at potential velar assimilation sites at T1 and consistently produced a majority of bilabial assimilations, she also produced a minority of velar assimilations at T1 and a minority of non-assimilated forms at potential bilabial assimilation sites. It may therefore be that the extent to which a specific phenomenon occurs in the input determines the age at which it is acquired. Thus, the relatively high proportion of bilabial assimilations compared with velar assimilations in M's speech may explain why bilabial assimilation was acquired and established earlier in Thomas' speech than velar assimilation. M's speech thus reflected the fact that assimilations are acceptable forms, along with a range of non-assimilation and open juncture forms, but that assimilation is not obligatory.

The close correspondence between M's and Thomas' proportions of assimilations during Thomas' periods of assimilation establishment indicates that Thomas was able to learn and apply information on the possible acceptable realisations of potential assimilation sites, by mirroring the input which he received. It does not seem plausible that Thomas could have acquired this information through an innately specified linguistic rule. Although there are speech conditions in which close juncture forms such as assimilation are more likely to occur, such as spontaneous speech, communication of given information, fast speech rate, high frequency words and unstressed syllables (Farnetani and Recasens, 2010; Shockey, 2003; Wells, 1994), there are no clear-cut linguistic rules specifying when they are permissible and when they are not.

In summary, it is concluded that assimilation was acquired through a two-way interactional process between Thomas and M. Thus, Thomas acquired both types of assimilation from the exemplars present in the input which he received from M, while M increasingly produced

assimilation in response to Thomas' phonetic, phonological and linguistic advancement. This conclusion is compatible with the suggestion of Vihman et al. (1994), that children filter in those aspects of phonology from the input which match their motor schemes and filter out those which are not currently within their capabilities. Thus, it may be that Thomas' assimilation increased as his phonological processing abilities increased and he was able to filter in more exemplars of assimilation from the input. In turn, M responded to these changes in Thomas' speech by producing further velar assimilations.

8.3. Other Phonetic Phenomena Occurring at M's Potential Assimilation Sites

8.3.1. Non-Assimilation Close Junctures

All of M's potential assimilation sites which were classified into the *Other Non-Assimilation Phenomena* category were in fact realised with close juncture, whereas this category was more heterogeneous for Thomas. All except one non-assimilation close juncture in M's speech occurred in constructions with *can't*, for instance, "can't be" [ˈkʰɑ̃nʔ bɨ̆] and "can't come" [ˈkʰɑ̃nʔ ˈkʰɔ̃̈m]. Although non-assimilation close juncture forms of *can* were frequent in Thomas' speech, only one instance was observed in M's speech at T3: "can put" [xə ˈpʊʔ].

8.3.2. False Starts and Revisions

As discussed in Chapter Six, several instances of false starts and revisions were observed at potential assimilation sites in Thomas' speech. Some of these were characterised by different phonetic behaviours in the false start and revision, for example an assimilated false start and open juncture revision in earlier months, or an open juncture false start and assimilated revision in later months. Only one occurrence of a false start and revision was sampled in M's

speech. M's realisation was more similar to those found for Thomas in later months, consisting of an open juncture false start and assimilated revision: "and we can (.) and we can maybe" [xn̩] [kʰm̩ ˈmɛbˈbɪ].

8.3.3. Phonetic Forms of *Can*

In contrast with the findings observed for Thomas, all except one of M's unstressed productions of *can* were realised as weak forms. The only exception occurred when M was reading from a book, in the construction "can push" [kʰæ̃n pʊʃ]. It is noteworthy that weak forms were also observed in reading and singing contexts, for instance "can make" [kʰm̩ ˈmeɪʔk] and "can play" [kʰə̃m ˈpleɪj]. Weak forms were also found in contexts in which a pause occurred, for instance, "can (0.5) move" [kn̩ (0.5) ˈmuv], "can (.) pretend" [kʰn̩ (.). (C)ɪˈtɛ̃nd] and "can (0.5) perhaps" [kˣə̃n (0.5) pʰəˈhæps].

A further phenomenon observed in M's speech, which also occurred in Thomas' speech, was the realisation of weak forms of *can* with an initial velar or palatal fricative. Examples include "can go" [xŋ̩ gəʊ], "can put" [xə ˈpʊʔ] and "can play" [çə̃m pleɪ]. Forms such as these have been previously documented in adult speech (Cruttenden, 2001), and were also identified in Thomas' speech (see Chapter Six).

8.4. A Comparison of Thomas' and M's Most Frequent Constructions Containing Can and Can't

As noted in Chapter Six, the most frequent constructions and main verbs which produced potential bilabial assimilation sites in Thomas' speech were *can be*, *can/can't put*, and *can play*. These constructions were also found to occur most frequently in M's speech, along with the additional construction *can make*. However, *can put* was the most frequent construction for

M, compared with *can be* for Thomas. The main verbs *go* and *get* were the most frequent constructions which produced potential velar assimilation sites in M's speech, as also found for Thomas. However, whereas *get* occurred more frequently than *go* in Thomas' speech, *go* was most frequent for M. These findings correspond with those of Bybee (2010), who also found that the verbs *put*, *get* and *go* frequently occurred in constructions with *can* in adult speech.

Two different constructions were found in Thomas' speech which contained *can* followed by an adverb: *can maybe* and *can (be)cause*. These were also observed in M's language, with the addition of *can because* (producing a potential bilabial site), *can probably* and *can perhaps*.

Figures for the most frequent constructions contributing to potential assimilation sites in M's speech are shown in Table 8.4 below.

Table 8.4. Frequent Constructions Occurring at M's Potential Assimilation Sites

Construction	Number of Occurrences Sampled
Can put	13
Can't put	3
Can play	5
Can be	4
Can make	4
Can go	15
Can't go	3
Can get	6
Can't get	4

In addition to the frequent constructions common to both Thomas and M, M also used 15 constructions leading to potential assimilation sites, which were not sampled in Thomas' speech. These include *can probably*, *can pretend*, *can bounce*, *can't be*, *can't move*, *can buy*, *can perhaps, can pop*, *can't believe*, *can't make*, *can't quite*, *can't cut*, *can't give*, *can't come* and *can*, *can't you?*. Seven of these constructions specific to M were sampled at T3, along with three instances of *can't go*, which were used to forbid Thomas from doing something.

These differences in usage between Thomas and M would further explain the divergence of Thomas's and M's assimilation patterns from age 3;8 to 4;0. It may therefore be that the period of reduction in Thomas' assimilation development was not a regression away from producing assimilation per se, but instead resulted from two usage factors. Firstly, he began to produce fewer instances of the constructions which had previously led to many potential assimilation sites in previous months. Secondly, further constructions which led to potential assimilation sites for M had not yet emerged in Thomas' language. It is therefore possible that if it had been feasible to sample Thomas' language over a longer period, a re-emergence of potential assimilation sites and assimilations may have been evident, as Thomas acquired novel constructions.

It is noteworthy that M occasionally used *can* in a negative context instead of *can't*. This phenomenon has been noted in adult usage (Bybee, 2010). However, no examples of this were sampled in Thomas' speech. The three utterances in which M produced this phenomenon are:

"So that nobody can get to it unless you really need to use it";

"Just (be)cause I say you're not bouncing on the settee, doesn't mean you can bounce on me";

"Oh you no, I don't think you can go out with my shoes on".

8.5. The Occurrence of Assimilation in Thomas' and M's Speech in the Context of Interaction

Appendix Six shows the segmental and prosodic transcriptions for the adjacent or near-adjacent pairs of utterances identified, in which either Thomas repeated M's potential assimilation sites, or M repeated Thomas' potential assimilation sites.

8.5.1. Thomas' Repetitions of M

Table 8.5 summarises the segmental and prosodic matches and non-matches for Thomas' three repetitions of M. A tick indicates a match, while a cross indicates a non-match. (See Appendix Six for information on Thomas's age for each utterance pair.)

Table 8.5. Summary of Thomas' Repetitions of M

Utterance	Potential Assimilation Site Realisation	Stress Pattern	Locus of Tonic Syllable	Nuclear Tone
M: …"ˈI ˈcan ˆmix it" (high falling)…[ˈkʰæ…m ˈmɪks] Thomas: "ˌI ˌcan ˆmix it" [kʲʰæ̃m ˈmĩ̃ks]	✓	✓	✓	✓
M: "You˜ can get it ˈout" [ˈkʰæ̃ŋ gɛt] Thomas: "˜Can get it ˈout" [ˈt̪ʰæ̃n d̪ɛt̪]	×	✓	✓	✓
M: "You can (0.5) colˈlect it ˆnow,ˆ can't you" [xn̩ (0.5) kʰəˈlekˀt] Thomas: "ˈI can colÎect it, ˆnow" [kʰæ̃ŋ kˈɬˈlɛʔtˈ]	×	×	×	×

There were three instances in which Thomas repeated M's potential assimilation sites. Two of these were exact imitations, while the third was a repetition in a similar utterance (henceforth known as a similar repetition). The first exact imitation was sampled early on at age 2;8;28 in the utterance "˜can get it ˈout". Although both M and Thomas produced the utterance with the same stress pattern and with a falling-rising tone on *can*, the segmental realisations of M's and Thomas' potential assimilation sites did not match (compare [ˈkʰæ̃ŋ gɛt] and [ˈt̪ʰæ̃n d̪ɛt̪]). The second was sampled at age 3;4;2 in Thomas' repetition of the formula "ˌI ˌcan ˆmix it" from *Bob the Builder*. M initially produced this utterance very rhythmically, as if singing. In this instance, M's and Thomas' utterances were matched for assimilation and all prosodic phenomena. It is interesting that M then repeated this utterance again after Thomas, but with an open juncture realisation [ˈkʰæ̃n ˈmɪks].

It is possible that this change may have resulted from increased emphasis on M's part, although it is surprising, considering that Thomas' and M's repetitions of "I can mix it" became increasingly less rhythmic, more closely resembling natural speech. It would therefore be predicted that this apparently more natural form of speech would be more conducive to assimilation than open juncture (Shockey, 2003). Thomas' single similar repetition was sampled at age 3;4;3. There was no segmental or prosodic matching between M's and Thomas' utterances.

It is striking that both of Thomas' two exact imitations matched M's preceding utterances on all prosodic variables: stress pattern, locus of tonic syllable and nuclear tone. In contrast, the similar repetition did not match with M's utterance on any of these variables. On a segmental level, only one of Thomas' exact imitations matched M's utterance. Although the sample size of Thomas' repetitions is extremely small, it appears from these preliminary observations that there exists a difference between the characteristics of exact imitations and similar repetitions. Thomas' exact imitations appear to match M's utterances prosodically, although segmental characteristics of potential assimilation sites may vary. In contrast, the single similar repetition seemed more susceptible to segmental and prosodic variability compared with M's original utterance. However, in order to substantiate this claim, it would be necessary to analyse a much larger sample and wider range of Thomas' exact imitations and similar repetitions over time.

8.5.2. Segmental Matching and Non-Matching in M's Repetitions of Thomas

The majority of M's repetitions of Thomas matched Thomas' preceding utterances in terms of the segmental characteristics of potential assimilation sites. A total of 15 out of 22 matches were found (68.2%). Ten of these occurred at potential bilabial assimilation sites in

constructions containing *can* and five of these occurred at potential velar assimilation sites in constructions containing *can't*. Assimilation was the most common segmental phenomena on which pairs were matched, accounting for ten out of 15 pairs.

Nine out of the ten matched pairs found at potential bilabial assimilation sites were matched for bilabial assimilation. These assimilation matches spanned the age period from 3;1 to 3;10. An example of such matching is:

Thomas: "ʷMummy, (hiccup) (..) ˈI can be a ˆpostman (.) ˈcan't (.) (hiccup) ˆI?" [kxsə̃m bij]

M: "ˌyou can be a ˆpostman?" [kʰm̩ bij].

The tenth matched pair of utterances involving a potential bilabial assimilation site was matched for open juncture. This was sampled at age 3;7:

Thomas: "ˈYou can ˈbe a ˆlady-ˈman" [kə̈n ˈbij]

Thomas: "I can ˌbe a ˆbig/bin man" [kʰʌm̊ ˈBij]

M: "ˈI can be a ˆlady-ˈman? ˈWhat's a ˆlady-ˈman?" (Chuckling) [kʰə̃n bĭj].

Four out of the five matched pairs occurring at potential velar assimilation sites were matched for other non-assimilation phenomena, which were in fact, non-assimilation close junctures. The first three of these occurred at age two and were the first matched pairs sampled. A further instance was sampled later at age 3;7. However, actual phonetic realisations varied to some extent, for example the presence versus the absence of a final glottal stop in *can't*. An example is given below:

Thomas: "I ˌcan't get ˆthrough,˜now" [ˈkʰə̃n gɛˀ]

M: "You ˈcan't get ˇthrough?" [kʰə̃nˀ gɛˀ].

Only one of the matched pairs involving a potential velar assimilation site was matched for assimilation. This was sampled at age 3;7:

Thomas: "You ˈcan't catch ˈme (be)cause ˈI'm a ˆsausage-ˈman" [ˈkʰə̃ŋˀ kætʃ]

M: (laughs)

Thomas: (laughs)

M: “You ˈcan’t catch ˈme (be)cause I’m a ˈsausage ˆroll” [ˈkʰɑ̃ŋʔ kʰæt͡ʃ].

The presence of only one pair which matched for velar assimilation is accounted for by the fact that three of the matched pairs with potential velar assimilation sites were sampled at age two, prior to the emergence of velar assimilation in Thomas’ speech. These findings also reflect the relatively low occurrence of velar assimilation in both Thomas’ and M’s speech, compared with bilabial assimilation.

A total of seven pairs with non-matched segmental phenomena were sampled (31.8%), all of which occurred within the period from age 3;2 to 3;7. Six of these occurred at potential bilabial assimilation sites in constructions with *can*. In five out of six of these instances, M produced an assimilation, whereas Thomas produced three open junctures, one non-assimilation close juncture and one unclassifiable realisation. This pattern of phonological reduction in M’s assimilated repetitions is a typical phenomenon in adult discourse (for example, Bybee, 2002). This finding indicates that M was not consciously trying to teach Thomas, but was reproducing a more mature, less disjunct form of Thomas’s prior utterance in her CDS, as an unconscious aid to Thomas’s future learning. An example in which Thomas produced an open juncture and M repeated with bilabial assimilation is given below:

Thomas: “And˜then, we can ˈplay with a ˆtractor” [kʰæn ˈpleɪ]

M: “No don’t touch! Don’t touch anything now Thomas please! Don’t touch anything”

M: “˜Yes you can ˈplay with your ˈtractor, when you come ˈdown” [kn̩ ˈpleɪ].

In the final non-matched pair involving a potential bilabial assimilation site, Thomas produced assimilation and M repeated the utterance with open juncture, possibly for the purpose of emphasis:

M: * On ‘ˌBob the˜Builder’, they`say things like, “ˈI ˈcan˜roll it”, “ˈI ˈcan ˆmix it” (high falling), “ˌI ˌcan ˆdig it (low falling)” [ˈKʰæ…m ˈmɪks]

Thomas: “ˌI ˌcan ˆmix it” [kʲʰæm ˈmĩks]

M: "ˌI ˌcan ˆmix it" [ˈkʰæ̃n ˈmɪks].

There was only one non-matched pair which involved a potential velar assimilation site in a construction with *can't*. Thomas realised this with velar assimilation, whereas M repeated the utterance with non-assimilation close juncture:

Thomas: "Ah you ˌcan't ¯get ˈme" [ˈkæ̆ŋ gɛʔ]

M: "¯Thoˈmas!"

Thomas: (laughs)

M: "ˆWhy? Are ˌyou the˜gingerbread-man?"

Thomas: "ˆYeah" (laughs)

M: "Is˜that ˈwhy I ˈcan't ˈget ˈyou?" [ˈkʰæ̃nʔ gɛʔ].

8.5.3. A Comparison of Segmental Matching with Matching of Prosodic Phenomena in M's Repetitions of Thomas

Table 8.6 summarises the prosodic matches and non-matches for all utterance pairs which showed segmental matching of potential assimilation site realisations. Only the two focal utterances are shown for each pair. All non-focal utterances are shown in Appendix Six. A tick indicates a match, while a cross indicates a non-match. (See Appendix Six for information on Thomas's age for each utterance pair.)

Table 8.6. Prosodic Characteristics of Utterance Pairs with Matched Potential Assimilation Site Realisations

Utterance	Potential Assimilation Site Realisation	Stress Pattern	Locus of Tonic Syllable	Nuclear Tone
Thomas: "Erm ˈyou can ˈbe a ˈgirl (.) ˈfast aˆsleep" [kʰæ̃mi] M: "ˈI can be a ˈgirl ˈfast aˆsleep" [kʰm̩ bɪ]	✓	✓	✓	✓
Thomas: "ˌI can (.) ˌmake some˜room ˈnow" [ˈaɪ kʰæ̰m (.) meɪ]	✓	×	✓	✓
M: "You can ˌmake some˜room ˈnow?" [kʰm̩ ˈmeɪʔkˀ]				
Thomas: "˜Mummy, ˈyou can ˌbe a ˆxxx (one syllable), to ¯set off, with ˌmy ˆhat on" [kʰæ̃m ˈbij]	✓	×	✓	✓
M: "¯Sorry? ˌI can (0.5) be ˆwhat love?" [kʰm̩ (0.5) bi]				
Thomas: "˜Mummy, (hiccup) (..) ˈI can be a ˆpostman (.) ˈcan't (.) (hiccup) ˆI?" [kx̬æ̃m bij]	✓	×	✓	✓
M: "ˌyou can be a ˆpostman?" [kʰm̩bij]				
Thomas: "ˈYou can ˈbe a ˆlady-ˈman" [kæ̤̃n ˈbij]	✓	×	✓	✓
M: "ˈI can be a ˆlady-ˈman?..." [kʰə̃n bĭj]				
Thomas: "Can't ˆclose it ˈproperly" [kɑ̃nʔ ˈkləʊz]	✓	✓	✓	×
M: "You ˌcan'tˇclose it ˈproperly?" [kʰɑ̃nʔ ˈkləʊz]				
Thomas: "I ˌcan't get ˆthrough,˜now" [ˈkʰɑ̃n gɛ̰ʔ]	✓	✓	✓	×
M: "You ˈcan't get ˇthrough?" [kʰɑ̃nʔ gɛʔ]				
Thomas: …"You can play ˆout at ˈspringtime" [gʲə̃m pleij]	✓	✓	✓	×
M: …"ˌshe said you can playˇout in springtime" [kʰm̩ pleɪj]				
Thomas: "You can be a ˈpostman ˆchanging it" [ɣæ̃m bij]	✓	✓	×	×
M: * "You can be a ˈpostman preˈtending …" [kʰə̃m bɪ]				

Thomas: " ˇMummy, I ˌcan't go to ˆsleep.ˆMum" [kʰɑ̃nʔ gəʊ] M: "What do you ˆmean you ˈcan't go to ˈsleep?" [ˈkʰɑ̃n gəu]	✓	✓	×	×
Thomas: "When I finished ˈdoing this ˆprintings, ˈwe can (0.5) ˈplayˇ fire-engines" [kʰə̃m (0.5) ˈpleɪ] M: * "Oh when we've ˌfinished the ˜printing we can ˈplay ˈfire-ˈengines. ˆYes. We ˋcan." [kʰm̩ ˈpleɪ]	✓	✓	×	×
Thomas: "I ˌcan't ˌget it ˆopen" [ˈkʰɑ̃nʔ (C,Vd)ɛt] M: * "ˆNo and ˌI ˆI ˈcan't get it ˈopen" [ˈkʰɑ̃nʔ gɛt]	✓	×	×	×
Thomas: "ˌThen you can ˌput it ˆin here" [kæm }pʊt] M: "And we can ˌput it inˆthere ˌlikeˆthat" [kʰm̩ ˈput]	✓	×	×	×
Thomas: "You ˈcan't catch ˈme (be)cause ˈI'm a ˆsausage-ˈman" [ˈkʰɑ̃ŋʔ kæʧ] M: "You ˈcan't catch ˈme (be)cause I'm a ˈsausage ˆroll" [ˈkʰɑ̃ŋʔ kʰæʧ]	✓	×	×	×
Thomas: "ˈI can ˈbe a ˈbig ˆgrabber, ˆcouldn't ˈI?" [kə̃m ˈmij] M: "You˜can be. ˆYes. [kʰə̃m bi]	✓	×	×	×

There did not appear to be any particular link between matching of prosodic phenomena and segmental matching of potential assimilation site realisations. This finding indicates that M's usage of prosody was determined by discourse factors, such as questioning Thomas's prior utterance for clarification purposes, or emphasising a different part of the utterance. Exploration of these factors is beyond the scope of the current study.

When the 15 pairs with segmental matching were grouped according to the prosodic variables on which they commonly matched, distributions across the groups were fairly equal, with three or four pairs falling into most categories (see Appendix Six). Only one complete match was found, in which Thomas' and M's utterances matched in terms of all four variables:

potential assimilation site realisation, stress pattern, locus of tonic syllable placement and nuclear tone:

Thomas: "Erm ˈyou can ˈbe a ˈgirl (.) ˈfast aˆsleep" [kʰæ̃mi]

M: "ˈI can be a ˈgirl ˈfast aˆsleep" [kʰm̩ bɪ].

Four pairs matched on all variables except stress pattern. This occurred in instances when the overall stress and intonation patterns of the two utterances were similar, but there was some variation. An example is given below:

Thomas: "ˌI can (.) ˌmake someˇroom ˈnow" [ˈaɪ kʰæ̰m (.) meɪ]

M: "You can ˌmake someˇroom ˈnow?" [kʰm̩ ˈmeɪʔkˀ].

Three pairs were matched on all variables except nuclear tone. These intonational differences were apparently pragmatic in origin. However, a detailed study of the relationship between potential assimilation matching and the pragmatics of intonation is beyond the scope of the current study. An example is given below:

Thomas: "ˌCan't ˆclose it ˈproperly" [kɑ̃nʔ ˈkləʊz]

M: "You ˌcan'tˇclose it ˈproperly?" [kʰɑ̃nʔ ˈkləʊz].

A further three pairs matched in terms of potential assimilation site realisation and overall stress pattern, but did not match for locus of tonic syllable or nuclear tone. An example is given below:

Thomas: "You can be a ˈpostman ˆchanging it. " [ɣæ̃m bij]

M: "Alright. Yes."

Thomas: "ˈI can ˈbe a ˆparcel" [kʰæ̃m bij]

M: * "You can be a ˈpostman preˈtending (.) that ˈit's ˆlunchtime" [kʰə̃m bɪ].

Finally, four pairs only matched in terms of potential assimilation site realisation and did not match on any of the prosodic variables. An example is as follows:

Thomas: "I ˌcan't ˌget it ˆopen" [ˈkʰɑ̃nʔ (C,Vd)ɛt]

M: * “ˆNo and ˌI ˆI ˈcan’t get it ˈopen” [ˈkʰɑ̃nʔ gɛt].

8.5.4. A Comparison of Segmental Non-Matching with Matching of Prosodic Phenomena in M’s Repetitions of Thomas

Table 8.7 summarises the prosodic matches and non-matches for all utterance pairs which showed segmental non-matching of potential assimilation site realisations. Only the two focal utterances are shown for each pair. All non-focal utterances are shown in Appendix Six. A tick indicates a match, while a cross indicates a non-match. (See Appendix Six for information on Thomas’s age for each utterance pair.)

Similarly, there does not appear to be a link between matching of prosodic variables and segmental non-matching of potential assimilation site realisations. There were seven pairs of utterances which were non-matched in terms of segmental realisations of potential assimilation sites. One pair did not match on any variables:

Thomas: “ˆHey! ˈWe˜can, ˌmake˜stickers ˈnow. ˆStickers” [ˈkʰæ̃(Nas) meɪʔk]

M: [kn̆ ˈmeɪkˀ].

Two pairs matched only on stress pattern. An example is given below:

Thomas: “And˜then, we can ˈplay with a ˆtractor” [kʰæ̃n ˈpleɪ]

M: “No don’t touch! Don’t touch anything now Thomas please! Don’t touch anything”

M: “˜Yes you can ˈplay with your ˈtractor, when you come ˈdown” [kn̩ ˈpleɪ].

Two pairs matched only on locus of tonic syllable. An example is given below:

Thomas: “ˈThis can ˈbe a˜train-ˈspotter. ˆAnd }this” [kʰæ̃ bɨj]

M: “Well you can be a ˆplane-spotter, as ˌwell as a˜car-spotter” [kʰn̩ bɪ].

One pair matched on both stress pattern and locus of tonic syllable:

Thomas: "ˌI can mess ¯hair" [kʰæ̃n mɛs]

M: "ˌI can mess ˆhair" [kʰm̩ mɛs].

Table 8.7. Prosodic Characteristics of Utterance Pairs with Non-Matched Potential Assimilation Site Realisations

Utterance	Potential Assimilation Site Realisation	Stress Pattern	Locus of Tonic Syllable	Nuclear Tone
Thomas: "ˆHey! ˈWe˜can, ˌmake ˜stickers ˈnow. ˆStickers" [ˈkʰæ̃(Nas) meɪʔk] M: "We can ˌmake˜stickers now ˈcan we?" [kn̆ ˈmeɪk˺]	×	×	×	×
Thomas: "And˜then, we can ˈplay with a ˆtractor" [kʰæ̃n ˈpleɪ] M: "˜Yes You can ˈplay with your ˈtractor, when you come ˈdown" [km̩ ˈpleɪ]	×	✓	×	×
Thomas: "Ah you ˌcan't ¯get ˈme" [ˈkæ̆ŋ gɛʔ] M: "Is˜that ˈwhy I ˈcan't ˈget ˈyou?" [ˈkʰæ̃nʔ gɛʔ]	×	✓	×	×
Thomas: "ˈThis can ˈbe a˜train-ˈspotter. ˆAnd }this" [kʰæ̃ bij] M: "Well you can be a ˆplane-spotter, as ˌwell as a˜car-spotter" [kʰm̩ bɪ]]	×	×	✓	×
Thomas: "˜I ˈcan ˈbe ˈone, and ˈyou can ˈsit ˆnext to me, I can (0.5) ˈbe a ˆdriver" [kʰæn (0.5) ˈbij] M: * "Oh but ˈyou can be a˜driver, and ˌI'll lookˆout, and ˈtell you what we ˈhave to liftˆup" [xm̩ bij]	×	×	✓	×
Thomas: "ˌI can mess ¯hair" [kʰæ̃n mɛs] M: "ˌI can mess ˆhair" [kʰm̩ mɛs]	×	✓	✓	×
Thomas: "ˌI ˌcan ˆmix it" [kʲʰæ̃m ˈmĩks] M: "ˌI ˌcan ˆmix it" [ˈkʰæ̃n ˈmɪks]	×	✓	✓	✓

Finally, one pair matched on all prosodic variables. It is noteworthy that this single pair which showed matching nuclear tone was highly formulaic for both Thomas and M:

M: * "On ˌBob theˇBuilder, they ˋsay things like, " ˈI ˈcanˉroll it", "ˈI ˈcan ˆmix it" (high falling), "ˌI ˌcan ˆdig it (low falling)"" [ˈKʰæ…m ˈmɪks]

Thomas: "ˌI ˌcan ˆmix it" [kʲʰæm ˈmĩks]

M: "ˌI ˌcan ˆmix it" [ˈkʰæn ˈmɪks].

8.5.5. The Relationship between Segmental and Prosodic Matching Over Time in M's Repetitions of Thomas

When the adjacent and near-adjacent utterance pairs are grouped according to their common segmental and prosodic matches and non-matches, each group contains utterances sampled across a broad age range. It can therefore be concluded that grouping the data in this way does not reveal any changes in the combinations of segmental and prosodic matching and non-matching over time. Although there is no pattern of gradual increase in segmental matching on M's part over time, it is noteworthy that no segmental non-matches were sampled beyond age 3;7;1. The age ranges of utterances within each data group are as follows:

Matched for all segmental and prosodic phenomena: 3;3

Matched for all phenomena except stress pattern: 3;1 to 3;7

Matched for all phenomena except nuclear tone: 2;9 to 3;10

Matched for potential assimilation site realisation and stress pattern, but not matched for locus of tonic syllable or nuclear tone: 3;3 to 3;10

Matched for potential assimilation site realisation, but non-matched on all prosodic phenomena: 2;9 to 3;7

Non-matched for all phenomena: 3;2

Matched only for stress pattern: 3;3 to 3;7

Matched only for locus of nuclear tone: 3;3 to 3;6

Matched only for stress pattern and nuclear tone: 3;4

Non-matched for potential assimilation site realisation, but matched for all prosodic phenomena: 3;4.

It is possible that grouping the segmentally matched and non-matched pairs according to their commonly matching prosodic variables was not the most valid way to determine overall co-occurrence of segmental and prosodic matching and non-matching. For this reason, the extent of prosodic matching was recalculated, this time by looking at the extent to which matching of each individual prosodic variable co-occurred with segmental matching. These findings show that segmental matching co-occurred with matched stress patterns in seven pairs (46.6%), matched locus of tonic syllable in eight pairs (53.3%) and matched intonation in five pairs (33.3%). These findings show that segmental matching of potential assimilation sites co-occurred with matched stress patterns and matched locus of tonic syllable in approximately half of instances and co-occurred with matched nuclear tone in a third of instances. These data were not explored further using statistical correlations between segmental and prosodic variables, owing to the small sample size.

The same recalculation was carried out for the pairs which were non-matched in terms of segmental realisation of potential assimilation sites. The aim was to determine the extent to which segmental non-matching co-occurred with prosodic non-matching. It was found that segmental non-matching co-occurred with non-matched stress patterns in three instances (42.9%), non-matching locus of tonic syllable in three instances (42.9%) and non-matching nuclear tone in six instances (85.9%). These findings show a substantially high co-occurrence between non-matching potential assimilation site realisations and non-matching nuclear tone. However, it is impossible to establish any causal link between these two variables using this level of analysis. In order to investigate this further, it would be necessary to carry out a detailed study of the relationship between segmental non-matching at potential assimilation sites and the pragmatics of the non-matching nuclear tones within the local interactional

context, using techniques for the analysis of conversation. This research is beyond the scope of the current study.

8.6. Summary and Conclusion

This chapter has compared the occurrence of bilabial and velar assimilation in Thomas' and M's speech. The extent of segmental matching of potential assimilation site realisations in adjacent utterance pairs has also been explored, in order to investigate the extent to which Thomas' and M's phonetic behaviours may have influenced each other within an interactional context. In addition, segmental matching and non-matching of potential assimilation sites were compared with matching and non-matching of prosodic phenomena.

Similarities between Thomas' and M's proportions of assimilations at potential sites strongly suggest that Thomas learned assimilation from the input which he received from M. However, this is not immediately apparent from the segmental and prosodic characteristics of the three adjacent and near-adjacent pairs of utterances, in which Thomas reproduced M's potential assimilation sites. The two exact imitations both matched M's utterances on all prosodic phenomena, but only one of these matched M's utterance in terms of segmental realisation of the potential assimilation sites. The single similar repetition did not match M's preceding utterance on any segmental or prosodic level. It appears from these preliminary findings that Thomas' exact imitations were more likely to share common segmental and prosodic characteristics with M's preceding utterances than his similar repetitions. However, the complete non-matching found in the similar repetition may still provide evidence that Thomas acquired assimilation from the input which he received. The fact that he produced velar assimilation following M's open juncture realisation indicates that Thomas had acquired velar assimilation and was perhaps less reliant on the immediate model in order to apply it appropriately. Finally, it must be remembered that the sample of Thomas' imitations and repetitions was too small to draw definite conclusions or make any strong claims.

M's adaptation to Thomas' ability to produce assimilation is more evident from the analysis of M's imitations and similar repetitions of Thomas. M's Realisations of potential assimilation sites matched those of Thomas in 15 out of 22 utterance pairs (68.2%). It is noteworthy that no segmental non-matches were sampled beyond age 3;7;1, which corresponds with Thomas' establishment periods for both bilabial and velar assimilation.

As concluded from the comparison of Thomas' and M's proportions of assimilations, these interactional findings also show more adaptation in M's speech at potential velar assimilation sites than at potential bilabial assimilation sites. M showed relatively consistent production of bilabial assimilation at potential sites over time in both matched and non-matched utterance pairs. The increased matching found at potential bilabial assimilation sites from age 3;7;2 is therefore attributed to increases in Thomas' bilabial assimilation, rather than adaptations in M's speech. Similarly, the single instance of velar assimilation matching occurred at age 3;7, during Thomas' velar assimilation establishment period. In contrast, the other four pairs involving a potential velar assimilation site were matched for non-assimilation close juncture. Three of these were sampled at age two, prior to Thomas' acquisition of velar assimilation. It therefore appears from this small sample that M unconsciously adapted her speech to match Thomas' phonological ability. This conclusion is supported by the finding that M produced only a minority of velar assimilations at T1.

Taken together, it can be concluded from these comparisons of Thomas' and M's speech that there existed a two-way interaction between Thomas' acquisition of assimilation and M's realisation of potential assimilation sites, as M adapted her speech to Thomas' phonological ability. M consistently produced many bilabial assimilations, which enabled Thomas to acquire bilabial assimilation shortly following the emergence of potential sites in his speech. It was therefore unnecessary for M to adapt her speech by producing more non-assimilated forms at potential bilabial assimilation sites. On the other hand, Thomas was slower to acquire velar assimilation, owing to a number of possible factors. Firstly, velar nasals present more motoric difficulty for children than bilabial nasals (for example, Newton and Wells,

2002; Dodd et al., 2003). There also exist cognitive factors and semantic factors relating to the specific words and constructions used at potential bilabial and velar assimilation sites, which are beyond the scope of this study. It appears that M adapted to Thomas's slower velar assimilation development by producing more non-assimilated forms at potential sites. However, it may also be that fewer occurrences of velar than bilabial assimilation in M's CDS further contributed to Thomas's slower velar assimilation development. It is therefore concluded that input frequency was an important contributing factor in Thomas's assimilation development.

There was no evidence that M produced prosodic matching of Thomas' speech in line with her segmental matching at potential assimilation sites. When the utterances were grouped according to phenomena on which they commonly matched or did not match, fairly equal numbers of pairs fell into each group. When the co-occurrence of segmental and prosodic matching was investigated, segmental matching was only found to co-occur with stress pattern and tonic syllable placement in approximately half of instances. There was a high co-occurrence of segmental non-matching and non-matching of nuclear tone (85.9%). However, there were only seven utterances which were non-matched for segmental realisation of potential assimilation sites. This sample is therefore too small for these percentages to be considered statistically meaningful. It would be necessary to conduct a detailed analysis of pragmatic and conversational factors in order to further investigate this co-occurrence. It is therefore concluded that the prosodic characteristics of M's utterances containing potential assimilation sites were governed by pragmatic and conversational factors, which are beyond the scope of the current study.

Chapter Nine

Discussion

The current study aimed to investigate the development of between-word assimilation at potential assimilation sites formed by the auxiliary verbs *can* and ca*n't*, in one typically developing child, Thomas, from age two to four years. It further aimed to investigate possible interactions between assimilation development and the acquisition of syntax. General quantitative advances in Thomas' syntactic development were investigated, as well as specific qualitative patterns of emergence in constructions containing *can* and *can't*, which were the primary focus of the current study.

In order to further investigate the potential influences on Thomas' assimilation development, patterns of assimilation were also measured for Thomas' mother (M) at three different points in time, which corresponded to three distinctive phases in Thomas' assimilation development. Thomas' and M's proportions of assimilations at potential sites were then compared. A more fine-grained interactional analysis was then carried out, which compared the segmental and prosodic characteristics of Thomas' and M's utterances containing potential assimilation sites in portions of interaction, when either Thomas repeated M's potential assimilation site in an identical or similar utterance, or (more often) when M repeated Thomas' potential assimilation site in a similar utterance.

9.1. Thomas' Acquisition and Usage of Can and Can't

The verbs *can* and *can't* appear to have followed different patterns of acquisition in Thomas' speech. *Can* was acquired earlier and in more formulaic utterances than *can't*, emerging from age 2;3 to 2;6. In contrast, when *can't* emerged three months later at 2;6, it appeared to occur more productively, with contextually congruent usage. When *can* and *can't* emerged alongside

main verbs from age 2;7 to 2;11, frequencies of *can't* became higher than those of *can*, and *can't* occurred in a much wider range of constructions than *can,* with 17 different main verbs. Usage of the auxiliary *can* was restricted to constructions with only five different main verbs. Phonetic forms of the auxiliary *can* were weak and showed high variability. These patterns were observed throughout the latter half of the third year. These differences indicate that *can't* was acquired more analytically as a single syntactic element and was therefore rapidly applied to a wide range of syntactic contexts. In contrast, *can* was acquired holistically within formulaic constructions and therefore could not initially be applied productively to novel contexts.

The different patterns of acquisition for *can* and *can't* support the individual acquisition of different verb forms, irrespective of their similarity in the adult grammar. This view is in line with the usage-based constructivist approach to language acquisition, which proposes that acquisition of verb forms is initially lexically-specific and limited to specific constructions; this learning is therefore not immediately generalised to other similar verb forms and syntactic structures. Generalisation only occurs when the child abstracts schemas for usage of specific constructions and lexical items, leading to less restricted and more generalised usage (Lieven and Tomasello, 2008). Specific evidence for this developmental trajectory has been found in relation to auxiliary verbs (Ambridge et al., 2006B; Rowland et al., 2005). Abstraction of the auxiliary *can* appeared to occur for Thomas at age three, at the point at which unstressed forms with full vowels emerged and occurred in an increasing range of constructions. The phonetic and phonological patterns of the strong and weak forms of *can* are discussed in more detail in a later section. The usage-based approach also greatly emphasises the importance of input frequency in determining a child's pattern of language acquisition (Lieven and Tomasello, 2008). A future study of frequency and usage of *can* and *can't* in M's speech would therefore be useful, in order to further investigate the extent to which the usage-based approach accounts for the current data.

The different emergence patterns for *can* and *can't* also support the proposal that analytic and holistic learning may be evident in the same child at different points in time, or may occur in parallel across different linguistic domains (Peters, 1977; Nelson 1981; Peters (1995). Analytic acquisition of *can* appears to have begun from age 3;1 to 3;2, when forms with a full vowel (full forms) emerged and abstraction of schemas for more generalised usage was evident. It therefore appears that the notion of analytic versus holistic language learning can be incorporated into the usage-based approach, to account for the learning processes which underpin initial specificity of constructions and later abstraction of schemas. Holistic acquisition of formulae leads to initial lexical specificity, while the transition to a more analytic learning style results in abstraction of schemas and generalisation of usage.

The opposite patterns of emergence of *can* and *can't* were observed by Richards (1990). He reports that *can't* was acquired earlier, occurred with a more limited range of main verbs and was used in more stereotyped utterances than *can* (Richards, 1990). However, one similarity between Thomas and Richards' participants is that when *can't* emerged, it began to occur more frequently than *can*. Richards concluded that the earlier acquisition of *can't* than *can* in his participants violated the complexity principle of nativist theories, i.e. the notion that grammatical forms should be acquired in order of complexity, starting with the simplest. On the other hand, the reverse patterns found for Thomas would initially appear to confirm the complexity principle. However, Thomas and Richards' participants could both be considered to have acquired these verbs in order of complexity, if complexity were to be viewed as the extent of syntactic productivity, rather than complexity of the internal grammatical structure. Thus, it appears that all of the children firstly learned the verb form which occurred more formulaically in their linguistic environment, before acquiring the second form, which occurred in fewer formulae and was learned more analytically.

The main finding common to both studies is that the children showed different patterns of acquisition for *can* and *can't,* indicating that they were acquired independently, as separate lexical items, in different usage contexts. This developmental pattern indicates that it is not

useful to view development in terms of the complexity principle. Although *can't* is a grammatically more complex form than *can* in adult language, containing an additional grammatical element (the negative morpheme *n't*), one cannot assume that the units identified in adult language have psychological reality or relevance for the child in early stages of language acquisition (Peters, 1983). A similar observation has been made with regard to the forms *wanna* and *gonna* which have become single lexical items and linguistic processing units in English over time, and which children initially learn without the knowledge that they each comprise two component parts: *want* and *to*, or *going* and *to* (Vihman, 1982; N. Ellis, 2002). The same is likely to apply in the case of *can't*, which children learn without having analysed its two component parts: *can* and *not*.

The individual differences observed between Thomas and Richards' participants can be interpreted in terms of individual variability between children and differences in the linguistic input which they received. For example, the earlier emergence of *can* in Thomas' speech was linked to the utterances "yes we can" and "can we fix it?" from the British television cartoon, *Bob the Builder*. It is evident from the context provided in the audio recordings that Thomas watched this cartoon and sang the song with his mother, showing evidence of high exposure to and usage of these utterances as formulae. These utterances were identified as formulae according to many of the criteria specified by Peters (1983); the utterances were used repeatedly without any alteration to the forms; the utterance "can we fix it" involved an interrogative construction with subject-auxiliary inversion, which did not emerge productively with other main verbs in Thomas' language until age 3;4; the utterances were phonologically coherent, with little phonological alteration across forms (except for some initial alveolars in *can*) and smooth intonation contours; the formulae existed for other members of the child's community, occurring in a popular television programme; the usage of these utterances was sometimes contextually incongruous.

Further evidence of the influence of maternal input on Thomas' usage of *can* and *can't* comes from a comparison of the constructions in which *can* and *can't* occurred, in both Thomas' and M's language. Analyses of the usage of *can* and *can't* in constructions leading to potential assimilation sites revealed striking similarities between Thomas' and M's usage. The most frequent main verbs contributing to potential assimilation sites in both Thomas' and M's speech were *be*, *put* and *play*. The only difference in relative proportions is that Thomas showed higher frequencies of *be*, whereas M showed higher frequencies of *put*. This is explained by Thomas' highly specific usage of *can be* to assign roles in imaginative play. In addition, M frequently used the construction *can make*, which was infrequent in Thomas' speech. The most frequent verbs occurring at potential velar assimilation sites in both Thomas' and M's speech were *get* and *go*. Whereas *get* was more frequent for Thomas, *go* was more frequent for M. Interestingly, Bybee (2010) found *put*, *get* and *go* to be the most frequent verbs co-occurring with *can* in a large-scale analysis of adult conversation.

In conclusion, the different patterns of emergence observed for *can* and *can't* in Thomas' language resulted from a combination of differences in both learning style and the type of input which he received. This conclusion corresponds with the results of empirical evaluations of the usage-based approach, which have found significant correlations between input frequency and age of auxiliary acquisition, as well as individual differences in the exact patterns of auxiliary emergence (Lieven, 2008; Theakston and Lieven, 2005).

9.2. The Relationship between Thomas' Assimilation Development and his Acquisition of Syntax in Constructions Containing Can and Can't

Bilabial assimilation emerged at a minority of potential sites in Thomas' speech at age 2;10. The first example occurred only five days following the emergence of the first construction containing a potential assimilation site: "I can't (re)member". Bilabial assimilation continued

to be produced at a minority of potential sites from this age until age 3;2. This was followed by a period of substantial assimilation development from age 3;3 to 3;7, when both numbers of potential assimilation sites and proportions of assimilations increased, leading to predominance of assimilation over other phenomena at potential sites. Assimilation predominance peaked at ages 3;6 and 3;7, occurring at 85.7% and 91.5% of potential sites respectively. This period of predominance is viewed as the phase during which assimilation became established as a phonological phenomenon.

In contrast, velar assimilation emerged much later than bilabial assimilation and showed less evidence of clear predominance over other phenomena at potential sites. Although constructions containing potential velar assimilation sites emerged from age 2;8, (mostly involving the verb *get*), velar assimilation did not occur until six months later at age 3;2. Velar assimilation then occurred at a minority of potential sites from age 3;2 until 3;5. Similarly to the findings for bilabial assimilation, velar assimilation then peaked at ages 3;6 and 3;7, occurring at 47.1% and 61.9% of potential sites respectively. This is viewed as the point at which velar assimilation became established as a phonological phenomenon. Velar assimilation showed a tendency towards predominance only at age 3;7, whereas bilabial assimilation showed predominance from age 3;3 to 3;7, with especially high proportions from age 3;4 to 3;7. It is noteworthy that close juncture forms predominated over open junctures at both potential bilabial and velar assimilation sites throughout the study, with only two exceptions for potential velar assimilation sites at ages 3;4 and 3;9.

The period during which each assimilation type became established also appears to have been the point at which Thomas learned to apply assimilation alongside other phonetic phenomena as they emerged. For example, weak forms of *can* re-emerged at age 3;5, when bilabial assimilation was being established. Consequently, assimilation immediately began to occur at potential sites with weak forms from the point of their re-emergence. In contrast, although weak forms of *can* occurred at potential velar assimilation sites from age 3;5, assimilation was not observed at these sites until the onset of its establishment at age 3;6.

There exist striking parallels between Thomas' assimilation development and his acquisition of syntax. During the period from 2;8 to 3;1, when only a minority of bilabial assimilations and no velar assimilations occurred, Thomas produced potential assimilation sites in constructions with only eight different main verbs: *(re)member*, *blow*, *put*, *bang*, *make*, *get*, *go* and *close.* This finding is not surprising, considering that Thomas was still acquiring the rudiments of syntax, including auxiliary verbs at this stage. He acquired a wide range of nouns, pronouns, verbs, adverbs, adjectives and prepositions and learned to combine these in sentences comprising the clause and phrase elements subject, verb, object and adverbial.

The period during which assimilation became an established phenomenon in Thomas' speech corresponds with the age at which many major quantitative and qualitative advances in his syntactic development took place. Substantial syntactic changes began to occur at age 3;2, slightly ahead of assimilation establishment. Age 3;2 was characterised by increases in both the frequencies of *can* and *can't* and the range of constructions with which they occurred. Thomas' mean length of utterance in morphemes (MLUm) increased to values typical of Brown's stage IV and Thomas' maximum length of utterance remained above 20 from this age onwards. These syntactic advancements explain why larger numbers of potential assimilation sites were found in a much wider range of constructions at this age, compared with earlier ages. Whereas many recordings from age 2;8 to 3;1 yielded no constructions containing potential assimilation sites with *can* or *can't*, these constructions were encountered in every recording from age 3;2 until 3;8. It is also noteworthy that velar assimilation emerged in Thomas' speech at age 3;2, in two instances of the most frequent construction leading to potential velar assimilation sites: *can't get*.

These syntactic developments at age 3;2 appear to have been the precursor for the onset of bilabial assimilation establishment at age 3;3. This establishment appears to be linked with the emergence of the construction *can be*, which Thomas frequently used from age 3;3 to 3;7, in order to assign roles in imaginative play. Further milestones in syntactic development were

evident alongside assimilation establishment. Whereas *can't* occurred more frequently than *can* from age 2;6 until 3;2, *can* occurred much more frequently than *can't* from age 3;3 onwards. Age 3;3 also marked the emergence of velar assimilation in constructions with *can.* Complex sentences and tag questions were evident from age 3;3. Interrogatives emerged at age 3;4 and novel interrogative constructions continued to emerge throughout the remainder of the study, especially from age 3;6 onwards. Ages 3;6 and 3;7 appeared to have marked the peak of Thomas' syntactic development in constructions with *can* and *can't*, as shown by an increased range of declaratives, tag questions and interrogatives and maximum frequencies of *can* and *can't.* In addition, Thomas' MLU reached stage V values and for the first time, was comparable with the norms reported by typically developing children (Rice et al., 2010). This peak of syntactic development corresponds exactly with Thomas' peak of assimilation establishment, when the highest numbers of potential assimilation sites were produced, and maximum numbers of bilabial and velar assimilations were identified.

It appears that the establishment of velar assimilation was especially dependent on advances in Thomas' syntactic development, following its later emergence and slower rate of progression. Thus, the optimum period for velar assimilation to become established occurred at ages 3;6 and 3;7, when Thomas was producing the maximum number of constructions containing potential assimilation sites. The relatively low level of velar assimilation predominance at age 3;7 (61.9%) indicates that velar assimilation never became as well established as bilabial assimilation during the period studied.

The parallels between increased assimilation and syntactic growth during this period can be explained in terms of changes in Thomas' language learning style (Perkins, 1999; Wray and Perkins, 2000). After Thomas had analytically acquired both *can* and *can't* as individual lexical items, he then adopted a more mature formulaic strategy. He was thereby able to store and retrieve frequent collocations for more efficient usage. This would also explain the re-emergence of weak forms from age 3;5 (see below).

It is proposed that Thomas' holistic storage of high frequency utterances in the lexicon led to phonetic changes in their production, in line with the usage-based, phonological approach of Bybee (2002, 2006). It is suggested that Thomas stored the neuromotor routines and retrieved articulatory gestalts for the production of high frequency constructions leading to potential assimilation sites, such as *can be*, *can/can't get* and *can/can't go*. Increased practice of these utterances through usage resulted in greater fluency and increased overlap between articulatory gestures (Bybee, 2002, 2006). This would account for the increases in assimilation observed over time, as well as the [m] gemination found in a minority of productions of the most frequent construction *can be*. However, it is also noted that open junctures and non-assimilation close juncture forms continued to occur in Thomas' speech alongside assimilations, even in these highly frequent utterances. This indicates that Thomas continued to undergo the process of forming holistic neuromotor routines for these utterances throughout the study. During this process, he sometimes produced a lesser degree of gestural overlap, leading to open juncture; at other times, the gestural overlap was greater than that required to produce assimilation, leading to phenomena such as the lack of a final nasal in *can* and nasal gemination in *can be*.

From age 3;8 until the end of the study at age 4;0, there was a dramatic decline in the numbers of constructions leading to potential sites for both bilabial and velar assimilation, as well as a reduction in the proportions of assimilations produced. This reduction was most substantial for bilabial assimilation, partly because figures were so high at age 3;7. Numbers of potential bilabial assimilation sites fell from 59 sites at age 3;7 to eleven sites at age 3;8 and were further reduced to single figures for the remainder of the study. Similarly, numbers of potential velar assimilation sites fell from 22 sites at age 3;7 to only three sites at age 3;8. Numbers of potential velar assimilation sites remained at single figures until the end of the study.

Proportions of bilabial assimilations fell from 91.5% at age 3;7 to levels between 63.6% and 66.7% from age 3;8 to 3;10. Proportions then fell further to 25% and 20% respectively for the

final two months of the study. Proportions of velar assimilations declined from 61.9% at age 3;7 to 33.3% at age 3;8 and 16.7% at age 3;9. However, a further rise in proportions of assimilations was then evident for the final three months of the study, reaching proportions of 60%, 75% and 66.7% respectively. These proportions suggest a trend towards predominance of velar assimilation during this period. However, the numbers of potential assimilation sites sampled from age 3;8 to 4;0 are too small for these percentages to be statistically meaningful. Conclusions regarding trends of assimilation occurrence during this period must therefore remain tentative. However, it is noteworthy that throughout the three distinctive periods of assimilation emergence, establishment and reduction, close juncture realisations clearly predominated both at potential bilabial and velar assimilation sites throughout the study.

There are several possible explanations for these reductions both in constructions containing potential assimilation sites and proportions of assimilations. It was noted that there were technical issues with four recordings made during this period; two were unavailable, one was only partially transcribed and one was extremely short. However, this does not wholly explain the reductions observed. Six complete recordings simply yielded no constructions containing *can* or *can't* at potential assimilation sites and numbers of potential assimilation sites in the available recordings were greatly reduced.

There was no overall quantitative change in the usage of *can* and *can't* at this stage, as shown by consistently high frequencies from age 3;2 until the end of the study. However, qualitative analysis of Thomas' syntactic advances during this period showed a great decline in the emergence of novel constructions containing potential assimilation sites with *can* and *can't*. There was also evidence of a reduction in usage of constructions which had previously produced potential assimilation sites in Thomas' speech, notably *can be*. Thomas' usage of *can* and *can't* appeared to shift away from high proportions of declaratives, towards higher proportions of interrogative constructions, which do not give rise to potential assimilation sites. The syntax of declarative constructions also appeared to change, resulting in constructions which were not the focus of the current study. Such constructions would involve

those in which *can* and *can't* were replaced by other verbs, and those in which *can* and *can't* occurred with main verbs and adverbials which do not have an initial bilabial or velar plosive. An example of the former was evident from age 3;7, when Thomas acquired the conditional form *could* and began to use it alongside *can* from 3;7 onwards. In fact, weak forms of *can* and *could* became phonetically and contextually difficult to distinguish at this stage. The available phonetic cues, such as vowel nasalisation, were used to ensure as far as possible that only forms of *can* were included in the analysis. Two recordings were eliminated from the analysis, because all utterances glossed as containing *can* or *could* were indistinguishable.

The analysis of assimilation in M's speech shed further light on Thomas' reduction period. It was evident that she produced potential assimilation sites in 15 constructions which Thomas had not yet acquired. Seven of these occurred for the first time at T3, corresponding with the period of assimilation reduction in Thomas' speech. It therefore appears that the apparent decline in Thomas' assimilation was not in fact a sign of regression, but instead indicates a transition in the usage of focal constructions. The constructions which had previously produced potential assimilation sites were now occurring less frequently, as Thomas' usage of non-focal constructions increased. In addition, he was not yet using other constructions which produce potential assimilation sites in adult speech. It is predicted that if the data set for Thomas aged 4;1 to 4;11 were to be analysed, there would emerge a further peak in assimilation development involving a novel set of verbs. M's patterns of assimilation usage are discussed in more detail below.

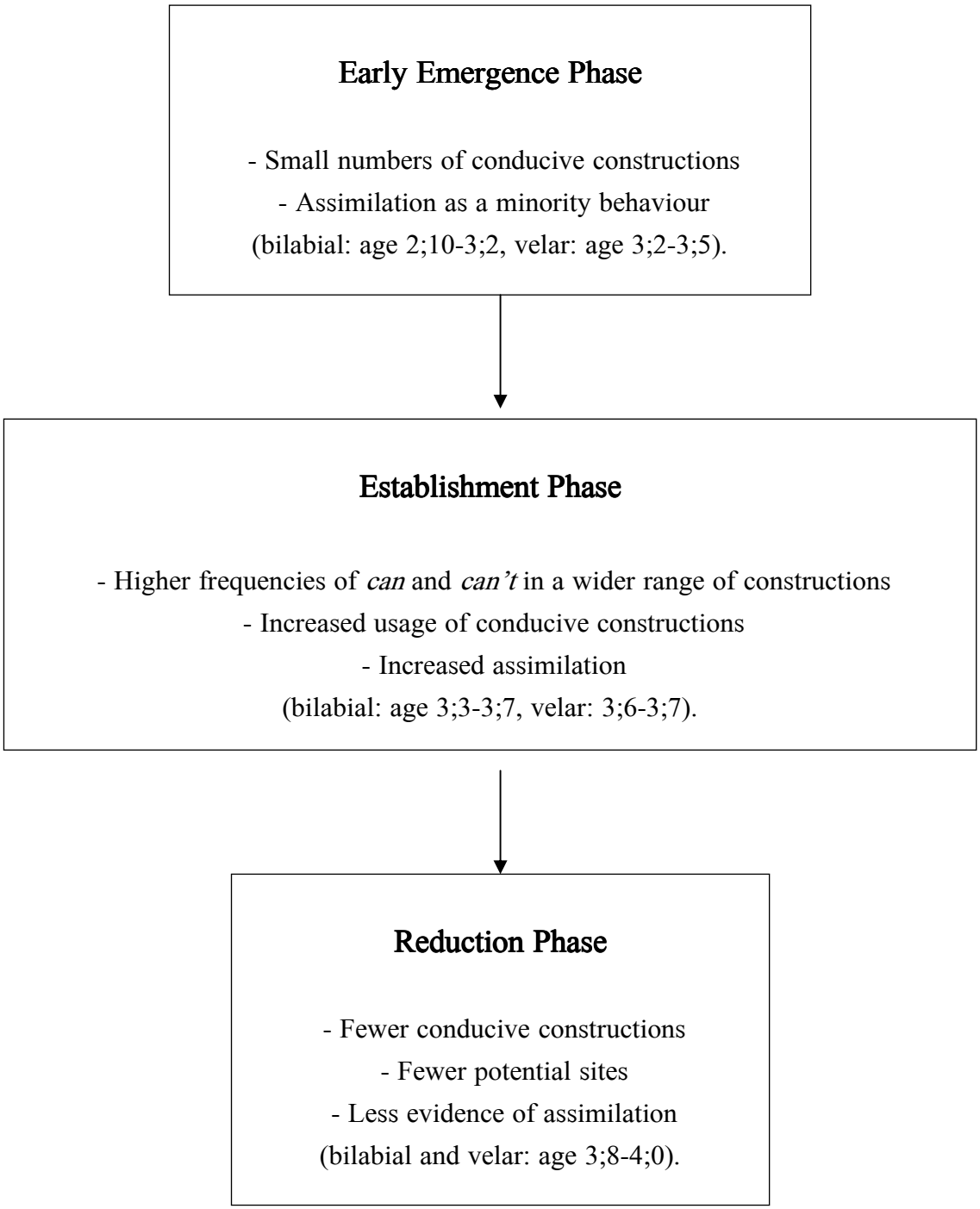
Taken together, these findings indicate that at least for this one child, there existed a critical period for the establishment of assimilation in constructions containing *can* and *can't* which was dependent on the acquisition of syntax. It was only when Thomas was producing constructions conducive to potential assimilation sites that he had sufficient opportunities to acquire assimilation in *can* and *can't*. Following the establishment of assimilation, it then occurred optionally alongside other possible phenomena at potential sites, including open juncture, non-assimilation close junctures, and a minority of other phenomena including

partial/gradient assimilation. When Thomas' usage of *can* and *can't* developed further, to include a different set of declarative and interrogative constructions, then utterances which had previously given rise to potential assimilation sites with *can* and *can't* declined, reducing the opportunities for further assimilation development. This developmental trajectory therefore provides evidence that interactions existed between different linguistic levels in Thomas' speech and language development. Specifically, it appears that the emergence and establishment of between-word assimilation in constructions containing *can* and *can't* was dependent on the syntactic development of these constructions.

9.3. A Phase Model of Thomas' Assimilation Development

Despite the differences in age of acquisition and possible differences in the underlying developmental processes governing bilabial and velar assimilation, the syntax-driven developmental trajectory described above is common to both. This trajectory of assimilation development can be described in terms of three developmental phases, which can be represented in the form of a simple box-and-arrow model (see Figure 9.1). The period during which assimilation initially emerged and occurred at a minority of potential sites can be viewed as the early emergence phase. This occurred from age 2;10 to 3;2 for bilabial assimilation and from 3;2 to 3;5 for velar assimilation. This early emergence phase was followed by an establishment phase, which was driven by advances in syntactic development and increasing numbers of potential assimilation sites occurring in a wide range of constructions. In the case of bilabial assimilation, establishment was evident from age 3;3 to 3;7 as clear predominance of assimilation over open junctures and other non-assimilation phenomena. For velar assimilation, establishment was evident during only a short period from age 3;6 to 3;7, with assimilation showing a trend towards predominance only at 3;7. The period from age 3;8 to 4;0 can be viewed as the reduction phase for both bilabial and velar assimilation, which was again driven by changes in syntactic development and usage.

Figure 9.1. A Phase Model of Thomas' Assimilation Development



9.4. A Comparison of the Current Study with Previous Research on Assimilation Development

Thomas' developmental trajectory for the emergence and establishment of assimilation compares interestingly with the findings of previous studies on assimilation development. The earlier emergence and establishment of bilabial than velar assimilation observed in Thomas' speech has consistently been reported in previous research (Bryan et al., 2010A; Newton and Wells, 2002; Thompson and Howard, 2007). It is striking that this is the opposite usage pattern from that reported in adult speech; regressive velar assimilation is reported to occur more frequently in English than regressive bilabial assimilation (Brown, 1990; Cruttenden, 2001). The earlier acquisition of bilabial than velar assimilation therefore appears to be at least partly explained by children's overall earlier acquisition of bilabial than velar nasals (for example, Dodd et al., 2003). It is possible that velar fronting in Thomas's speech contributed to his relatively late development of velar assimilation, a factor which may also be relevant for the children in previous research.

Most studies of assimilation development have reported parallel increases in numbers of potential assimilation sites, occurrence of assimilation and syntactic growth, as well as links between assimilation and the emergence of specific lexical items and constructions (Howard, Methley et al., 2008; Newton and Wells, 2002; Thompson and Howard, 2007). Such relationships were also identified in the current study and have been explored in detail, in order to further investigate the relationship between assimilation development and syntax.

Despite these similarities between the findings of the current study and those of previous studies of assimilation development, there also exist striking differences. Newton and Wells (1999, 2002) found that mature assimilation consistently predominated over open junctures, partial assimilations and idiosyncratic forms, both in a cross-sectional study of children aged three to seven years and in a longitudinal case study of a child aged from 2;4 to 3;4. These

researchers therefore conclude that assimilation (along with other CSPs) is automatic and coarticulatory in origin. They therefore suggest that CSPs emerge in early multi-word utterances, similarly to other immature forms, being simpler for the child to produce than open juncture forms (Newton and Wells, 1999, 2002).

The findings of the current study contrast sharply with those of Newton and Wells and therefore do not support this conclusion. Thomas showed clear progression from no occurrence of assimilation in constructions with *can* and *can't*, to assimilation at a minority of potential sites, to assimilation establishment, followed by reduction in both potential assimilation sites and assimilations. Predominance was only found for bilabial assimilation during its period of establishment, from age 3;3 to 3;7. This pattern suggests that phonological learning plays a greater role in assimilation development than Newton and Wells suggest. The earlier establishment of bilabial assimilation indicates either that it may be more coarticulatory in origin than velar assimilation, or that it simply reflects the fact that bilabial nasals are acquired earlier than velar nasals (Dodd et al., 2003). However, it is clear from these findings that both bilabial and velar assimilation of *can* and *can't* were acquired in the context of relevant constructions, a process of phonological acquisition which was dependent on syntactic development. Velar assimilation appears to have been especially dependent on Thomas' prior abstraction of schemas for *can* and *can't*. It emerged at age 3;2, at a time when it was evident that Thomas had recently abstracted the schema for *can* and was beginning to use both *can* and *can't* in a wider range of constructions. It became established only when advances in Thomas' syntactic development of relevant constructions peaked at age 3;6 and 3;7.

This conclusion corresponds more closely with the findings of Thompson and Howard (2007), who reported higher occurrence of open juncture in two-year-olds and higher occurrence of assimilation (and other CSPs) in three-year-olds. They similarly conclude that CSPs are learned over time as a result of gradual phonological refinement. The current study adds to

this proposal by placing greater emphasis on the interactions between syntactic and phonological acquisition in the development of assimilation.

The exact nature of the interaction between CSP development and syntax appears to vary across different children and at different developmental stages. Howard et al. (2008) report that increased usage of CSPs coincided with a reduction in MLU growth in their participant. They conclude that this resulted from competition between different linguistic processing demands. In contrast, Thomas' period of assimilation establishment appears to have been directly driven by advances in his syntactic development. These differences in the nature of interactions between assimilation and syntax may be a function of differential influences between linguistic levels at different developmental stages, and/or individual variability across children. It is also noteworthy that the study of Howard et al. focused on a wider range of CSPs and construction types than the current study.

It appears that the differences between the current findings and those of previous research may partly result from methodological differences between the two studies. Because this study focused only on *can* and *can't*, it was possible to track the trajectory of assimilation development in relation to specific constructions containing these auxiliaries. In contrast, Newton and Wells (2002) and Howard Methley and Perkins (2008) focused more globally on the phenomena occurring at all potential CSP sites in all relevant constructions. Their method permitted a more global investigation of assimilation development, but would not have revealed specific patterns of assimilation emergence within individual constructions over time. A further methodological difference which may explain the variable findings is that the current study spanned a longer period of two years, from age 2;0 to 4;0. In contrast, the studies of Newton and Wells (2002) and Howard et al. (2008) began later and finished earlier, spanning the periods from age 2;4 to 3;4 and 2;3 to 2;10 respectively.

In order to further explore the possibility that a global study of all CSP sites may obscure the specific patterns of assimilation emergence in individual constructions, the findings of the

current study are compared with the previous study of Thomas' assimilation development (Bryan et al., 2010A; Bryan et al., 2010B). Bryan et al. investigated the phonetic phenomena occurring at ten potential assimilation sites each month from the age of 2;0 to 4;0 (subject to availability of data). Assimilation was found to occur in Thomas' speech at a range of different potential sites from the age of 2;0, prior to the emergence of any constructions with *can* or *can't*. Variable trends towards predominance and non-predominance of assimilation over open junctures and other phenomena were observed throughout the study. Periods showing trends towards predominance increased in length over time relative to periods showing trends towards non-predominance.

The authors concluded from this pattern that Thomas increasingly produced assimilation over time, although no clear pattern of consistent predominance was evident by the end of the study. The evidence for increasing assimilation over time corresponds well with the findings of assimilation establishment in the current study. However, the overall patterns of emergence noted in this study and that of Bryan et al. (2010A) are quite different. It is suggested here that the variable trends towards assimilation predominance and non-predominance observed by Bryan et al. (2010) reflect the fact that assimilation was developing at different points in time in different constructions. Thus, the trends towards predominance may have corresponded with the emergence and establishment of assimilation in specific constructions, while the trends towards non-predominance may have indicated reduction in one construction, prior to emergence in another. The fact that assimilation was always evident in Thomas' speech to some extent indicates some overlap between emergence in some constructions and reduction in others. In order to investigate this more fully, it would be necessary to carry out further analyses of assimilation development in a selection of constructions, such as those containing *in*, *on* and *one* at potential assimilation sites.

Previous research has shown that constructions containing *can* and *can't* give rise to potential assimilation sites in the speech of two-year-olds (Bryan et al., 2010A; Newton and Wells, 2002). It is predicted that similarly to the findings observed for *can* and *can't*, other

constructions would show similar patterns of emergence and establishment. Reduction may also be evident, if the child's usage moved away from these constructions towards constructions which do not give rise to potential assimilation sites.

It is also interesting to compare the occurrence of open juncture and non-assimilation phenomena observed in the current study with those reported in previous research. Newton and Wells (2002) reported the sudden emergence of open juncture at potential assimilation sites in their participant at age 2;9, following previous predominance of mature assimilation at potential sites. In contrast, Thomas and the participants of Thompson and Howard (2007) produced open juncture forms from the onset of potential assimilation sites. In the case of potential velar assimilation sites in Thomas' speech, open junctures and non-assimilation close junctures persisted for six months prior to the emergence of assimilation in forms of *can* and *can't*. Overall, open juncture was extremely rare in Thomas' speech, as also reported by Howard et al. (2008) in relation to their participant. Instead, Thomas produced many non-assimilation close junctures. These appear to have been important in the current study because of the specific constructions under investigation. Firstly, Thomas produced many tokens of *can't* at age two containing [nʔ] clusters. Secondly, Thomas often omitted the final nasal in *can* leading to non-assimilation close junctures. Such forms would have been classified as idiosyncratic by Newton and Wells (2002).

9.5. The Relationship between Assimilation in Thomas' Speech and Assimilation in the Maternal Input

Assimilation in M's speech was studied at three points in time; T1 corresponded with Thomas aged 2;6, prior to the production of any potential assimilation sites in his speech; T2 corresponded with Thomas aged 3;3, at the beginning of his establishment phase for bilabial assimilation; T3 corresponded with Thomas aged 4;0, during the reduction phase. This analysis revealed that bilabial assimilation predominated over other phenomena at potential

sites in M's speech at all three points in time. Proportions of assimilations were extremely similar at T1 and T2, reaching levels of 78.9% and 78.3% respectively. There was a slight reduction in predominance at T3 to 72.7%. This reduction may have occurred because only four audio recordings were available for this month, one of which was cut short. In contrast, M produced only a minority of velar assimilations at T1 and produced assimilations at approximately half of potential sites at T2 and T3.

Following the findings of a large-scale study of 39 mothers' child-directed speech (CDS), (Foulkes et al., 2005), it was predicted in the current study that M would produce more non-assimilated forms during Thomas' third year, in order to maximise his exposure to the canonical alveolar coda consonants in *can* and *can't*. It was also predicted that her proportions of assimilations would increase over time, as Thomas' linguistic abilities increased and he no longer required this emphasis on canonical forms from the input. This prediction was not borne out for bilabial assimilation, which consistently predominated in M's speech at all three points in time. However, it was confirmed for velar assimilation, which increased in M's speech from only a minority of occurrences at T1 to occurrence at approximately half of potential sites at T2 and T3.

M's patterns of bilabial assimilation usage correspond more closely with the phonological patterns observed in CDS by Shockey and Bond (1980). These authors reported high levels of within-word phonological reduction in CDS and concluded from these findings that children are exposed to multiple exemplars of canonical and reduced forms in the input (Shockey and Bond, 1980). However, Shockey and Bond also draw a conclusion which is not supported by the current findings. They conclude that children learn the canonical phonological forms of words through their prior knowledge of the distribution of contrastive phones within the language which they are learning. The example which they give is that children learn that [t] in word coda position is "correct", because it occurs more frequently than the vernacular glottal stop. This conclusion has already been questioned (see chapter four). The current findings further refute this conclusion, because the nasals [m], [n] and [ŋ] are all highly

frequent in word coda position in English. Furthermore, both the alveolar nasal and the assimilated form (either bilabial or velar) produce acceptable forms of *can* and *can't* at potential assimilation sites. In this case, it was therefore impossible that Thomas acquired the canonical citation forms of *can* and *can't* purely from the relative distributions of the three nasals in English.

In a comparison of Thomas' and M's assimilation usage, it was noted that proportions of assimilations in M's speech at most points in time were similar to those observed for Thomas during his establishment phase. Thus, it is evident that Thomas' and M's proportions of bilabial assimilation were divergent at T1, converged at T2, before diverging again at T3, during Thomas' reduction phase. The pattern for velar assimilation was somewhat different; assimilation patterns for Thomas and M were convergent at both T1 and T2, because M produced only a minority of assimilations at T1. Thomas' and M's patterns diverged during Thomas' reduction phase at T3, as observed for bilabial assimilation. These patterns of convergence during Thomas' establishment phase indicate that he learned that both assimilated and non-assimilated forms are acceptable at potential assimilation sites and that his usage reflected the variety which he received from M's input. Again, this supports the conclusion of Shockey and Bond (1980).

These findings indicate that M may unconsciously have adapted her usage of velar assimilation according to Thomas' phonetic, phonological and syntactic abilities. Thus, she produced only a minority of velar assimilations at a time when Thomas had not yet acquired the appropriate constructions for velar assimilation. An increase in her velar assimilation was then evident at T2; this was a point at which velar assimilation had emerged in Thomas' speech, but had not yet become established. It is possible that this increase in velar assimilation in M's speech may have facilitated Thomas' establishment phase at age 3;6. This change in the phonetic and phonological characteristics in M's usage of velar assimilation over time is a novel finding, which could not have been detected in previous studies of CDS, owing

to their purely cross-sectional methodologies (for example, Shockey and Bond, 1980; Vihman et al., 1994; Foulkes et al., 2005).

Adaptations to Thomas' emerging phonological and syntactic abilities were not evident in M's speech for bilabial assimilation, which was produced at the majority of potential sites at T1, even when potential sites had not yet emerged in Thomas' speech. M's highly consistent production of assimilation at this early stage may partly explain why Thomas acquired bilabial assimilation four months earlier than velar assimilation. In contrast, the relatively low proportions of velar assimilations in M's speech at T1 may have been an adaptation to Thomas' linguistic level, but may also partly explain why Thomas acquired velar assimilation relatively late. As noted earlier, velar fronting of alveolar consonants during Thomas's third year may also have contributed to the lack of velar assimilation during this period. This is especially evident in instances when progressive dental/alveolar assimilation appeared to occur. In sum, it is suggested that Thomas' acquisition of assimilation was at least partly driven by exemplars provided in the input. It appears that only a minority of realisations are necessary for acquisition to occur, but that a higher number of occurrences may contribute to earlier emergence.

It appears that Thomas' acquisition of assimilation in constructions with *can* and *can't* resulted from a two-way interactional process between Thomas' increasing cognitive and linguistic abilities and M unconsciously adapting her input accordingly. This conclusion is compatible with the suggestion of Vihman et al. (1994) that children filter in those aspects of phonology from the input which match their motor schemes and filter out those which are not currently within their capabilities. If velar assimilation develops more slowly than bilabial assimilation, then Thomas may have filtered out exemplars of velar assimilation from the input, because he had not yet acquired the appropriate motor schemes. As Thomas gradually acquired the motor schemes for velar assimilation, M then responded to the changes in Thomas' speech by producing higher proportions of velar assimilations. The fact that Thomas then eventually mirrored the proportions of assimilations which he received in the input best

supports a usage-based approach to language acquisition. Such approaches emphasise the role of input frequency, the child's cognition and the child's usage in the acquisition of specific syntactic constructions and phonological phenomena (Lieven and Tomasello, 2008; Sosa and Bybee, 2008).

Further insight into the relationship between Thomas' learning and M's input was gained from an analysis of the relationship between Thomas' and M's realisations of potential assimilation sites, in the context of interaction. Portions of conversation were identified in which both Thomas and M repeated the same two words leading to a potential assimilation site with either *can* or *can't*, occurring either in identical or similar utterances.

There were only three instances in which Thomas immediately repeated a potential assimilation site produced by M. Two of these were exact repetitions of M's utterances, both of which matched M's preceding utterances in terms of the three prosodic variables tested: stress pattern, locus of tonic syllable and nuclear tone. In contrast, the single utterance which was not an exact imitation of M did not match M's preceding utterance on any prosodic variables. Only one of the two exact imitations matched M's utterance with regard to segmental realisation of the potential assimilation site, which notably occurred in a formula and was realised with assimilation.

This preliminary evidence therefore indicates that Thomas' exact imitations of M were more likely to prosodically match her previous utterances than his similar repetitions, possibly because such utterances have been learned holistically from the input and stored as formulae. However, there was no evidence that Thomas' segmental realisations of potential assimilation sites would necessarily mirror those occurring in M's previous utterance. This observation contrasts sharply with the findings of a case study of an autistic boy, who displayed immediate echolalia (Local and Wooton, 1995). This boy, Kevin, was found to produce exact articulatory matching of his mother's immediately preceding utterance, as well as prosodic matching of stress, duration and pitch contour. It would be necessary to conduct further studies of the

imitations and repetitions produced by typically developing children and children with autism, in order to investigate whether exact articulatory matching within an imitation is a specific feature of autism or results from individual differences across children.

The lack of consistent segmental matching in Thomas' imitative utterances demonstrates that although maternal input was instrumental in Thomas' acquisition of assimilation, he did not necessarily learn from the immediate model provided by M in the preceding utterance. It is therefore suggested that acquisition of assimilation is dependent on cumulative exposure in the input over time and does not rely on immediate modelling. This observation is in line with that of Ambridge et al. (2006B). These authors investigated the effect of prior training on children's acquisition of complex constructions. They found that training distributed over several weeks was more effective than massed exposure in enabling children to learn complex constructions in an experiment. They concluded that temporally distributed training optimises language acquisition and more closely resembles the input which the child naturally receives from caregivers (Ambridge et al., 2006B). It should be noted however, that the sample of Thomas' imitations and repetitions in the current study was extremely small. In order to confidently substantiate this claim, it would therefore be necessary to identify the same patterns occurring in a larger sample of Thomas' imitations and repetitions of M, in a wider range of construction types. A cross-sectional investigation of these phenomena across several children would also be valuable in order to determine whether individual differences exist. It may be that children who learn more slowly or who have learning difficulties (such as autism) may be more dependent on immediate modelling and massed exposure and less able to learn from temporally distributed exposure.

M's repetitions of Thomas provide clearer evidence that Thomas' assimilation development was facilitated by the input which he received. M's realisations of potential assimilation sites matched those of Thomas in 15 out of 22 utterance pairs (68.2%). This is a further, clearer piece of evidence showing that M adapted her realisations of potential assimilation sites according to Thomas' behaviours and therefore his underlying linguistic and speech processing

abilities. Most of the matched pairs occurring at potential bilabial assimilation sites were realised with assimilation, while most of the pairs occurring at potential velar assimilation sites were realised with non-assimilation close juncture forms. This finding further reflects the lower occurrence of velar than bilabial assimilation in both Thomas' and M's speech, apparently because of slower progress on Thomas' part and adaptation on M's part. Most instances of non-matching between Thomas' and M's potential assimilation site realisations occurred when Thomas produced a non-assimilated form and M produced a repetition with bilabial assimilation. This is further evidence for M's consistent predominance of bilabial assimilation throughout the study. It is also noteworthy that no non-matching of potential assimilation site realisations occurred beyond age 3;7;1, reflecting Thomas' more consistent production of assimilation during the establishment phase. In conclusion, the findings of this interactional analysis further support the suggestion made above, that assimilation developed in Thomas' speech as the result of a two-way interaction between Thomas' linguistic abilities and adaptation of M's speech.

There is clear evidence that M's realisations of potential assimilation sites matched those produced by Thomas in the preceding utterance. However, there is no evidence that she produced prosodic matching of Thomas' utterances to the same extent. When the co-occurrence of segmental and prosodic matching was investigated, segmental matching was found to co-occur with matching of stress pattern and locus of tonic syllable only in approximately 50% of instances. Co-occurrence of segmental matching and matching of nuclear tone occurred in only a third of instances. It was therefore concluded that matching of potential assimilation sites in the context of interaction is an independent phenomenon, which does not interact with prosodic matching. When the co-occurrence of non-matching segmental and prosodic phenomena was investigated, the only substantial finding was that segmental non-matching co-occurred with non-matching nuclear tone in 85.9% of instances. It is likely that this pattern resulted from interactions between pragmatic and conversational factors, rather than from a negative pattern of interaction between segmental and prosodic phenomena. Further investigation of this pattern is therefore beyond the scope of the current study.

9.6. Pauses, False Starts and Revisions at Potential Assimilation Sites in Thomas' Speech

It has been noted that age 3;2 marked the onset of substantial syntactic advancements for Thomas. In addition, he began to produce false starts and revisions at this age, which interrupted the word junctures at potential assimilation sites. It is possible that these were a manifestation of increased challenges in linguistic processing, at a time of great change. Two types of false starts occurred; firstly, there were those in which the whole potential assimilation site was repeated, for example "I can be, you can be a bear"; secondly, there were those in which the utterance was interrupted prior to the production of the potential assimilation site, for example "you can, I can be". There was evidence of phonetic changes in the realisations of false starts and revisions throughout Thomas' fourth year. From age 3;2 to 3;6, false starts were realised with assimilation and revisions were realised with open juncture. The peak of Thomas' assimilation establishment phase at age 3;6 and 3;7 was characterised by more instances of assimilation in both false starts and revisions. This appears to have been a period of transition, which was followed by a further change at age 3;7 towards more open juncture false starts and assimilated revisions. This pattern continued until the end of the study at age 4;0.

In instances when the false start involved interruption of the word juncture at the potential assimilation site, it was interesting to note that regressive place assimilation was sometimes evident in *can*, despite the pause which then followed prior to revision of the utterance. In other words, assimilation occurred in the absence of a potential site, as there was no following word-initial consonant with which the final /n/ in *can* could assimilate. Reasons for the revision were sometimes evident, such as the change of a pronoun, for example "I can be, you can be". However, in other instances, the utterance was interrupted and revised without any syntactic changes. In these cases, the interruption may have occurred because Thomas was

planning a later part of the utterance or because of interactional factors, which are beyond the scope of the current study.

Assimilations were also observed in a minority of utterances which contained a pause at the potential assimilation site, but which did not involve revision of the whole utterance. The first instance of this was sampled at age 3;1, for example "I can (.) make some room now" [ˈaɪ kʰæ̟̃m (.) meɪ]. Although assimilation is traditionally viewed as a close juncture phenomenon along with other CSPs, (Wells, 1994), these findings clearly show that assimilation may occur at an open juncture with a pause and in utterances spoken with low syntagmatic fluency. Again, these pauses are explained in terms of utterance formulation and interactional factors.

These findings contrast with those of Local and Kelly (1986), who found that regressive bilabial assimilation did not occur at potential sites at which there was a pause. However, these researchers were investigating the speech of an adult female speaker from East London (Local and Kelly, 1986). It is therefore possible that Thomas' production of assimilation in false starts and at word junctures with pauses was related specifically to developmental speech patterns, which do not occur in adults. There may also be effects of age, sex and regional accent upon these phenomena. It would be necessary to investigate these phenomena in larger samples of both adults and children, in order to explore these possibilities.

The introduction of pauses, false starts and revisions in Thomas' speech at a time of increasing syntactic complexity corresponds more closely with the developmental findings of Howard et al. (2008). These researchers noted that their participant produced familiar utterances with higher syntagmatic fluency and close juncture, while less familiar utterances were produced with lower syntagmatic fluency and more open junctures. They concluded that these less familiar utterances relied more upon segmental phonological processing, requiring analytic learning. It therefore appears that the introduction of false starts, revisions and pauses in Thomas's constructions with can at age three marked the onset of analytic learning of these constructions.

9.7. Phonetic Forms of Can in Thomas' Speech

Another developmental pattern highlighted in this study involved changes in Thomas' realisations of *can,* which shed further light on the processes by which *can* and *can't* were acquired. When *can* emerged in utterance-final position from 2;3 to 2;6, notably in the *Bob the Builder* formula "yes we can" or the modification "we can", it was realised as a strong form with variable initial alveolar, velar, or occasionally palatal or palatalised velar plosives. There was no clear trend at this stage towards predominance of either alveolar or velar initial plosives in *can*. When *can* emerged alongside main verbs from age 2;7 to 2;11, it was noted that the minority of stressed auxiliary forms occurring were usually realised with an initial /t/, while the more frequently occurring weak forms were realised with a variety of velar consonants, including voiceless, voiced, aspirated and unaspirated plosives, plosives with affricated release, palatalised plosives and fricatives. These findings are strikingly similar to those reported both in typical development (Inkelas and Rose, 2008) and in a five-year-old boy with an immature phonological system (Chiat, 1983). Inkelas and Rose (2008) report velar fronting of word-initial consonants and those forming the onsets of stressed syllables. This pattern persisted from age 1;0;27 until 2;2;28. Chiat (1983) also reports the occurrence of these patterns in delayed speech, although the interactions between segmental phonology and prosody were more complex, involving factors relating to word boundary characteristics. From the age of 2;9 onwards, Thomas predominantly produced initial velar plosives in *can*, although initial alveolars continued to occur throughout the remaining months of his third year.

The relative proportions of initial alveolar and velar plosives in Thomas' productions of *can* at age two reflect his global realisation of target velar plosives at this age. A control analysis of target velars in a variety of words and word positions revealed that Thomas realised less than half of target velar plosives with velar articulation at age 2;3 and 2;6. The most common realisations were alveolar plosives, although there were also a minority of other places of articulation, including palatal, uvular, glottal and bilabial. However, as specifically reported

for productions of *can*, the global analysis of target velar plosives in a variety of contexts showed a change towards consistent predominance of velar realisations from age 2;9 onwards.

Alveolar realisations of target velar plosives were much rarer from the age of three onwards. All except three forms of *can* from age three to four were realised with initial velar consonants. This indicates diffusion of velar consonants from weak forms at age two, to a range of stressed and unstressed forms with full vowels at age three. Again, these findings reflect those of the global analysis of target velar plosives. This showed that whereas 20 target plosives sampled at age two were produced with initial alveolar consonants, only three instances were sampled at age three and only one at age four. Thomas' pattern of velar fronting is typical of that described in previous research on phonological development. Fronting is considered a typical phenomenon occurring in children's speech at the age of two and three years, which is usually eliminated before the age of four (Grunwell, 1987; Hewlett, 1988; Ingram, 1976). It is noteworthy that the single instance of fronting sampled for Thomas at age four occurred within a cluster in the multi-syllabic word *excavator*.

Other sources of phonetic variation were also observed in Thomas' weak forms of *can* at age two. As well as the variable realisations of the initial consonant, there also existed variable vowel realisations, presence or absence of final nasals and the occurrence of syllabic final nasals instead of nuclear vowels. From age three to four, there continued to exist some phonetic variation in the voicing and manner of articulation of initial velar consonants in *can* as observed at age two. Examples include voicing, palatalisation and fricative production. It was predicted that some of these velar variants may also occur in M's speech, as they have previously been reported in adult connected speech (Cruttenden, 2001; Shockey, 2003). An analysis of M's speech confirmed this prediction, by showing similar variability in the initial consonants of *can,* especially the occurrence of palatals and fricatives.

There are several possible explanations which may account for the fact that Thomas produced weak forms of *can* with initial velar consonants at age two, at a time when alveolar

realisations of target velar plosives were frequent. It is possible that the pattern of initial alveolar consonants in strong forms and velars in weak forms may have resulted from coarticulatory influences of the nuclear vowels on the initial consonants. Thus, the front vowel [æ] in strong forms of *can* may have been more conducive to a preceding alveolar consonant, whereas a more central schwa or syllabic consonant may have been more conducive to a velar consonant. Similar consonant-vowel interactions have previously been reported in both typically developing (TD) children and children with speech impairments (Scobbie, 2002).

A further explanation for the presence of initial alveolars in strong forms and initial velars in weak forms is that Thomas initially learned weak forms formulaically as part of whole utterances, which he stored and retrieved as single units, as suggested by Wray and Perkins (2001). An analysis of M's speech revealed that the vast majority of her productions of *can* were weak forms. It is therefore likely that Thomas' high exposure to these forms in various constructions in the input enabled him to store the most frequent of these individually. It is suggested that he had not analysed the word *can* in these contexts. Future research could further investigate the point at which forms such as these cease to be filler syllables and develop a lexical representation linked with that of the strong form.

In contrast, Thomas first used strong forms of can in the formulae "yes we can" and "can we fix it?", which occurred frequently in the input. These formulae were spoken rhythmically as part of a song, with relatively equal stress placed on each word. A small minority of strong forms also occurred in other constructions in M's spontaneous speech, enabling Thomas to acquire these forms, although Thomas did not produce these frequently in non-formulaic contexts until age three. Similarly to the conclusions drawn for velar assimilation development, it is concluded that Thomas' acquisition of strong forms relied on only a minority of instances in the input, but that this reduced input may have contributed to later acquisition.

From age 3;1 until 3;5, all productions of *can* occurring at potential assimilation sites and all except two productions sampled in non-assimilation environments were realised with full vowels, regardless of whether or not they were stressed. In addition, these full forms occurred in a much wider variety of constructions than the limited range observed at age two. This provides evidence that Thomas analysed the auxiliary verb *can* at this point and began to use it more productively and less formulaically. Wray and Perkins (2000) define syntactic productivity as the ability to use the structure of language analytically in the segmentation of sentences into words and morphemes and the combinatorial construction of sentences from words and morphemes. They argue that this analytic learning phase follows on from the child's initial holistic learning phase. Both segmentation and construction were evident in Thomas' language during this period.

Weak forms and partial weak forms then re-emerged alongside full forms from age 3;5 onwards. This pattern supports the proposal of Wray and Perkins (2000) and Perkins (1999) that once children have learned to analyse individual syntactic elements and to combine them productively, they then proceed to a stage of more adult-like formulaicity, whereby they synthesise and store frequent utterances and collocations for efficient retrieval and usage. Wray and Perkins (2000) argue that this second, more mature formulaic phase relies upon both analytic and creative proficiency. They argue that although adults mostly employ holistic, formulaic language processing, they are able to employ more analytic processing when necessary, such as in adverse communicative conditions. Thomas' productions of both full and weak forms of *can* from age 3;5 indicates that he was employing both analytic and mature holistic language processing styles in parallel at this stage. The gradual reduction in full forms and increase in weak forms indicates a shift from less analytic to more mature, holistic language processing over time. This observation is in line with the proposal of Wray and Perkins (2000), that the two language acquisition strategies should be viewed as continuous and complementary, rather than as discrete.

It is evident that Thomas' acquisition of full and weak forms of *can* is in line with the proposal that children progress from an initial phase of rudimentary holistic learning, to a phase of analytic language learning, to a more mature holistic phase, which is complemented by analytic learning when necessary. Thomas's pattern of progression from phonetically variable weak forms, through to full forms and back to weak forms corresponds closely with the findings of a prosodic study by Behrens and Gut (2005). These authors similarly found that German-speaking children's early noun phrases progressed from a mature stress pattern with segmental errors on weak syllables, through to a phase of equal stress placement on each word, back to a mature stress pattern. It therefore appears that once Thomas had analysed *can* as an independent grammatical element, he was then able to integrate it as a weak form into the prosodic patterns of whole utterances. Thomas's later acquisition of mature weak forms is also in line with prosodic research at the single word level, which shows that reducing the stress of target weak syllables is a major challenge in the acquisition of prosody (Allen and Hawkins, 1980).

Thomas' pattern of acquisition of full and weak forms is also compatible with the usage-based, constructivist approach to language acquisition. According to this approach, usage of a word is initially restricted to the limited range of lexically-based constructions of which the child has experience (Lieven and Tomasello, 2008). High frequency and distribution of exemplars in the input which the child receives have been found to be important factors in determining the child's pattern of acquisition (Theakston and Lieven, 2005). These authors found that constructions which occurred frequently in the input were acquired earlier, appeared to be stored as wholes (as shown by a lack of generalised usage of auxiliaries) and were used frequently. Thomas was mainly exposed to weak forms of *can* in the input available in a range of samples, which explains his early acquisition of weak forms in a limited range of constructions. Substantial changes were evident in Thomas' usage of *can* from age 3;1 to 3;5, when weak forms were almost eliminated and were replaced by full forms, which occurred in an increasing range of constructions. The usage-based approach would explain this change in terms of abstraction of a more general schema for the usage of *can*; this occurred only when

Thomas gradually learned the linguistic relationships between different constructions containing *can*. Evidence of abstraction has been reported specifically for *can*, as shown by increased correct usage in a wider range of constructions over time (Rowland and Theakston, 2009).

In contrast with the findings for Thomas' assimilation development, Thomas' production pattern of full and weak forms does not reflect the patterns found in M's speech. Whereas definite periods of transition are evident for Thomas, from mainly weak forms, to full forms, to the re-emergence of weak forms, all of M's unstressed productions of *can* in spontaneous speech were weak forms. M even showed a tendency to produce weak forms in reading and singing contexts. Reading contexts are not spontaneous and singing contexts are more rhythmic than natural speech; such conditions might therefore be predicted to produce more open juncture phenomena, including strong forms.

A further analysis investigated whether there appeared to be any link between Thomas' full and weak forms of *can* and the presence or absence of assimilation at potential sites with *can*. Weak forms began to be realised with assimilation as soon as they re-emerged at age 3;5, during Thomas' bilabial assimilation establishment phase. Weak forms and assimilation also appeared to be CSPs which were independent of each other in Thomas' speech. Evidence for this comes from the study of false starts and revisions. As already noted, a range of full and weak forms of *can* were found at the sites of false starts and revisions. There was evidence of assimilation and non-assimilation behaviours occurring in both full and weak forms, with no evidence of any pattern of interaction between these phenomena. Both full and weak forms were also involved in instances of regressive bilabial assimilation, progressive nasal assimilation and resultant gemination. It therefore appears that Thomas's development of assimilation in these constructions was not affected by changes in his production of stress at potential assimilation sites from age 3;5.

Thomas produced many non-assimilation close junctures with *can*, when there was no evidence of a final nasal. The nuclear vowel was often nasalised in such instances, although this was not always the case. These were evident in weak forms at age two, full forms in early months of the fourth year and were especially evident in re-emerging weak forms from 3;5. Although non-assimilation close juncture forms of *can* occurred frequently in Thomas' speech, these were extremely rare in M's speech.

9.8. A Comparison of Thomas' Data with Phonological Research on Auxiliary Verbs

In an attempt to explain Thomas' pattern of auxiliary development in the light of linguistic and phonological theory, his data are compared with specific phonological and grammatical theories proposed in the literature on auxiliary verbs. Zwicky (1970) proposes that in adult speech, weak forms of auxiliaries are derived from strong forms via generative phonological rules. With regard to speech development, this theory would predict either that strong and weak forms would emerge in parallel, or at least, that weak forms would emerge very soon after the emergence of strong forms (Zwicky, 1970).

Thomas' pattern of emergence for *can* does not support this prediction. At the age of two, Thomas used strong and weak forms of *can* in entirely different contexts and may not have formed the lexical links between them. There is no evidence that Thomas' early weak forms of *can* had clear lexical or phonological representations. It is proposed above that the high phonetic variability observed in these early weak forms indicates that Thomas may have employed these as filler syllables, rather than as lexical items at this stage. There is more evidence that Thomas had lexical and phonological representations for strong forms of *can* at age two, as shown by the greater phonetic consistency across different tokens. The emergence of weak forms from Thomas' lexical representations of *can* was not evident in Thomas' speech until age 3;5, when weak forms re-emerged alongside stressed and unstressed full forms and occurred in the same constructions. This age therefore marked the point at which

Thomas formed the lexical links between full and weak forms and recognised both forms as permissible in unstressed contexts. Thomas's increasing pragmatic knowledge and linguistic processing capacity are probable contributing factors which enabled him to form these links.

Zwicky (1970) argues that the strong and weak forms of an auxiliary have the same lexical representation, but different phonological representations, the weak form being of secondary status. In contrast, Ogden (1999) argues that strong and weak forms should be viewed as phonologically contrastive forms of equal status, which are stored together in the lexicon. Thomas' pattern of acquisition of *can* and *can't* sheds no further light on whether weak forms of *can* have equal or secondary status compared with full forms. However, the data indicate two possibilities concerning his phonological representations of *can.*

Firstly, it is possible that Thomas had separate phonological representations for strong and weak forms, which were stored together for the same lexical item, as suggested by Zwicky (1970) and Ogden (1999). This seems unlikely, owing to the phonetic similarity between strong and weak forms of *can*, which suggests that they are closely phonologically related, rather than suppletive allomorphs, as suggested by Kaisse (1983). Furthermore, weak forms such as [k^hə̃n] and [k^hn̩] cannot occur in isolation and might not even be recognised as words if presented to a listener in isolation. In contrast, the full citation form /kæn/ would be recognised as a word in isolation. Therefore, an alternative explanation is that /kæn/ has a single phonological representation and that the weak forms are phonetically derived from this representation, rather than having separate representations. Thus, the single phonological representation of *can* is modified phonetically during speech production to produce either the strong or weak form, as appropriate for the context.

Further evidence for this claim is the fact that Thomas did not always produce a discrete phonetic distinction between unstressed full and weak forms, when weak forms first re-emerged. Instead, there appeared to be a continuum spanning from full forms at one end, weak forms at the other and gradient forms in-between, which showed partial reduction of the

vowel towards a weak form, such as [kʰæ̈n]. Numbers of gradient forms reduced over time and were eliminated by the end of the study. Numbers of fully reduced weak forms with either a nuclear schwa or syllabic nasal increased over time, with the reduction of gradient forms. Gradient or intermediate auxiliary forms have also been observed in adult speech (Mackenzie, 2012). However, this research involved different verbs and consequently, different types of phonetic variation. MacKenzie (2012) distinguishes between three distinct categories of auxiliary forms: full, intermediate and contracted. This categorical distinction between forms may be appropriate in the study of auxiliary forms in adult speech. However, the idea of phonetic continuity between forms is preferred here, in order to account for the wide range of subtly different forms observed in Thomas' speech. This view is in line with a recent phonological study of auxiliary development, in which continuity between forms was also identified (Dye, 2011). If this variation were to be driven by the selection of an appropriate phonological representation, then each auxiliary would require multiple phonological representations for full, weak and gradient forms. It is therefore concluded that this high phonetic variability reflects phonetic modification of a single phonological representation during connected speech production, which is driven by the speaker's phonological, grammatical and contextual knowledge. This knowledge is acquired over time, as shown by the variable patterns of full and weak forms in Thomas' speech and the gradual elimination of gradient forms.

Ogden (1999) argues that the occurrence of strong and weak forms is determined by grammatical factors and CSPs. There was no evidence that grammatical factors determined the occurrence of strong and weak forms in Thomas' speech. When weak forms of *can* re-emerged alongside stressed and unstressed full forms at age 3;5, they co-occurred in the same types of constructions and appeared to be interchangeable. This finding suggests that if there exist grammatical influences on the occurrence of strong and weak forms, these may rely on more advanced grammatical knowledge than Thomas had attained by the end of the study and may therefore be more evident in the speech of older children and adults.

With regard to Ogden's (1999) proposal that connected speech factors also determine the occurrence of strong and weak forms, Thomas' data show no evident effect of assimilation in determining whether full or weak forms occurred; that is, instances of assimilation, open juncture and other non-assimilation phenomena were found in both full and weak forms. However, as discussed above, it is evident that Thomas gradually acquired appropriate usage of full and weak forms over time. This process is likely to have resulted from the effect of increasing phonological and linguistic knowledge on connected speech production.

9.9. Phonetic Forms of Can't in Thomas' Speech

As observed for the auxiliary *can*, the auxiliary *can't* showed high phonetic variability at age two. There existed the same variable initial alveolar and velar realisations at age two as found for *can* and as observed globally in Thomas' speech at this age. However, in contrast with *can*, no initial alveolars were found in *can't* from age three onwards, whereas a minority were still observed in *can*. This may be the result of coarticulation, whereby the front nuclear vowel in *can* was more conducive to alveolar articulation, whereas the back vowel in *can't* was more conducive to velar articulation (see also Scobbie, 2002).

Whereas open juncture forms of *can* with a perceptible final alveolar nasal were frequently observed, open juncture forms with a final /nt/ cluster were extremely rare for *can't*. When *can't* first emerged in utterance-final position, it was frequently realised with a fully released /nt/ cluster. However, when it emerged as an auxiliary alongside a main verb, it was realised with a variety of non-assimilation close juncture forms. These included a final [nʔ] cluster, a final singleton [n] or the absence of a final consonant, with nasalisation of the nuclear vowel. The latter was not regularly observed beyond age 3;5.

9.10. Evaluation of the Current Study

Following the summary and discussion of all of the major findings, it remains to evaluate both the strengths and weaknesses of the methods employed in the current study. It appears that although some factors might be construed as obvious weaknesses which may have confounded the study, the same factors can also be viewed as strengths, which have enabled the data to be examined in a unique and novel way.

9.10.1. Strengths and Weaknesses of the Dense Database and the Longitudinal Case Study

The current study involved a longitudinal case study of one typically developing child, Thomas. The period of investigation spanned over two years from age two to age four. From age 2;0 to 2;11, approximately hour-length recordings of spontaneous mother-child interaction were sampled five days per week, every week throughout the year. Although this highly intensive sampling was continued until age 3;2 for the purposes of the Dense Database (Lieven et al., 2009), sampling in the current study was reduced from age 3;0 to include five hour-length recordings during one week of every month. This reduction was implemented in the current study three months earlier than it was employed in data collection by Lieven et al. (2009), in order to increase the time-efficiency of the study, at a time when numbers of potential assimilation sites with *can* and *can't* began to increase in Thomas' speech. This was especially evident from age 3;2.

The reduced sampling rate of five hours during one week of every month is a creative sampling technique, designed to enable researchers to reach a compromise between the realistic time limits involved in data analysis and the need to produce accurate estimates of developmental phenomena (Tomasello and Stahl, 2004). Thus, while sampling of Thomas' language remained much more intensive than that employed in most other studies of language

acquisition, the overall sample size was reduced to the equivalent of that collected in a smaller scale study, while individual sampling periods were sufficiently dense to capture the emergence of both high and low frequency phenomena (Tomasello and Stahl, 2004).

To the current author's knowledge, the Dense Database provides an unprecedentedly dense and rich sample of natural child speech and language. This has enabled the detailed investigation of many different linguistic phenomena as they emerge naturally, especially aspects of syntax (Dabrowska, Rowland and Theakston, 2009; Lieven et al., 2003; Lieven et al., 2009; Rowland and Theakston, 2009; Theakston and Rowland, 2009). However, the current study is the first which has employed this database in the investigation of phonetic and phonological phenomena in language acquisition.

In the current study, this dense longitudinal data sample has enabled a detailed investigation of the interactions between assimilation and syntax. The results have shown that the emergence and establishment of both bilabial and velar assimilation are dependent upon quantitative and qualitative advances in syntactic development, which give rise to higher numbers of potential assimilation sites within an increasing range of constructions containing *can* and *can't*. Equally, reductions in the usage of constructions leading to potential assimilation sites led to reductions both in numbers of potential assimilation sites and proportions of assimilations. Although previous longitudinal and cross-sectional studies have produced evidence of interactions between CSP development and syntax, this is the first study in which such dense data has been sampled over a sufficiently long period to provide detailed information regarding these interactions.

One limitation of longitudinal case studies such as this is that the findings reported and the interpretations made are based on observations of only one child. It is important to remember that individual differences exist in children's patterns of language acquisition (Peters, 1977), an observation which has been made specifically in relation to the development of auxiliaries (Lieven, 2008) and assimilation (Thompson and Howard, 2007). However, the advantage of

the single case study method employed in this study is that it has enabled in-depth phonetic and linguistic analyses, which have yielded uniquely detailed information on the interactions between assimilation development and syntax (see above). This level of analysis would not have been possible in a cross-sectional study, owing to constraints on research resources. It is also noteworthy that a number of seminal research findings which have contributed to linguistic understanding of phonological and syntactic development were obtained using in-depth case studies (for example, Bloom, 1970; Brown, 1973; Peters, 1977; Donahue, 1986; Stemberger, 1988; Matthei, 1989) This fact demonstrates the high value of the case study methodology. Nevertheless, it would be worthwhile to expand on the current findings by conducting similar longitudinal case studies on further children's data from the Dense Database, in order to investigate whether the patterns observed for Thomas occur in other children. It is predicted that similar patterns of assimilation emergence, establishment and possibly reduction would be observable in specific constructions in other children. However, there may exist individual differences in the age at which these phases are evident and the duration of each phase, as a function of the child's linguistic ability and the input and usage frequencies of the constructions under investigation.

The current study has focused only on constructions which contain the verbs *can* and *can't* at potential assimilation sites, whereas other studies have focused on CSP sites in a range of different constructions. On the one hand, this is an obvious limitation, because it narrows the focus of the study and potentially produces findings which are not generalisable to other verbs and constructions. On the other hand however, this study reveals a pattern of assimilation emergence, establishment and reduction within these specific constructions. Such a pattern has not been observed in previous longitudinal studies of assimilation development (Bryan et al., 2010A; Howard et al., 2008; Newton and Wells, 1999, 2002), even though the most recent of these studies was conducted on the same child's data as the current study. It is therefore suggested that the broader focus on potential assimilation sites in previous research may have obscured patterns of assimilation emergence in individual constructions, whereas the current study's specific focus on constructions with *can* and *can't* has highlighted this pattern and

proposed the first developmental model of assimilation development. It would be necessary to further test the current claims regarding the developmental trajectory of assimilation, by investigating its development in a wider range of constructions.

9.10.2. Strengths and Weaknesses of the Methods Employed in Data Analysis

Most of the data analyses conducted in this investigation involved impressionistic phonetic transcription. As discussed in Chapter Four, this methodology provides valuable insight into the production of speech, although there exist many challenges and sources of bias of which the researcher should be aware.

The data analysed in the current study consisted of digital audio computer files, which were derived from digital minidisk recordings. The microphone was stationed in the room with Thomas and his mother. Subject to good working order of the recording equipment and appropriate setting of the recording level, the resultant recordings were generally of the best available quality at the time of data collection. However, there were instances in which either Thomas wandered away from the microphone, the recording level was set too high or low, or the recording equipment was not in full working order. These factors may have adversely affected transcription accuracy in a minority of utterances which were rendered difficult to transcribe.

Although it is considered optimal that recordings for phonetic transcription be obtained in a quiet room (Ladefoged, 2003), this was not always possible for the recordings in the Dense Database, which were produced in the children's own homes. The advantage of this is that the speech and language data were entirely natural and not elicited. However, the disadvantage is that there was often extraneous noise in the recording environment from other people, the

sounds of household activities such as cooking, and particularly from Thomas playing with toys.

The transcriber attempted to minimise the potentially confounding effect of these factors on the quality of the transcriptions by transcribing only those utterances which were sufficiently clear for confident interpretation. Any utterances which were present in the recordings, but which the researcher could not confidently interpret were marked as untranscribable.

It is likely that phonetic expectation bias confounded the current study to some extent. The researcher worked with full orthographic transcriptions of Thomas' speech alongside the audio recordings. The usage of glosses such as these is considered disadvantageous in phonetic transcription, because it may interfere with the transcriber's ability to detect information in the signal through bottom-up perceptual processing (Heselwood and Howard, 2008). In the current study, it is possible that the presence of the glosses, which were the basis of both the phonetic analysis and the quantitative analyses using CLAN, led to misidentification of utterances. Thus, it is possible that a minority of utterances which were glossed as containing *can* and *can't* did not in fact contain either of these focal forms. This may account for the minority of unusual forms which occurred. For example, the utterance glossed as "look this, you can put in" and transcribed as [ˈkʰɛə̃ʔ pʊʔ] was interpreted by one of the current author's supervisors as "you get put in", although the former was considered by the current author to be grammatically more typical of Thomas' language.

The current transcriber attempted to avoid the negative effects of expectation bias as far as possible, by remaining aware of possible sources of bias and being open to unexpected phonetic phenomena in the data as advised in the phonetic transcription literature (Howard and Heselwood, 2002; Kent, 1996). Evidence that this has been achieved to an extent comes from instances in which the researcher interpreted utterances differently from the meanings given in the glosses (see notes on transcription discrepancies in Appendices Three and Four). These different interpretations arose either directly from perceptible phonetic aspects of the utterance

or from the interactional context surrounding the utterance. In cases of uncertainty, both the original and new interpretations were recorded as possible options in the transcription data. In instances of uncertainty as to whether or not the focal word was present, either the possible alternatives were shown in the transcript and/or notes, or in highly doubtful instances, the utterance was eliminated from the data set for this study. These instances in which the current transcribers' interpretations differed from those given in the gloss indicate that the transcription process was carried out as autonomously as possible, using the glosses only for guidance.

A further consideration is those factors which may have interfered with the transcriber's correct perception of assimilation versus non-assimilation. It may be that small deviations from the expected phonetic forms of nasals in *can* and *can't* at potential assimilation sites led the transcriber to misidentify their place of articulation and thus perceive them as different, contrasting segments. This phenomenon has been labelled phonemic false evaluation (Buckingham and Yule, 1987). A similar phenomenon is observer drift, which occurs when the transcriber's criteria for the identification of phenomena change over the course of the transcription process (Shriberg and Lof, 1991). The transcriber attempted to avoid these biases, by keeping an open mind as to the identity of all sounds occurring in Thomas' speech, being especially aware of the high phonetic variation which can occur in developmental speech. Evidence that this was successful to an extent comes from a minority of instances in which unexpected phenomena were perceived, which had features of both assimilation and non-assimilation and therefore were not clearly classifiable as one or the other. Examples include a palatal nasal in *can go* [kʰæ̃nʲ (..) ˈgəʊ] and a lengthened nasal with an audible transition from alveolar to velar articulation in *can get* [kʰɑ̃nŋ ˈgɛt̪ˢ]. Further evidence that the transcriber's perceptions were mostly accurate comes from the high level of agreement obtained between the current author and her supervisors during the consensus exercise.

It has been reported here that assimilation and non-assimilation phenomena were perceived as distinct categories, with little evidence of assimilation occurring as a partial or gradient

phenomenon. It is possible that these perceptions resulted from phonemic false evaluation. The consensus exercise showed that the second and third raters classified a minority of instances as gradient, which had been classified into one of the discrete categories by the current author. However, the conclusion that partial assimilation was rare in Thomas' speech is in line with an articulatory study of adult assimilation, which also found that instances of partial assimilation were rarer than clear instances either of assimilation or non-assimilation (L. Ellis and Hardcastle, 2002). Furthermore, not all studies of CSP development have included a category of partial assimilation in their analyses, suggesting that it did not emerge in the children's speech (Newton and Wells, 2002; Thompson and Howard, 2007). It is therefore concluded that although the current author may have classified a small minority of gradient forms too readily into one of the discrete categories, such forms are rare and therefore do not affect the overall validity of the study.

In spite of these potential pitfalls, the greatest strength of employing phonetic transcription in this study is that it is the only methodology which could have yielded the current findings. Instrumental methods, such as acoustic analysis, could not easily have been used in the current study owing to the naturalistic research conditions and the huge quantity of spontaneous speech data obtained. However, the investigation of spontaneous speech data in naturalistic conditions has yielded some of the richest data on the acquisition of phonology, including the acquisition of connected speech (Donahue, 1986; Stemberger, 1988; Matthei, 1989; Newton and Wells, 2002; Howard, Methley and Perkins, 2008). It is therefore concluded that phonetic transcription was the only realistic methodology which could be employed in an investigation of this kind. Furthermore, phonetic analysis through transcription has revealed the developmental trajectory of assimilation and its relation to syntactic acquisition, findings which may not have come to light in more controlled experimental conditions. However, it is possible that future instrumental analysis of high quality portions of the audio data could be used to further investigate phenomena highlighted in this study.

9.11. Conclusion

The current study has employed the dense naturalistic speech and language data obtained from one typically developing child, in order to explore in detail the interactions which exist between assimilation development and acquisition of syntax. The longitudinal nature of the study and the focus on specific syntactic constructions have revealed parallel developmental trajectories in the development of assimilation and syntax, which have not previously been detected. The findings shed light on two theoretical debates; the process by which assimilation emerges as a phonological phenomenon and the processes which underlie the acquisition of auxiliary syntax.

The findings have shown that there have existed evident parallels between assimilation and syntactic acquisition in Thomas' speech and language development, for constructions containing the auxiliaries *can* and *can't*. Assimilation appears to be a phonological phenomenon, rather than simply a phonetic behaviour resulting from coarticulation. It gradually emerges in individual constructions at different points in time. There exists evidence that it is dependent on and driven by syntactic advancements. Thus, high usage of a wide range of constructions which give rise to potential assimilation sites provides optimal opportunities for assimilation to become established. There appears to be a sensitive period in which this establishment takes place. Conversely, reduced usage of such constructions produces a decline in assimilation, at least for a short period, while the child's phonological system is still undergoing refinement. Assimilation is never obligatory and continues to occur optionally alongside open junctures and other non-assimilation phenomena, including non-assimilation close juncture forms. The developmental trajectory for assimilation can be expressed as a phase model showing an emergence phase, establishment phase and at least for some constructions, a reduction phase. Comparison of the current findings with previous assimilation research indicates that individual trajectories exist for different construction types. This means that a child may simultaneously be in different phases of assimilation development

for different constructions. It therefore appears that assimilation development is lexically specific. This conclusion is in line with usage-based approaches to language acquisition, especially those which emphasise the holistic storage of neuromotor routines and retrieval of articulatory gestalts for utterances in phonological development (Bybee, 2002, 2006).

Usage-based approaches greatly emphasise the importance of the input which the child receives in determining patterns of language acquisition. The current study has confirmed this at a phonological level for the emergence of assimilation. Comparison of Thomas' and M's realisations of potential assimilation sites has revealed that Thomas' proportions of assimilations relative to other phenomena were similar to those produced by M. This resulted in a pattern of convergence between Thomas and M during Thomas' establishment phase, in which bilabial assimilation predominated and velar assimilation was produced at approximately half of potential sites. Comparison of Thomas' and M's realisations of potential assimilation sites in the local context of interaction appears to show that distributed exposure in the input was more important than immediate modelling in Thomas' acquisition of assimilation. This has previously been shown in relation to the acquisition of word order, (Ambridge et al., 2006B).

Non-assimilatory forms also continued to be reinforced in Thomas' speech by the exemplars in M's input. Assimilation was therefore acquired as an optional phenomenon alongside other acceptable non-assimilated forms. It is concluded from Thomas' late acquisition of velar assimilation and M's low frequency of velar assimilation at T1, that acquisition requires only a minority of exemplars in the input. However, lower frequency of a phenomenon in the input may lead to later acquisition. This conclusion is in line with those drawn from previous usage-based research on the development of auxiliary syntax (Lieven, 2008; Rowland and Theakston, 2009; Theakston and Rowland, 2009). The relationship between Thomas's and M's realisations of potential assimilation sites was not straightforward. There exists evidence that pragmatic and interactional factors may also determine the realisation of potential assimilation sites. Future research into the exact nature of these factors would be valuable.

It is suggested that once assimilation has become established in a child's speech, the occurrence of assimilated versus non-assimilated forms depends on whether the child is processing the relevant construction in a more analytic or holistic way. Thus, holistic processing produces assimilation at the vast majority of potential sites, in order to maximise articulatory efficiency. However, analytic processing remains available for the production of non-assimilated forms, such as open juncture, in adverse communicative situations. This prediction is in line with the balanced and complementary nature of the analytic and holistic learning styles proposed by Wray and Perkins (2000). Further research should test this prediction, by tracking the occurrence of assimilation in later childhood and adolescence. The first step would be to continue to analyse Thomas' assimilation, using the remaining data available from age 4;1 to 4;11 (Lieven et al., 2009).

The patterns observed for the acquisition of the auxiliary verbs *can* and *can't* provide evidence for the usage-based constructivist approach to language acquisition. The relative formulaicity and restricted usage of *can* compared with the relative productivity and restricted usage of *can't* at age two support individual, lexically-specific learning of items which are closely related in the adult grammar. It therefore appears that grammatical relationships between similar forms and their relative grammatical complexity do not determine patterns of acquisition.

It should be acknowledged that other communicative factors besides relative analyticity and formulaicity may have contributed to the different developmental trajectories of c*an* and *can't*. For example, Thomas produced frequent tokens of *can't* in isolation during his third year, whereas isolated tokens of ca*n* were extremely rare. This demonstrates a communicative preference for the negative form of the auxiliary during this period. This preference provided Thomas with more opportunities to abstract a schema for the usage of can'*t* as a single lexical item and grammatical element.

Previous empirical research based on the usage-based approach has shown that the input which the child receives plays an important role in the acquisition of auxiliary syntax (Lieven, 2008; Rowland and Theakston, 2009; Theakston and Rowland, 2009). This has also been confirmed in the current study. Thomas' earlier acquisition of *can* than *can't* seems at least partly to be related to formulae occurring in a song to which he was frequently exposed. It is also predicted that *can* was more frequent overall than *can't* in the maternal input. Further research should test this prediction by calculating frequency counts for *can* and *can't* in M's speech. It is noted that the same constructions containing *can* and *can't* contributed to the highest number of potential assimilation sites in both Thomas' and M's speech; these were constructions in which the auxiliaries *can* and *can't* modified the main verbs *be*, *put*, *play*, *get* and *go*.

Thomas' differential emergence patterns for *can* and *can't* also show that analytic and holistic language learning may occur in the same child at different points in time and across different domains (Peters, 1977, 1995; Wray and Perkins, 2000). In fact, these two complementary learning strategies can explain the patterns of acquisition proposed in usage-based theories. The child's initial tendency towards a holistic strategy leads to the lexically-specific, formulaic acquisition of forms and constructions. A subsequent transition towards a more analytic strategy enables the child to recognise the relationships between similar forms occurring in different syntactic contexts and to abstract schemas for more generalised usage of individual forms.

Wray and Perkins (2000) propose that children progress from an initial, immature holistic language acquisition style, to an analytic style, to a more mature holistic style, which requires syntactic creativity. This pattern was evident at a phonetic level in Thomas' changing realisations of *can*. Most of his early unstressed forms of *can* were weak, reflecting the vast majority of weak forms which he received from the maternal input. The high phonetic variability of these forms and the limited range of constructions in which they occurred indicate that these constructions were learned as formulae and that Thomas had not yet linked

these forms lexically with the strong form *can.* The introduction of unstressed forms of *can* and the almost complete exclusion of weak forms from age 3;1 to 3;5 indicates that Thomas had made this lexical link, through analytic learning. Thomas' relatively late acquisition of unstressed full forms is explained by their relative infrequency in the maternal input. Further evidence for analytic learning comes from Thomas' usage of *can* in an increasing range of constructions and the onset of pauses, false starts and revisions. A gradual transition to more mature, holistic learning was evident as weak forms re-emerged and began to occur increasingly frequently from age 3;5.

When the interpretations of the phonetic, phonological and syntactic analyses in this study are brought together, the interdependence of these three domains becomes apparent. The main theme of this study has been the interaction between increasing syntactic creativity in a range of constructions (including complex sentences) and the increase in opportunities for Thomas to develop assimilation as part of his phonological system. Once Thomas had acquired assimilation as a permissible phonological form at potential sites, it then seemed to develop further at a phonetic level during its peak of establishment from age 3;6 to 3;7. The evidence for this is that as weak forms re-emerged through a return to more holistic learning, assimilation was immediately applied to these constructions, which were apparently produced holistically. This suggests that while assimilation and weak forms were independent CSPs at a phonological level, holistic language processing enabled further phonetic development of both CSPs. Thus, gestural overlap at potential assimilation sites increased, through the storage of neuromotor routines and the retrieval of articulatory gestalts for whole utterances (Bybee, 2002; 2006). This would explain the especially high incidence of assimilation in the highly frequent constructions *can be*, *can/can't get* and *can/can't go*. It appears that initial assimilation emergence was a gradual process and is therefore assumed to be phonological in origin, perhaps linked with early syntactic creativity and analytic learning. However, its establishment at the majority of potential sites then involved further phonetic, construction-specific refinement, driven by holistic storage and retrieval of constructions. Open juncture at potential assimilation sites always remained an option for Thomas. This confirms the proposal

of Wray and Perkins (2000) that analytic processing always remains an option, even in adulthood.

Thomas's tendencies towards more analytic or holistic learning at different points in time also appear to have interacted with his phonetic development at a prosodic level. Both early immature linguistic processing and later mature processing coincided with mature stress patterns, while the intervening period of schema abstraction through analytic learning led to deviations from mature stress patterns and the elimination of weak forms. These examples show how grammatical development and learning at the segmental phonetic level may drive the acquisition of mature prosody.

These connections between the domains of phonetics, phonology and syntax demonstrate that it is most valuable to investigate parallel developments in these areas, rather than exploring a single linguistic domain in isolation, in order to gain a more complete picture of speech and language development. This study has also provided preliminary evidence of interactions between these linguistic levels and factors relating to pragmatics, discourse and interaction. These factors are beyond the scope of the current study, but warrant further investigation in future research.

In conclusion, the current study is the first to have explored the interactions between the development of assimilation and syntax in detail. A phase model for the development of assimilation has been proposed. The major phonological and syntactic findings and their interactions support the predictions made by usage-based, constructivist approaches to language acquisition. In addition, the concepts of analytic and holistic learning have been incorporated into the usage-based approach, in order to explain the phonological and syntactic patterns observed. Further research should employ similar methods to investigate a wider range of CSPs and constructions in further children speaking a variety of languages. This will enable researchers to ascertain whether the conclusions and predictions proposed here can be

generalised to a wider population, or whether theoretical revisions will be necessary to account for individual and cross-linguistic differences.

References

Abbot-Smith, K., Lieven, E. V., & Tomasello, M. (2001). What pre-school children do and do not do with ungrammatical word orders. *Cognitive Development, 16*(2), 679 - 692.

Abbot-Smith, K., Lieven, E. V., & Tomasello, M. (2004). Training 2;6-year-olds to Produce the Transitive Construction: the Role of Frequency, Semantic Similarity and Shared Syntactic Distribution. *Developmental Science, 7*(1), 48 - 55.

Akhtar, N. (1999). Acquiring basic word order: evidence for data-driven learning of syntactic structure. *Journal of Child Language, 26*(2), 339 - 356.

Allen, G. & Hawkins, S. (1980). Phonological Rhythm: Definition and Development. In G. Yeni-Komshian, J. F. Kavanagh & Ferguson, C. A. (eds.) *Child Phonology: Volume 1. Production* (pp. 227 – 256). London: Academic Press.

Ambridge, B., Rowland, C. F., Theakston, A. L., & Lieven, E. V. (2006A). Comparing different accounts of inversion errors in children's non-subject wh-questions: What experimental data can tell us? . *Journal of Child Language, 33*, 519 - 557.

Ambridge, B., Theakston, A. L., Lieven, E. V., & Tomasello, M. (2006B). The Distributed Learning Effect for Children's Acquisition of an Abstract Syntactic Construction. *Cognitive Development, 21*, 174 - 193.

Amorosa, H., Vonbenda, U., Wagner, E., & Keck, A. (1985). Transcribing Phonetic Detail in the Speech of Unintelligible Children: a Comparison of Procedures. *British Journal of Disorders of Communication, 20*, 281 - 287.

Andruski, J. E., Kuhl, P. K., & Akiko, H. l., P. K. (1999). *The Acoustics of Vowels in Japanese Women's Speech to Adults and Infants*. Paper presented at The 14th International Congress of Phonetic Sciences.

Ashby, M., Maidment, J. E., & Abberton, E. (1996). Analytic Listening: a New Approach to Ear Training. *Speech, Hearing and Language, 9*(1), 10 - 10.

Baker, C. L. (1971). Stress Level and Auxiliary Behavior in English. *Linguistic Inquiry, 2*(2), 167 - 181.

Barnes, S., Gutfreund, M., Satterly, D., & Wells, G. (1983). Characteristics of Adult Speech which Predict Children's Language Development. *Journal of Child Language, 10*(1), 65 - 84.

Barry, W., & Andreeva, B. (2001). Cross-language similarities and differences in spontaneous speech patterns. *Journal of the International Phonetic Association, 31*(1), 51 - 66.

Bates, E., Dale, P., & Thal, D. (1995). Individual Differences and their Implications for Theories of Language Development. In P. Fletcher & B. MacWhinney (Eds.), *The Handbook of Child Language* (pp. 96 - 151). Cambridge, MA: Blackwell.

Bernhardt, B., & Stöl-Gammon, C. (1994). Nonlinear phonology: Introduction and clinical application. *Journal of Speech and Hearing Research, 37*, 23 - 43.

Bernstein Ratner, N. (1984). Patterns of Vowel Modification in Mother-Child Speech. *Journal of Child Language, 11*, 557 - 578.

Bloom, L. (1970). *Language Development: Form and Function in Emerging Grammars.* Cambridge, MA, London: MIT press.

Bloom, L., Lightbown, P., & Hood, L. (1975). Structure and Variation in Child Language. *Monographs of the Society for Research in Child Development, 40*(2), 1 - 97.

Bowen, C. (1998). Brown's Stages: the Development of Morphology and Syntax Retrieved from http://www.speech-language-therapy.com/BrownsStages.htm

Braine, M. (1963). The Ontogeny of English Phrase Structure: the First Phrase. *Language, 39*(1), 1 - 14.

Braine, M. (1976). Children's First Word Combinations. *Monographs of the Society for Research in Child Development, 41*(1), 1 - 104.

Behrens, S. & Gut, U. (2005). The Relationship between Prosodic and Syntactic Organization in Early Multiword Speech. *Journal of Child Language*, 332(1), 1 – 34.

Boysson-Bardies, B. de, Vihman, M. M., Roug-Hellichius, L., Durand, C., Landberg, I. & Arao, F. (1992). Material Evidence of Infant Selection from the Target Language: a Cross-Linguistic Study. In Furguson, C., Menn, L. and Stol-Gammon, C. *Phonological Development: Models, Research, Implications* (pp. 369 – 391). Timonium, MD: York Press Inc.

Bretherton, I., McNew, S., Snyder, L., & Bates, E. (1983). Individual Differences at 20 Months: Analytic and Holistic Strategies in Language Acquisition. *Journal of Child Language, 10*, 293 - 320.

Browman, C. P., & Goldstein, L. (1987). *Tiers in Articulatory Phonology, with some Implications for Casual Speech. Haskins Laboratories Status Report on Speech Research*, *92*, 1-30

Brown, G. (1990). *Listening to Spoken English*. London: Longman.

Brown, R. (1973). *A First Language: the Early Stages*. Cambridge, MA: MIT Press.

Bryan, S., Howard, S., & Perkins, M. (2010A). *An Investigation of Connected Speech Process Development in One Typically Developing Child*. Paper presented at the 2010 Child Language Seminar.

Bryan, S., Howard, S. J., & Perkins, M. (2010B). *An Investigation of Word Juncture Development in One Typically Developing Child*. Paper presented at the The 31st colloquium of the British Association of Academic Phoneticians.

Buckingham, H. W., & Yule, G. (1987). Phonemic False Evaluation: Theoretical and Clinical Aspects. *Clinical Linguistics and Phonetics, 1*(2), 113 - 125.

Buckley, B. (2003). *Children's Communication Skills: from Birth to Five Years*. New York: Routledge.

Bybee, J. L. (2000). The Phonology of the Lexicon: Evidence from Lexical Diffusion. In M. Barlow & S. Kemmer (Eds.), *Usage-Based Models of Language* (pp. 65 - 86). Junior Leland, Stanford, CA: Centre for the Study of Language and Information Publications.

Bybee, J. L. (2002). Phonological Evidence for Exemplar storage of multiword sequences. *Studies in Second Language Acquisition, 24*(2), 215 - 221.

Bybee, J. L. (2006). From Usage to Grammar: the Mind's Response to Repetition. *Language, 82*(4), 711 - 733.

Bybee, J. L. (2010). *Language, Usage and Cognition*. Cambridge: Cambridge University Press.

Bybee, J. L., & McClelland, J. L. (2005). Alternatives to the Combinatorial Paradigm of Linguistic Theory based on domain general principles of human cognition. *Linguistic Review, 22*(2 - 4).

Catford, J. C. (1977). *Fundamental Problems in Phonetics*. Edinburgh: Edinburgh University Press.

Chiat, S. (1983). Why Mikey's right and My Key's Wrong: the Significance of Stress and Word Boundaries in a Child's Output System. *Cognition, 14*(3), 275 - 300.

Childers, J. B., & Tomasello, M. (2001). The Role of Pronouns in Young Children's Acquisition of the English Transitive Construction. *Developmental Psychology, 37*(6), 739 - 748.

Chomsky, N. (1957). *Syntactic Structures*. The Hague: Mouton.

Chomsky, N. (1964). The Development of Grammar in Child Language. *Monographs of the Society for Research in Child Development, 29*(1), 35 - 42.

Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge, Massachusetts: MIT press.

Chomsky, N. (1966). Current Scene in Linguistics - Present Directions. *College English, 27*(8), 587 - 595.

Chomsky, N. (1967A). Recent Contributions to Theory of Innate Ideas. *Synthese, 17*(1), 2 - 11.

Chomsky, N. (1967B). Some General Properties of Phonological Rules. *Language, 43*(1), 102 - 128.

Chomsky, N. (1976). Conditions on Rules of Grammar. *Linguistic Analysis, 2*(4), 303 - 351.

Chomsky, N. (1980). On Binding. *Linguistic Inquiry, 11*(1), 1 - 46.

Chomsky, N. (1985). *Knowledge of Language*. New York: Praeger.

Chomsky, N. (1995). *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, N., & Halle, M. (1965). Some Controversial Questions in Phonological Theory. *Journal of Linguistics, 1*(2), 97 - 138.

Chomsky, N., & Halle, M. (1968). *The Sound Pattern of English*. New York: Harper and Row.

Clark, R. (1973). Performing without Competence. *Journal of Child Language, 1*(1), 1 - 11.

Cordes, A. (1994). The Reliability of Observational Data 1: Theories and Methods with Speech and Language Therapy. *Journal of Speech and Hearing Research, 37*, 264 - 278.

Cruttenden, A. (1981). Item-Learning and System-Learning. *Journal of Psycholinguistic Research, 10*(1), 79 - 88.

Cruttenden, A. (2001). *Gimson's Pronunciation of English* (6 ed.). London: Edward Arnold.

Crystal, D. (1984). *Linguistic Encounters with Language Handicap*. Oxford: Blackwell.

Crystal, D. (1992). *Profiling Linguistic Disability*. London: Whurr.

Cucchiarini, C. (1996). Assessing Transcription Agreement: Methodological Aspects. *Clinical Linguistics and Phonetics, 10*(2), 131 - 155.

Dabrowska, E., & Lieven, E. V. (2005). Towards a lexically specific grammar of children's question constructions. *Cognitive linguistics, 16*(3), 437 - 474.

Dabrowska, E., Rowland, C. F., & Theakston, A. L. (2009). The Acquisition of Questions with Long-Distance Dependencies. *Cognitive linguistics, 20*(3), 571 - 598.

Dodd, B., Holm, A., Hua, Z., & Crosbie, S. (2003). Phonological Development: a Normative Study of British English-Speaking Children. *Clinical Linguistics and Phonetics, 17*(8), 617 - 643.

Donahue, M. (1986). Phonological Constraints on the Emergence of Two-Word Utterances. *Journal of Child Language, 13*(2), 209 - 218.

Duckworth, M., Allen, G., Hardcastle, W., & Ball, M. (1990). Extensions to the International Phonetic Alphabet for the Transcription of Atypical Speech. *Clinical Linguistics and Phonetics, 4*(4), 273 - 280.

Dye, C. D. (2011). Reduced auxiliaries in early child language: converging observational and experimental evidence from French. *Journal of Linguistics, 47*(2), 301 - 340.

Eastwood, M. P. (1981). Junctural Variations of English Plosive Consonants. *English Language Teaching Journal, 35*(4), 415 - 418.

Ellis, L., & Hardcastle, W. J. (2002). Categorical and gradient properties of assimilation in alveolar to velar sequences: evidence from EPG and EMA. *Journal of phonetics, 30*(3), 373 - 396.

Ellis, N. (2002). Frequency Effects in Language Processing: a Review with Implications for Theories of Explicit and Implicit Language Acquisition. *Studies in Second Language Acquisition, 24*(2), 143 - 188.

Faircloth, M. A., & Faircloth, S. R. (1970). An analysis of the articulatory behavior of a speech-defective child in connected speech and in isolated-word responses. *Journal of Speech and Hearing Disorders, 35*(1), 51 - 61.

Farnetani, E., & Recasens, D. (2010). Coarticulation and Connected Speech. In W. J. J. Hardcastle, J. Laver & F. E. Gibbon (Eds.), *The Handbook of Phonetic Sciences* (2 ed., pp. 316 - 352). Oxford: Blackwell.

Fitch, W. T. (2000) The Evolution of Speech: a Comparative Review. *Trends in Cognitive Sciences,* 4(7), 258 - 267

Foulkes, P., Docherty, G., & Watt, D. (2005). Phonological Variation in Child-Directed Speech. *Language, 81*(1), 177 - 206.

Gertner, Y., Fisher, C., & Eisengart, J. (2006). Learning words and rules: abstract knowledge of word order in early sentence comprehension. *Psychological Science, 17*(8), 684 - 691.

Gick, B. (1999). A gesture-based account of intrusive consonants in English. *Phonology, 16*(1), 29 - 54.

Grunwell, P. (1987). *Clinical Phonology* (2 ed.). London: Croom Helm.

Harris, J. (2003). *Grammar-Internal and Grammar-External Assimilation* Paper presented at the Proceedings of the 15th International Congress of Phonetic Sciences.

Heselwood, B. (2009). A Phenomenalistic Defence of Narrow Impressionistic, Phonetic Transcription as a Clinical and Research Tool. In V. Marrero & I. Pineda (Eds.), *Linguistics: the Challenge of Application.* (pp. 25 - 31). Madrid: Euphonia Ediciones.

Heselwood, B., & Howard, S. J. (2008). Clinical Phonetic Transcription. In M. Ball, J., M. R. Perkins, N. Müller & S. J. Howard (Eds.), *The Handbook of Clinical Linguistics* (pp. 381 - 399). Oxford: Blackwell.

Hewlett, N. (1985). Phonological versus Phonetic Disorders: some Suggested Modifications to the Current Use of the Distinction. *British Journal of Disorders of Communication, 20*, 155 - 164.

Hewlett, N. (1988). Acoustic properties of /k/ and /t/ in Normal and Phonologically-Disordered Children. *Clinical Linguistics and Phonetics, 2*(1), 29 - 45.

Howard, S. J. (2004). Connected speech processes in developmental speech impairment: Observations from an electropalatographic perspective. *Clinical Linguistics and Phonetics, 18*(6 - 8), 405 - 417.

Howard, S. J. (2007). The Interplay between Articulation and Prosody in Children with Impaired Speech: Observations from Electropalatographic and Perceptual Analysis. *Advances in Speech-Language Pathology, 9*(1), 20-35.

Howard, S. J. (2010) Children with Speech Sound Disorders. In J. Damico, N. Müller and M. Ball (eds.) *The Handbook of Language and Speech Disorders* (pp. 339 – 361). Oxford: Wiley-Blackwell.

Howard, S. J. & Heselwood, B. (2002). Learning and Teaching Phonetic Transcription for Clinical Purposes. *Clinical Linguistics and Phonetics, 16*, 371-401.

Howard, S. J., Methley, M., & Perkins, M., R. (2008). *Emergence of Word Juncture in a Typically Developing Child: Evidence of Multiple Interactions in Speech and Language Development*. Paper presented at the 12th Annual Symposium of the International Clinical Phonetics and Linguistics Association.

Howard, S. J., Perkins, M., R., & Raine-Killeen, H. (2008). *Emergence of Word Juncture in a Child with SLI: Evidence of Multiple Interactions in Speech and Language Development*. Paper presented at the 12th Annual Symposium of the International Clinical Phonetics and Linguistics Association.

Howard, S. J., Wells, B., & Local, J. (2008). Connected Speech. In M. J. Ball, M. R. Perkins, N. Müller & S. J. Howard (Eds.), *The Handbook of Clinical Linguistics* (pp. 583 - 602). Oxford: Blackwell.

Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., & Lyons, T. (1991). Early Vocabulary Growth: Relation to Language Input and Gender. *Developmental Psychology, 27*, 236 - 248.

Inkelas, S. & Rose, Y. (2008). Positional Neutralization: a Case Study from Child Language. *Language* 83(4), 707 – 736.

Ingram, D. (1976). *Phonological Disability in Children*. New York: Elsevier.

International Phonetic Association. (2008). IPA Homepage. Retrieved 4th March, 2008, from http://www.arts.gla.ac.uk/IPA/index.html

Kaisse, E. M. (1983). The Syntax of Auxiliary Reduction in English. *Language, 59*(1), 93 - 122.

Karmiloff, K., & Karmiloff-Smith, A. (2001). *Pathways to Language: From Fetus to Adolescent*. Cambridge, MA: Harvard University Press.

Kelly, J., & Local, J. (1989). *Doing Phonology*. Manchester: Manchester University Press.

Kent, R. D. (1996). Hearing and Believing: some Limits to the Auditory-Perceptual Assessment of Speech and Voice Disorders. *American Journal of Speech-Language Pathology, 5*(1), 7 - 23.

Kent, R. D. (1997). Gestural Phonology. In M. J. Ball & R. D. Kent (Eds.), *The New Phonologies: Developments in Clinical Linguistics* (pp.247-268). San Diego, California: Singular Press.

Khattab, G. (2006). *Does Child-Directed Speech Really Facilitate the Emergence of Phonological Structure? the Case of Gemination in Adult CDS*. Paper presented at the Tenth Laboratory Phonology Conference.

Kehoe, M. (1997). Support for Metrical Stress Theory in Stress Acquisition. *Clinical Linguistics and Phonetics*, 12(1), 1 – 23.

Kehoe, M. & Stöl-Gammon, C. (1997). The Acquisition of Prosodic Structure: an Investigation of Current Accounts of Children's Prosodic Development. Language, 73 (1), 113 – 144.

Kidd, E., Lieven, E. V., & Tomasello, M. (2006). Examining the role of lexical frequency in children's acquisition and processing of sentential complements. *Cognitive Development, 21*(1), 93 - 107.

Kuhl, P. K., Andruski, J. E., Chistovich, I. A., Chistovich, L. A. K., E. V., Ryskina, V. L., Stolyarova, E. I., et al. (1997). Cross-Language Analysis of Phonetic Units in Language Addressed to Infants. *Science, 277*(5326), 684 - 686.

Ladefoged, P. (2003). *Phonetic Data Analysis: an Introduction to Fieldwork and Instrumental Techniques*. Oxford: Blackwell.

Langacker, R. W. (2000). A Dynamic Usage-Based Model. In M. Barlow & S. Kemmer (Eds.), *Usage-Based Models of Language* (pp. 1 - 60). Stanford, CA: CSLI Publications.

Lieven, E. V. (2008). Learning the English Auxiliary: a Usage-Based Approach. In H. Behrens (Ed.), *Corpora in Language Acquisition research: History, Methods and Perspectives* (pp. 60 - 98). Amsterdam: John Benjamins.

Lieven, E. V., Behrens, H., Speares, J., & Tomasello, M. (2003). Early Syntactic Creativity: a Usage-based Approach. *Journal of Child Language, 30*(2), 333 - 370.

Lieven, E. V., Pine, J. M., & Baldwin, G. (1997). Lexically-Based Learning and Early Grammatical Development. *Journal of Child Language, 24*, 187 - 219.

Lieven, E. V., Pine, J. M., & Barnes, H. D. (1992). Individual Differences in Early Vocabulary Development: Redefining the Referential - Expressive distinction. *Journal of Child Language, 19*(287 - 310).

Lieven, E. V., Salomo, D., & Tomasello, M. (2009). Two-Year-Old Children's Production of Multiword Utterances: a Usage-Based Analysis. *Cognitive Linguistics, 20*(3), 481 - 508.

Lieven, E. V., & Tomasello, M. (2008). Children's first language acquisition from a usage-based perspective. In P. Robinson & N. Ellis (Eds.), *Handbook of Cognitive Linguistics and Second Language Acquisition* (pp. 168 - 196). Mahwah, NJ: Lawrence Erlbaum Associates.

Local, J. (2003). Variable domains and Variable Relevance: Interpreting Phonetic Exponents. *Journal of Phonetics, 31*, 321 - 339.

Local, J., & Kelly, J. (1986). Projection and Silences: Notes on Phonetic and Conversational Structure. *Human Studies, 9*, 185 - 204.

Local, J., & Walker, G. (2005). Methodological imperatives for investigating the phonetic organization and phonological structures of spontaneous speech. *Phonetica, 62*, 120 - 130.

Local, J., & Wooton, T. (1995). Interactional and Phonetic Aspects of Immediate Echolalia in Autism: a Case Study. *Clinical Linguistics & Phonetics, 9*, 155 - 184.

MacKay, D., Wulf, G., Ying, C., & Abrams, L. (1993). Relations between Word Perception and Production: New Theory and Data on the Verbal Transformation Effect. *Journal of Memory and Language, 32*(5), 624 - 646.

Mackenzie, L. (2012). *English Auxiliary Contraction as a Two-Stage Process: Evidence from Corpus Data*. Paper presented at the 29th West Coast Conference on Formal Linguistics.

MacWhinney, B. (1996). The CHILDES System, *American Journal of Speech-Language Pathology, 5*(1), 5-14.

MacWhinney, B. (2000). *The CHILDES Project: Tools for Analysing Talk* (3rd ed.). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Marcus, G. (1993). Negative Evidence in Language Acquisition. *Cognition, 46*(1), 53-85.

Matthei, E. H. (1989). Crossing Boundaries: More Evidence for Phonological Constraints on Early Multi-Word Utterances. *Journal of Child Language, 16*(1), 41 - 54.

Matthews, D., Lieven, E. V., Theakston, A. L., & Tomasello, M. (2005). The role of frequency in the acquisition of English word order. *Cognitive Development, 20*, 121 - 136.

McGurk, H., & McDonald, J. (1976). Hearing Lips and Seeing Voices. *Nature, 264*, 746-748.

Miccio, A. W., & Scarpino, S. E. (2008). Phonological Analysis, Phonological Processes. In M. Ball, J., M. Perkins, R., N. Müller & S. J. Howard (Eds.), *The Handbook of Clinical Linguistics* (pp. 412 - 422). Oxford: Blackwell.

Morgan, J., & Demuth, K. (1996). *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition*. Mahwah, NJ: Lawrence Erlbaum.

Munson, B., & Brinkman, K. N. (2004). The Influence of Multiple Presentations on Judgements of Children's Phonetic Accuracy. *American Journal of Speech-Language Pathology, 13*(4), 341 - 354

Nelson, K. (1973). Structure and Strategy in Learning to Talk. *Monographs of the Society for Research in Child Development, 38*(1 - 2).

Nelson, K. (1981). Individual Differences in Language Development: Implications for Development and Language. *Developmental Psychology, 17*(2), 170 - 187.

Newton, C., & Wells, B. (1999). The development of between-word processes in the connected speech of children aged between 3 and 7. In B. Maassen & P. Grönen (Eds.), *Pathologies of Speech and Language: advances in clinical linguistics and phonetics* (pp. 67 - 75). London: Whurr.

Newton, C., & Wells, B. (2002). Between word junctures in early multi-word speech. *Journal of Child Language, 29*(2), 275 - 299.

Ogden, R. (1999). A Declarative Account of Strong and Weak Auxiliaries in English. *Phonology, 16*, 55 - 92.

Oller, D. & Eilers, R.E., (1975). Phonetic Expectation and Transcription Validity. *Phonetica 31*, 288-304.

Perkins, M. R. (1999). Productivity and formulaicity in language development. In M. Garman, C. Letts, B. Richards, C. Schelleter & S. Edwards (Eds.), *Issues in Normal and Disordered Child Language: From Phonology to Narrative* (pp.51-57). Reading: University of Reading.

Perkins, M. R., & Howard, S. J. (1995). Principles of Clinical Linguistics. In M. R. Perkins & S. J. Howard (Eds.), *Case Studies in Clinical Linguistics* (pp.?). London: Whurr.

Peters, A. M. (1977). Language learning strategies: Does the whole equal the sum of the parts. *Language, 53*(3), 560 - 573.

Peters, A. M. (1983). *The Units of Language Acquisition*. Cambridge: Cambridge University Press.

Peters, A. M. (1995). Strategies in the Acquisition of Syntax. In P. Fletcher & B. MacWhinney (Eds.), *The Handbook of Child Language* (pp. 462 - 482). Cambridge, Massachusetts: Blackwell.

Peters, A. M. (2001). Filler Syllables: What is their Status in Emerging Grammar? *Journal of Child Language, 28*, 229 - 242.

Pine, J. M. (1992). How Referential are Referential Children? Relationships between Maternal Report and Observational Measures of Vocabulary Composition and Usage. *Journal of Child Language, 19*, 75 - 86.

Pine, J. M., & Lieven, E. V. (1993). Reanalysing Rote-learned utterances: Individual Differences in the Transition to Multi-Word Speech. *Journal of Child Language, 20*, 551 - 571.

Pinker, S. (1994). *The Language Instinct.* London: Penguin.

Pitt, M., A. (1995). The Locus of the Lexical Shift in Phoneme Identification. *Journal of Experimental Psychology: Learning, Memory and Cognition, 21*(4), 1037 - 1052.

Plunkett, K. (1993). Lexical Segmentation and Vocabulary Growth in Early Language Acquisition. *Journal of Child Language, 20*, 43 - 60.

Pye, C., Wilcox, K., & Siren, K. A. (1988). Refining Transcriptions: the Significance of Transcriber Errors. *Journal of Child Language, 15*(1), 17 - 37.

Rechziegel, A. (2001). *Consonants in Contact: on Assimilation and Cross-Language Contrast*. Paper presented at the IFA Proceedings.

Rice, M., Smolik, P., Perpich, D., Thompson, T., Rytting, N., & Blossom, M. (2010). Mean Length of Utterance Levels in 6-Month Intervals for for Children 3 to 9 Years With and Without Language Impairments. *Journal of Speech, Language and Hearing Research, 53*(2), 333 - 349.

Richards, B. (1990). *Language Development and Individual Differences: a Study of Auxiliary Verb Learning*. Cambridge: Cambridge University Press.

Rowland, C. F., Pine, J. M., Lieven, E. V., & Theakston, A. L. (2005). The incidence of error in young English children's wh-questions. *Journal of Speech, Language and Hearing Research, 48*(2), 384 - 405.

Rowland, C. F., & Theakston, A. L. (2009). The Acquisition of Auxiliary Syntax: a Longitudinal Elicitation Study. Part 2: the Modals and Auxiliary Do. *Journal of Speech, Language and Hearing Research, 52*(6), 1471 - 1492.

Scobbie, J. (2002). Context-Conditioned Error Patterns in Disordered Systems. In M. Ball & F. E. Gibbon (Eds.), *Vowel Disorders* (pp. 145 - 185). Boston: Butterworth Heinemann.

Shockey, L. (2003). *Sound Patterns of Spoken English*. London: Blackwell.

Shockey, L., & Bond, Z. S. (1980). Phonological Processes in Speech Addressed to Children. *Phonetica, 37*, 267 - 274.

Shriberg, L. D., KwiatKowski, J., & Hoffman, K. (1984). A Procedure for Phonetic Transcription by Consensus. *Journal of Speech and Hearing Research, 27*, 456 - 465.

Shriberg, L. D., & Lof, G. L. (1991). Reliability Studies in Broad and Narrow Phonetic Transcription. *Clinical Linguistics & Phonetics, 5*(3), 225 - 279.

Simpson, A. (1992). The Phonologies of the English Auxiliary System. In R. Tracy (Ed.), *Who Climbs the Grammar Tree* (pp. 209 - 219). Tübingen: Niemeyer.

Sosa, A. V., & Bybee, J. L. (2008). A Cognitive Approach to Clinical Phonology. In M. J. Ball, M. R. Perkins, N. Müller & S. J. Howard (Eds.), *The Handbook of Clinical Linguistics*. (pp. 480 - 490). Oxford: Blackwell.

Stampe, D. (1979). *A Dissertation on Natural Phonology*. New York: Garland Publishing inc.

Stemberger, J. P. (1988). Between-Word Processes in Child Phonology. *Journal of Child Language, 15*(1), 19 - 61.

Stephens, M. I., & Daniloff, R. G. (1977). A Methodological Study of Factors Affecting the Judgement of Misarticulated /s/. *Journal of Communication Disorders, 10*(3), 207 - 220.

Stöl-Gammon, C., & Sosa, A. V. (2007). Phonological Development. In E. Hoff & M. Shatz (Eds.), *The Handbook of Language Development* (pp. 238 - 256). Oxford: Blackwell.

Stöl-Gammon, C. (2007). Variability in Speech. In S. McLeod (Ed.), *The International Guide to Speech Acquisition* (pp. 55 - 60). New York: Thompson Delmar Learning.

Studdert-Kennedy, M., & Goodell, E.W. (1992). Gestures, Features and Segments in Early Child Speech. *Haskins Laboratories Status Report on Speech Research, 111-112*, 1-14.

Theakston, A. L., & Lieven, E. V. (2005). The Acquisition of Auxiliary Syntax: an Elicitation Study. *Journal of Child Language, 32*, 589 - 616.

Theakston, A. L., & Lieven, E. V. (2008). The influence of discourse context on children's provision of auxiliary BE. *Journal of Child Language, 35*(1), 129 - 158.

Theakston, A. L., Lieven, E. V., Pine, J. M., & Rowland, C. F. (2001). The Role of Performance Limitations in the Acquisition of Verb-Argument Structure: an Alternative Account. *Journal of Child Language, 28*(1), 127 - 152.

Theakston, A. L., Lieven, E. V., Pine, J. M., & Rowland, C. F. (2005). The acquisition of auxiliary syntax: BE and HAVE. *Cognitive linguistics, 16*(1), 247 - 277.

Theakston, A. L., & Rowland, C. F. (2009). The Acquisition of Auxiliary Syntax: a Longitudinal Elicitation Study. Part 1: Auxiliary Be. *Journal of Speech, Language and Hearing Research, 52*(6), 1449 - 1470.

Thompson, J., & Howard, S. (2007). Word juncture behaviours in young children's spontaneous speech production. *Clinical Linguistics & Phonetics, 21*(11-12), 895-899.

Tomasello, M. (1992). *First Verbs: a Case Study of Early Grammatical Development*. Cambridge: Cambridge University Press.

Tomasello, M. (1995). Language is not an Instinct. *Cognitive Development, 10*, 131 - 156.

Tomasello, M. (1998). The Return of Constructions. *Journal of Child Language, 25*(431 - 432).

Tomasello, M. (2000A). Do Young Children have Adult Syntactic Competence? *Cognition, 74*, 209 - 253.

Tomasello, M. (2000B). First Steps toward a Usage-Based Theory of Language Acquisition. *Cognitive linguistics, 11*(1 - 2), 61 - 82.

Tomasello, M., & Stahl, D. (2004). Sampling Children's Spontaneous Speech: how Much is Enough? *Journal of Child Language, 31*(1), 101 - 121.

Van Lieshout, P., & Goldstein, L. (2008). Articulatory Phonology and Speech Impairment. In M. Ball, M. Perkins, N. Müller & S. J. Howard (Eds.), *The Handbook of Clinical Linguistics* (PP. 467 - 479). Oxford: Blackwell.

Vihman, M. M. (1982). Formulas in First and Second Language Acquisition. In

Vihman, M. M. (1996). *Phonological development: The origins of language in the child*. Oxford: Blackwell.

Vihman, M. M. (1999). The Transition to Grammar in a Bilingual Child: Positional Patterns, Model Learning and Transitional Words. *The International Journal of Bilingualism*, 3(2), 267 – 299.

Vihman, M. M., Kay, E., Boysson Bardies, B. de, Durand, C., & Sundberg, U. (1994). External sources of individual differences? A cross-linguistic analysis of the phonetics of mothers' speech to 1-yr-old children. *Developmental Psychology, 30*(5), 651 - 662.

Vihman, M. M., DePaolis, R. and Davis, M. (1998). Is There a "Trochaic Bias" in Early Word Learning: Evidence from Infant Production in English and French. *Child Development*, 69(4), 935 – 949.

Warren, R. M., & Obusek, C. J. (1971). Speech Perception and Phonemic Restoration. *Perception and Psychophysics, 9*, 358 - 362.

Wells, B. (1994). Junction in developmental speech disorder: a case study. *Clinical Linguistics and Phonetics, 8*(1), 1 - 25.

Wells, G. (1979). Learning and Using the Auxiliary Verb in English. In V. Lee (Ed.), *Language Development* (pp. 250 - 270). London: Open University Press.

Wray, A. (2002). *Formulaic Language and the Lexicon*. Cambridge: Cambridge University Press.

Wray, A., & Perkins, M. R. (2000). The Functions of Formulaic Language. *Language and Communication, 20*(1), 1 - 28.

Wright, S., & Kerswill, P. (1989). Electropalatography in the analysis of connected speech processes. *Clinical Linguistics and Phonetics, 3*(1), 49 - 57.

Zwicky, A. M. (1970). Auxiliary Reduction in English. *Linguistic Inquiry, 1*(3), 323 - 336.