German and English Comparison of Fluency Development and Stuttering

Katharina Dworzynski

Thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

University College London

ABSTRACT

The thesis aimed explored disfluencies in stuttering through linguistic contrasts between English and German. Part one focused on the analysis of speech samples of German speaking adults and children who stutter. Part two analysed bilingual German-English language development aiming to examine whether increased cognitive load of two languages is related to language errors. Since people who stutter do not speak disfluently all the time, but have stretches where speech is fluent, research has investigated whether there is a consistent pattern predicting fluency breakdown (Brown, 1937, 1938a, 1938b, 1938c, 1945; Johnson & Brown, 1939). Inconsistencies across different languages would weaken purely motoric accounts of stuttering. This was analysed with spontaneous speech samples of German speaking people who stutter, from a wide age range (2 years to adult). Previously, mainly in English, an exchange pattern of function and content word disfluencies with age was reported (see Au-Yeung, Howell and Pilgrim, 1998, and Rommel, 2000, for increased function word disfluencies in German children). This pattern was more prominent in German speakers changing from more function words in children to predominantly content word disfluencies in adults. Larger amounts of content word disfluencies in German adults were related to their higher phonetic complexity in comparison to English. Part two of the thesis dealt with factors affecting language development in bilingual infants (language onset to school age) based on the link, identified in the literature, between bilingualism and onset of disfluencies. This part aimed to explore differences in lexical and syntax development (both are influential in the onset of disfluencies in infants who stutter - Bernstein Ratner, 1997), naming errors, and lexical access. The results showed differences concerning the lexicon (i.e. compound nouns – complex content words) and syntax (i.e. word order).The last chapter highlighted aspects of findings that differentiate current models / theories of fluency failure.

CONTENTS

FIGURES

TABLES

ACKNOWLEDGMENTS

There are numerous people who have contributed extensively to the work that is presented here. I would like to thank Prof. Peter Howell for his help and guidance – his level of commitment to, and knowledge of the research is incredible. I am also indebted to Dr. James Au-Yeung for his support and whose in depth knowledge of linguistics was invaluable. Also a 'thank you' to Steve Davis and Stevie Sackin for their encouragement.

During the time that this research was conducted I was funded by a Departmental Research Demonstratorship from the University College London – Psychology department.

Both Dr. Dieter Rommel and Dr. Ulrich Natke have helped me obtaining speech samples of German speakers who stutter, without which this research could not have been carried out. In addition, there were self-help groups and Sprachheilzentren (speech therapy centres - particularly Werscherberg) in Germany where recordings were made. Peter Schneider at the University Clinic in Aachen was also particularly helpful. A special thanks goes to all the speakers who were recorded and therefore enabled me to carry out this research. In chapters four and six there is English control data which was not collected by myself, but which I have re-analysed in these chapters – a thank you to those researchers / students who collected this data.

There were a number of schools that have made time and space to help make this research possible and I am indebted to the help of staff and students:

Deutsche Schule, London
European School, Culham
Kreuzschule, Münster (Germany)
Elterninitiative Kotenbeis

I greatly enjoyed the work with the bilingual infants (chapter seven) and thank both them and their parents for their time, effort and hospitality.

Finally, VIELEN DANK to my two nieces, Maja and Anna who played an integral part in making the connection with the school in Germany and the play group of the Elterninitiative Kotenbeis (tested in chapters 6 and 8).

On a personal level there was invaluable support and help by Helen, Tim, my parents, family and friends (particularly Des, Joachim, Conny, Gra, Catherine, Lynda, Sean + Cath + the boys, Carsten, and Simone).

1 Investigating stuttering using a psycholinguistic perspective – the application of cross-linguistic and comparative research

"One must let people talk, since fish can't."
Proverbial (Polish)

1.1 Introduction and Aims

This thesis investigated stuttering from a comparative linguistic perspective. In the first part of this work (chapters 2, 3 and 4) the words on which stuttering occurred were analysed in samples of German speakers who stutter and who were from different age groups. This was then compared to established findings in English samples. In the latter part of the thesis (chapters 6, 7 and 8) bilingual children and their language development were the focus. This line of investigation was based on a link between the onset of stuttering in early childhood and bilingualism. Furthermore the same linguistic structures that were found to be problematic in the first part were also analysed in bilingual children. The current chapter is outlining the background to the linguistic examination of stuttered speech and is providing the rationale for the study of fluent bilingual infants' language development.

One of the puzzling facts about stuttering is that people who stutter are not disfluent all the time. In other words, there are stretches of fluent, interspersed with, disfluent speech. Consequently, research over many years has investigated whether there are consistent properties in speech that is stuttered that distinguish it from a stretch of speech that is not stuttered (work pioneered by Brown, 1937; 1938a; 1938b; 1938c; Johnson & Brown, 1935; 1939). A second important issue about stuttering is that relatively little is known about the causes of the disorder and what contributes to the onset of stuttering in childhood (as described by Bloodstein, 1995 for instance). There is also no known pattern predicting which children recover from their stutter spontaneously early on and which ones persist in speech problems (a point which has been raised by e.g. Yairi & Ambrose, 1992).

The main approach, when examining these issues, has been to investigate the linguistic properties of words that are stuttered across different age groups. In the following review the original work on this line of investigation is given (Brown's work with adults who stutter, 1945). This work was interpreted as suggesting that the individual linguistic factors that affect stuttering operate because they all affect the semantic difficulty of an utterance (Brown, 1945). Subsequent work that evaluated and extended the linguistic determinants on stuttering is then described. This work was largely empirical and has been predominantly analysed with English speaking people who stutter. The latest research is then outlined that ties the findings to particular speech production models. Two illustrative models of fluency failure are outlined (the EXPLAN model, Howell & Au-Yeung, 2002; and the covert repair hypothesis, Postma & Kolk, 1993) that have both been applied to stuttering.

New avenues to investigate processes that operate at the time of stuttering onset in childhood are then outlined. The reasons for cross-linguistic and comparative analysis, which is the approach taken throughout this thesis, are then introduced. It is highlighted what the study of normal fluency development at the age that is equivalent to stuttering onset in some children, can contribute to the knowledge of the disorder. More specifically the rationale is given for an investigation of language development of bilingual children and how this can shed light on the linguistic factors that operate in people who stutter. This is then linked to factors that have been suggested to lead to the onset of stuttering.

The last part of the review describes the differences between English and German that potentially afford ways of differentiating between the competing theories.

1.2 Brief Background to Stuttering

The literature on stuttering in general is vast, ranging from investigations into stuttering in songbirds (Rosenfield, Viswanath & Helekar, as reported by Walker, 1999), to the analysis of cartoon characters who stutter (such as the clinical study of Porky Pig, Johnson, 1987). Necessarily, the review needs to focus on definitions and methods of study that can be related to linguistic determinants.

There have been numerous attempts at defining stuttering. Early definitions of this condition relied heavily on the theorists' assumptions about the underlying causes of the stutter. Quite often researchers did not even include a description of the relevant speech characteristics (that precludes them from detailed investigation in this review). Some of the causes invoked were emotional conflict, fear, and conditioned emotion (see Hahn, 1956 for an overview of early theories). More recently the focus has shifted to a description of the nonfluency types displayed by persons who stutter, such as Van Riper's so-called core behaviours (Van Riper, 1971; 1981) – discussed below. Subsequent definitions of stuttering also reflect this shift and provide descriptions of the nonfluency types displayed by people who stutter. Peters and Guitar (1991), for instance, write that stuttering is characterised by an abnormally high frequency and/or long duration of stoppages in the forward flow of speech. This definition was inspired by previous work by Andrews and Harris' (1964) and Wingate (1964). For instance, Wingate (2002) thinks "(1) its cause is unknown, (2) its essential nature is not understood, and (3) there is no known cure" and, furthermore, suggests that there is no developmental aspect to stuttering (as put emphatically on p.372 "Stuttering does not 'develop'"). Most researchers consider that the onset of stuttering may occur at any time during childhood, between the beginning of multiword utterances (18 month) and puberty (Bernstein Ratner, 1997). It has been found most likely to emerge between ages 2 and 5 (Andrews et al., 1983). On the prevalence of the disorder, (i.e. the occurrence in the population at large), Bloodstein (1987) analysed the results of 38 studies world-wide and concluded that 1% of children stutter at any given time (this same percentage was given by Andrews et al., 1983). Andrews et al. (1983) also found that there is a decline in prevalence after puberty that continues into adulthood where the percentage is less than 1%. Incidence figures indicate how many people have stuttered at any point in their lives. This depends on how long stuttering has to persist to be counted as an "incident". When only stuttering that lasted longer than 6 month is included, the incidence is about 5% (Andrews et al. 1983), whereas when children who only stutter for brief periods are included the incidence is 15% (Bloodstein, 1987). The high remission rates have practical implications for treatment. Peters and Guitar (1991) found that between 50 and 80% of children who stutter will recover with or without professional treatment, most before puberty. With respect to sex ratio, results from numerous studies of people who stutters over a wide age range (for instance see Leske, 1981; Louttit & Halls, 1936; Wallin, 1916) and from many cultures (research dating from 1890 to the present) put the ratio at about three male to every female person who stutters. It has also been reported that the sex ratio imbalance increases as children get older. Bloodstein's (1987) review of research literature indicated that the sex ratio is about 3:1 in the first grade and 5:1 by the fifth grade. Research indicates that girls begin to stutter a little earlier (Yairi, 1983), but also recover earlier (Andrews et al., 1983). Reviewing the data from several previous studies Andrews et al (1983) in respect of hereditary aspects of stuttering, the authors estimated that the incidence of stuttering among first degree relatives of people who stutter was more than three times that of the general population. The finding of a sex ratio favouring girls together combined with research indicating that people who stutter have often relatives with the same disorder suggests some form of genetic transmission of the disorder. However, a chromosome responsible for, or a model which is entirely compatible with, the observed familial pattern has not yet been identified (see Kidd, 1980 ; 1984 for genetic models and; Yairi, Ambrose, & Cox, 1996 for a critique).

1.3 Types of Disfluencies in Stuttering

As described above the shift in the definitions of stuttering has placed the focus on the so-called core behaviour of stuttering (Van Riper, 1971, 1981) – i.e. a description of the actual speech disfluencies occurring during stuttering. Wendell Johnson was the original authority who gave impetus to evaluating speakers with respect to the type and frequency of the disfluencies they showed. He described eight features of the speech of people who stutter that would create stoppages / delays to the forward flow of speech, (Johnson, 1961):

1. Interjections of sounds, syllables, words, or phrases.
2. Part-word repetitions.
3. Word repetitions.
4. Phrase repetitions.
5. Revisions
6. Incomplete phrases.
7. Broken words.
8. Prolonged sounds.

This list of descriptors has been extensively used in research on speech disfluencies, as it has in comparative work on the performance of people who stutter and normally fluent speakers (e.g. Adams, 1982; Bloodstein & Grossman, 1981; Hedge, 1982; Silverman, 1974).

1.4 Background to the early Psycholinguistic Investigation of Stuttering

In this section the term *stuttering* is referring to all the events on Johnson's (1961) list. As highlighted in the introductory paragraph, one important aspect of stuttering is that people who stutter do not do so all the time, but have episodes where their speech is fluent. Consequently research into stuttering has tried to clarify why this is the case and where in an utterance disfluency occurs. Typically these use units for linguistic analysis longer than single words. The linguistic analysis of disfluent speech is a relatively recent area in the research of stuttering. The interest in this field started in the 1930s when there was a widespread interest in semantics, i.e. words and their meaning. Broadly this work was inspired by the influential writings of Korzybski (1933) and Whorf (1940). Korzybski coined the term 'semantogenic' which translates into 'caused by words', whereas Whorf hypothesised that speakers view the world along lines laid down by their native language (linguistic relativity or Sapir-Whorf hypothesis). This focus on word meaning seems to be specifically related to content word disfluency which is discussed in a later section. Wendell Johnson, one of the most prominent figures in stuttering research during this time, was strongly influenced by this work in semantics (in fact his theory of stuttering is called the semantogenic theory). He particularly focused on how peoples' evaluations of events influences how they deal with them. He also used these evaluations in his approach to counselling people who stutter. Two phenomena that followed from the semantogenic viewpoint were the 'adaptation effect' and the 'consistency effect'. The reason for this was that the meaning of a text becomes less central when repeated. Adaptation refers to the fact that, when individuals who stutter read out a text repeatedly, their stuttering decreases over the course of sequential readings (adaptation effect), as does also the tendency for stutters to recur at the same places (consistency effect). Words such as adaptation and consistency stem from the behaviourists' approach and Johnson interpreted both of these phenomena in light of a stimulus – response (S-R) approach. In this respect he suggested that the person who stutters feared 'difficult sounds' (consistency effect) and that habituation to these stimuli would set in after several trials (adaptation effect). The first author who investigated this consistency effect, in respect of linguistic factors of the words that were more frequently stuttered, was Spencer F. Brown – one of Johnson's students. Although his series of studies/articles did not spark major interest at the time, his seminal work can be viewed as the first formal investigation of language factors that affect stuttering. Though there were some excellent individual articles (for instance see Conway & Quarrington, 1963) Brown's work was publicised and gave an impetus to this area of research 30 years on with the publication in 1988 of Marcel Wingate's "Structure of stuttering".

Word factors were the main focus of examination in Brown's perspective (Brown, 1937; 1938a; 1938b; 1938c; 1943; Brown, 1945; 1942; Johnson & Brown, 1935; Johnson & Brown, 1939). His analysis of stuttered speech revealed that there were four main word factors that make it more likely that a word would be spoken disfluently (though his work had found others too, such as stress). Later these became known as Brown's 'four factors'. These factors are that disfluencies are more likely to occur when: 1) The grammatical category of the word which was later noted to correspond to the content and function word distinction; 2) the word is long; 3) words appear in early positions in a sentence; 4) the word starts with a consonant. It is particularly noteworthy here that all of these aspects make specific reference to words, and influences outside these bounds are not considered (even in the case of sentence position where an ordinal measure is used rather than a syntactic criterion). He argued, consistent with the semantogenic theory, that the underlying cause of all the factors (apart from factor 4) he identified was semantic difficulty. This last point is taken up again when the different patterns of children and adult individuals who stutter are discussed in a later section.

The psycholinguistic review of stuttering by Wingate (1988) raised two important conceptual points. First, in contrast to 'semantic difficulty' as an underlying cause for stuttering (as suggested by Brown), Wingate emphasised linguistic stress (which as pointed out earlier, Brown 1938 had found to have an effect but did not include in his main factors) as the principal factor that leads to disfluencies in adult individuals who stutter. This raises the wider issue of confounds – some researchers use content as a more heuristic category pushing the effects of word length and consonantal factors (e.g. Quarrington, Conway & Siegel, 1962) others focus more on word frequency (Dayalu, Kalinowski, Stuart, Holbert & Rastatter, 2002). Heuristic in the sense that these are the words that carry the meaning of the message (i.e. the content) and therefore it is hard to distinguish cause and effect. These would be the words most frequently carrying stress which, Wingate (1988) also argued, was the underlying reason why Brown's factors increased the likelihood of disfluencies. Second, Wingate also highlighted the need to relate observations of stuttered disfluency to the work on fluent speakers' speech errors. Fluent speaker's speech errors are revisited in the section discussing the way stuttering changes across age groups. In respect to Brown it can be argued that he was not looking centrally at linguistic factors, but was more concerned with the relationship between introspective reports by persons who stutter and linguistic aspects of the words. In other words Brown highlighted "struggle" that can be in anticipation of a "feared word" (see Wingate, 2002). In other words Brown interpreted the linguistic factors not as a way of specifying speech motor difficulty but rather interpreted stuttering as a conditioned response to certain types of words.

1.5 Stuttering and the Continuity Hypothesis (Bloodstein, 1987; 1995)

Fluency breaks down in fluent speakers on occasions. These speech errors, e.g. phonological exchange patterns, such as spoonerisms for instance, word finding problems, repetitions of monosyllabic words, and pauses have been extensively studied in psycholinguistics to provide insights into language processing (see for instance Fromkin, 1971; Garrett, 1975; 1980). The relationship between fluent speakers' speech errors and disfluencies of people who stutter, was emphasised by Wingate (1988) and is also highlighted by Bloodstein (1987; 1995) in the 'Handbook on Stuttering'. Since the latter part of the thesis is looking at fluent speakers' early speech acquisition a bit more detail is given here. Bloodstein (1987;1995) assumed that the relationship between speech errors of fluent speakers and those experienced by people who stutter could be placed along one spectrum on severity (i.e. stuttered events as severe cases of normally occurring speech errors), which was to be called his 'continuity hypothesis' of stuttering. His view on the continuity hypothesis of stuttering is best exemplified by the following quote:

"In all probability, the question of how to differentiate between stuttering and normal nonfluency can never have an absolute answer. The only distinction which one can validly make appears to be a purely relative one between struggle reactions which are mild and occasional and those which are more severe and persistent (Bloodstein, 1961)."

Bloodstein (1995) also stressed two further points in his review, which are relevant to what is described in later sections. One is his interpretation that stuttering might occur because of a difficulty with a word yet to be produced (which is the struggle hypothesis), the other point is that higher speech rates can also result in stuttering.

The way stuttering changes across age groups (as described in more detail in the next section) can be reinterpreted in light of one of the conceptual points Wingate (1988) raised (i.e. his focus on the relationship between fluent speakers' disfluencies and observations about stuttered speech). Supporters of the avoidance theory of stuttering disfluencies (avoidance theory as first proposed by Johnson & Associates., 1959) would assume that there is a clear distinction between stuttering and normal nonfluencies. However, an underlying assumption for a comparison between the disfluencies fluent speakers exhibit and those of people who stutter, is that what is referred to as a stutter is seen as a somewhat extreme degree of certain types of disfluencies that can be found in fluent speakers (the continuity hypothesis – Bloodstein, 1995, Brown, 1945). This perspective can help identify similarities between fluent speakers and children who stutter, and the processes that change when children persist in their stutter (Conture, 1990).

Properties of disfluencies in <u>fluent</u> speakers' speech (an appropriate basis of comparison according to the continuity hypothesis) were described by Maclay and Osgood (1959) and later by Clark and Clark (1977). Both noted that whole word repetitions and pausing occur frequently in fluent speakers' speech. Clark and Clark (1977) proposed that such repetition and hesitation occurs when the speech plan for a later word is not ready for execution. This could mean that both repetitions and pausing prior to a problem word buy time by delaying the execution of the subsequent word. How the latter idea is related to general psycholinguistic findings and how it led to a recent model of fluency breakdown is described in the following sections.

More recent research led to a re-evaluation of these early findings and the conclusions that were drawn. A phonological concept that was used in the analysis of disfluent speech in chapter four and has direct implications for models of fluency failure, is the phonological word (PW). This is briefly outlined in the next section.

1.6 The phonological word (PW)

The *phonological* (often also called *prosodic*) *word* is introduced very briefly in this section since it is described in depth in chapter four, where this concept was used in the investigation of stuttered speech. In linguistic theory the PW is a concept which is distinct from the grammatical / morphological word (see Hall & Kleinhenz, 1999). It is generally viewed as an important stage between the planning and execution of speech, since it provides the domain for syllabification (Wiese, 1996). As such it has to be finalised before syllables are assigned and speech production can commence. The theory of fluency failure (the EXPLAN theory Howell, 2002; Howell & Au-Yeung, 2001; Howell & Au-Yeung, 2002) that is central to this thesis assumes that disfluencies of certain grammatical word types (*function words*) are a delaying strategy for the planning of a later more complex grammatical word (a *content word*) to be finalised. To study this process the phonological word is used as a research tool since it provides a concept that extends beyond a single word. Function words are put in direct relation to surrounding content words in PWs. With this concept researchers can make clear predictions about where disfluencies ought to occur on function words in an utterance and where such disfluencies should be unlikely.

The background and description of the EXLAN model of fluency failure is given in section 1.11.2. It now follows a more in depth introduction to the two main grammatical classes of words. How these word classes were central to the re-evaluation of the early linguistic research is also highlighted in the next section.

1.7 Reinterpretation of the Early Research

One property highlighted by Brown's work was word class. Though he looked at a variety of word classes, the main feature of his data that has been noted by subsequent researchers is the difference between stuttering rates on content and function words. For his adults, stuttering rate was higher on words that would be classified as content words than function words. Definitions of function and content words are given to commence looking at this in more detail. Function words (pronouns, articles, prepositions, conjunctions and auxiliary verbs) are a closed class of words that do not carry a full lexical meaning. They have a grammatical or functional role (Hartmann & Stork, 1972; Quirk, Greenbaum, Leech, & Svartvik, 1985). Content words (nouns, main verbs, adverbs and adjectives) are an open class of words, and play a crucial role in conveying semantic information. These are the words that convey meaning to a sentence, whereas function words are the grammatical fillers. This again links in with the early linguistic analysis of stuttering and its focus on semantics.

The class of function words contains a set of highly practised high frequency words which in English are predominantly monosyllabic. Thus disfluencies on these types of words are surprising, because, at least in the case of English, they are linguistically simpler than the content words. In this respect the function/content word distinction is one of the key issues and provides a challenge for a unified theory of stuttering that is relevant for both the developmental aspect of stuttering and its established form. Content words as a grammatical word class can take the role of a heuristic, since it contains most of the other factors associated with stuttering such as linguistic stress, complex onsets, word length, word frequency and their different syntactic roles (confounds highlighted at an early date by Quarrington, Conway and Siegel, 1962).

In contrast with Brown's (1945) observations about adults, Bloodstein and Gantwerk (1967) and Bloodstein and Grossman (1981) have reported that children who stutter are more likely to be disfluent on function words than on content words. For example in an utterance like 'I stretch it', the childhood stuttering pattern could be 'I I I stretch it' whereas an adult who stutters could say 'I sssssstretch it' or 'I s s stretch it'. In this example it can also be seen that the content word, 'stretch', is linguistically more complex than the function words surrounding it, i.e. it starts with a complex consonant string. Findings like these also have consequences for the underlying causes suggested by Brown (1945) and Wingate (1988). Brown's hypothesis, that the underlying cause why his factors operate is semantic difficulty, cannot explain why children have difficulties on the semantically simpler function words. Even though these word classes appear later developmentally (see for instance Bates & Goodman, 2001) they are a word class of syntactic rather than semantic complexity. Similarly Wingate's emphasis on the role of stress can only account for the pattern observed in adults who stutter. In English, stress is related to the function and content word type. Function words are not usually stressed, while every content word carries a word stress. As such Wingate's emphasis on the role of stress is less applicable to children's speech, due to the prevalence of disfluency on function words (these are rarely stressed in these speakers). However, in dismissing word repetitions as stuttering events he gets around this problem. In general the fact that children stutter on different types of words is ignored by most developmental theories of stuttering. For instance, the theories that view stuttering as a learned response to difficult words (such as the anticipatory avoidance theory, Johnson, 1938; approach avoidance conflict theory, Sheehan, 1953; or the preparatory set, Van Riper, 1937 to name but a few). Again, these theories cannot explain why children would predominantly stutter on the easier function words.

1.8 Cross-linguistic Research into Stuttering

The heuristic value of the function-content word distinction for English has been pointed out already (i.e. content words have different properties to function words and the properties of content words make them more likely to be the focus of stalling or advancing disfluencies as described in more detail in the theories of fluency failure section 1.11). Howell, Au-Yeung and Sackin (1999) analysed the structure of function words in spontaneous speech of several different age groups to examine the simplicity of English function words over ages. They found that 95% of function words used were monosyllabic, 89% of content words started with consonants whereas only 54% of function words did (these are two of the phonological factors Throneburg, Yairi, & Paden, 1994, investigated - this will be given more detail in a later section). Howell et al. also reported that 99% of function words have no primary stress.

As described in the introduction to the second chapter Bloodstein (1995) in his review highlighted the fact that there is a lack of structured investigations into linguistic factors predicting stuttering in languages other than English. Van Borsel and colleagues also stress that most of the research has been conducted with monolingual, usually English speaking, speakers (Van Borsel, Maes, & Foulon, 2001). Cross-linguistic or comparative research may be ideal to validate theories that propose that specific linguistic factors precipitate stuttering (as pointed out by Bernstein Ratner & Benitez, 1985).

Another reason to carry out cross-linguistic research is to find out whether these stuttering patterns are universally associated with linguistic structures irrespective of their surface form; or whether difficult motor outputs lead to stuttering independent of the linguistic unit they occur in. This dissociation between motor and linguistic aspects can be achieved because the same phonetic structures may appear in different linguistic units in different languages. Even though other languages allow scope for separating motor properties from the linguistic units in which they occur in English, no previous studies have made such comparisons. Thus the main purpose of this work is to address the general issue concerning why disfluency occurs on certain words and also, more specifically, focus on the question whether linguistic or motor factors affect disfluency.

Although only in its infancy, cross-linguistic research has now been started by Howell and colleagues that investigates these differences using the phonological word (PW – a linguistic concept that is described in detail in chapter four). For instance in

English function words are usually short and unstressed, whereas this is not always the case in Spanish. In Spanish function words can carry primary stress and can sometimes be longer than content words. A study has shown that the function/content word distinction predicted the location of a disfluency better than word stress (Au-Yeung, Vallejo Gomez & Howell, 2003). Other work with Spanish looked at stress (as opposed to content) word definitions of PW that should assess whether one or both factors operate in leading to stuttering (Howell, submitted). It is also investigated whether the possibly more complex structure of German content words leads to differences in stuttering patterns (Dworzynski, Howell, Au-Yeung, & Rommel, 2003).

1.8.1 Bilingualism and Stuttering

There have only been two recent studies that have looked at the relationship between bilingualism and stuttering (Au-Yeung, Howell, Davis, Charles, & Sackin, 2000; Van Borsel et al., 2001). Some researchers draw a direct causal link between the occurrence of mixed utterances (by the parents and onset of stuttering in children (Lebrun & Paradis, 1984). Lebrun and Paradis (1984) based their theory on prior work by Pichon and Borel-Maisonny (1964) who suggested that a lack of verbal immediacy in bilinguals leads to disfluencies and also on research suggesting a higher percentage of people who stutter in bi- or multilingual populations (1.80% vs. 2.80% and 2.38%, respectively Travis, Johnson, & Shover, 1937). Bloodstein (1995) cited a similar finding by Stern (1948) where prevalence of stuttering in the monolingual group was 1.66% whereas it was 2.16% in the bilingual group. In a case study by Karniol (1992), the author also drew a direct causal link between bilingualism of a child and stuttering onset. She describes how stuttering started when the bilingual child (a child with Hebrew and English speaking parents - from language onset 1;2 to 3;0 years) started to produce grammatical sentences. When the non dominant language was dropped the child became a non-stuttering speaker in his first language. Her conclusion was that stuttering onset is precipitated by syntactic overload which is particularly high in bi- or multilingual language acquisition. Such points of view have implications for the assessment and therapeutic intervention for these children and numerous researchers have advised in such cases that the number of linguistic systems to which a bilingual child is exposed be temporarily reduced (see for instance Biesalski, 1978; Eisenson, 1986; Rustin, Botterill, & Kelman, 1996). Details of methodological difficulties in the collection of prevalence data, assessment and therapeutic issues are beyond the scope of this thesis and readers are referred to the above mentioned survey and review.

The demands and capacities framework (see also Adams, 1990; Andrews et al., 1983; and for a critique Siegel, 2000; Starkweather, 1987; Starkweather, 1997; Starkweather & Gottwald, 1990; Starkweather, Gottwald, & Halfond, 1990; 2000) has been used to highlight the multidimensional nature of fluency development and is a useful way of characterising early bilingual development. According to this model fluency failure occurs when the challenges (or demands such as speech rate, continuity of production etc) exceed the capacities (for instance speech motor control, language formulation, social and emotional maturity, and cognitive skill) of the child – similar to the lack of immediacy suggested by Pichon and Borel-Maisonny (1964). For children who are brought up in bilingual environments – it is estimated that nearly half of the world's population is functionally bilingual (as pointed out by de Houwer, 1995) – the task of lexical and syntactical development is doubled for the two languages acquired. If, as Bernstein Ratner (1997) suggested, increases in vocabulary interact with syntax, leading to retrieval problems and disfluencies, this should be even more likely in bilingual children since they have to learn concurrently two different lexicons and grammars. To phrase it in the language of the demands and capacities model, it could be assumed that children growing up with two languages possess the same capacity for language competence but are faced with higher demands since they have to acquire two linguistic rule systems.

For the purpose of the current work a brief outline is provided of studies that have looked at the linguistic pattern of stuttering in individuals with more than one language. The majority of cases in the literature, reporting on patterns of disfluencies in more than one language, indicate differences in frequency, distribution and / or disfluency types between the two languages spoken. This is referred to as the difference-hypothesis (Nwokah, 1988).

1.8.2 Linguistic patterns in bilingual stuttering research

Ten bilingual English – Kannada (a Dravidian language spoken in South India) speakers who stutter were studied by Jarayam (1983). Although no difference in sound pattern distribution was found, subjects were reported to stutter more frequently in Kannada than in English. Another study looked at disfluency patterns of an English – French speaking infant (Shenker, Conte, Gingras, Courcey, & Polomeno, 1998). Here more word repetitions were observed in French but more part-word repetitions in English. An adult Spanish – English bilingual person who stutters was described by Bernstein Ratner & Benitez (1985). This subject was nearly twice as disfluent in Spanish as compared to English. A more detailed analysis showed that, particularly words starting in vowels (although stuttered more frequently in both languages) were stuttered twice as often in Spanish. Their explanation for this was based on differences of sentence structure between Spanish and English. Cabrera and Bernstein Ratner (2000) reported results from the analysis of the speech of a five-year old Spanish – English bilingual boy. Here higher proportions of stuttering on reflexives was shown in Spanish whereas in English adjectives were stuttered more frequently. This was also explained to be a positional effect since reflexives occur before a noun in Spanish whereas adjectives occupy this position in English. The only study that showed a difference in patterns and distribution of stuttered sounds analysed sixteen English – Igbo (a Nigerian language) speakers (Nwokah, 1988). All of them stuttered more in one language (either English or Igbo). On the whole English words that had initial consonants were stuttered more frequently than those that started with a vowel, whereas the opposite pattern was observed for Igbo. There is one thing all of the above studies have in common. This is that stuttering was observed to be more frequent in the dominant language (L1). In fact only two studies showed a pattern where the frequency of stuttering was higher in the less proficient language (L2 - an English - Afrikaans speaking adult, as described by Jankelwitz & Bortz, 1996; and a Spanish - English speaking adult, Scott Trautman & Keller, 2000).

As an interim summary, the review so far introduced facts that have been established about stuttering relevant to the current work. This led into a detailed description of findings regarding the linguistic factors involved in stuttering and their subsequent re-interpretation in light of the pattern observed in young children who stutter. A description of cross-linguistic and bilingual research into linguistic variables of stuttering followed. The causal link between bilingual children experiencing mixed utterances in their environment and the onset of stuttering was highlighted. Links were established between linguistic aspects that precipitate stuttering onset (rapid lexical and syntax acquisition) and might cause fluency problems in bilingual language acquisition. It has been described in section 1.2 that the literature on stuttering in general is vast and that some definitions of stuttering have not even included a description of the speech characteristics of stuttering. Since the approach of this thesis is to explain some aspects of stuttering from a speech production background the next section is taking a step back and outline different theories of fluent speech production. This is followed by a introduction to processing issues in bilingual language development. Then the two models of fluency failure that are contrasted in this thesis are described. The literature reviewed is then linked to these theories and leads into a detailed plan and rationale for the current work.

1.9 Psycholinguistic Models of Speech Production

As pointed out by Levelt (2000, p. 844) '*the functional architecture of language use, i.e., its organisation as a behavioural system, is the domain of psycholinguistics*'. Such an enterprise would address the particular issue examined in this thesis, namely, the way the behavioural system functions during fluent speech and its breakdown. The continuity hypothesis implies that findings from models of fluent speech control should have important theoretical implications for models of disfluency. One widely accepted finding from recent psycholinguistic research, that is important for stuttering research, is the strong evidence in support of the view that lexical access in speech production takes place in a two-stage process (see for instance Butterworth, 1989; Garrett, 1975; Garrett, 1980; Levelt, 1992; Martin, Dell, Saffran, & Schwartz, 1994; Schriefers, 1990).

1.9.1 Levelt, Roelofs & Meyer's (1999) Model

An example of such a dual-process model is Levelt, Roelofs, & Meyer's (1999) theory of lexical access in speech production, which is currently one of the most prominent models in the psycholinguistic field. One process is concerned with the selection of words, i.e. the 'lemmas'. It involves the selection of an item from the mental lexicon that appropriately expresses the speaker's intention. At this stage gender information is also retrieved in languages for which this is relevant (Caramazza & Miozzo, 1997; Jescheniak & Levelt, 1994). The second system prepares the appropriate articulatory gestures for these selected words in their utterance contexts.

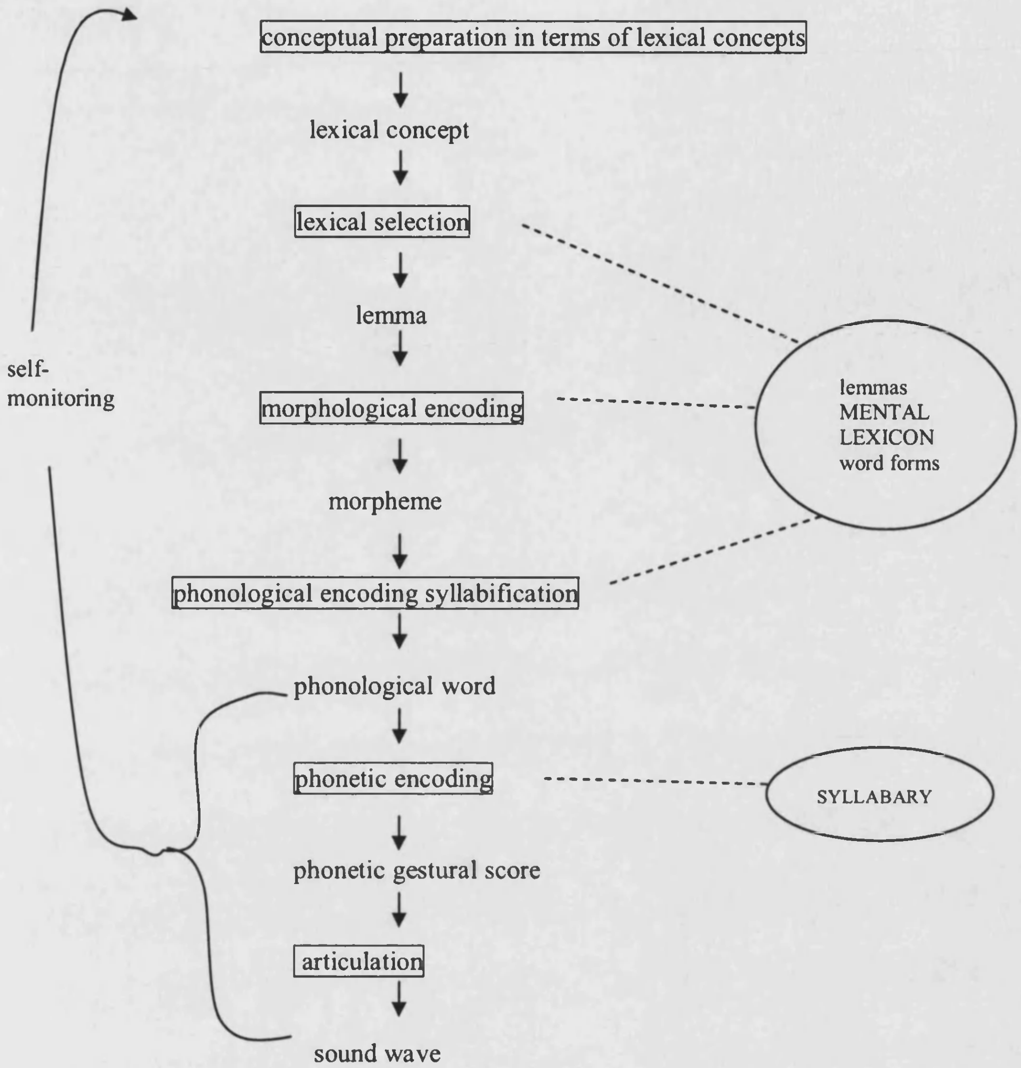


Figure 1: Diagram of Levelt et al's (1999) theory. Before articulation can be initiated, word preparation proceeds through stages of conceptual preparation, lexical selection, morphological and phonological encoding and phonetic encoding. Output monitoring involving the speaker's normal speech comprehension mechanism occurs in parallel.

A brief summary of the six processing stages of Levelt et al's (1999) model is now given.

1.9.1.1 Description of Processing Stages:

I. Conceptual Preparation: This is the stage where a lexical concept is activated. A difficulty to overcome in this stage is the *verbalisation problem.* This is concerned with the question of how a speaker achieves a transition from the notion to be expressed to a message comprised of lexical concepts. This stage involves choosing a lexical concept when there are multiple ways of referring to the same object (levels of specification as in *pig*, female *pig*, *sow*, for instance). Another problem at this stage is the so-called *hyperonym problem.* According to their model once a word's semantic features are active then the feature sets for all of its hyperonyms / superordinates are also active by definition. This refers to the problem of choosing a target word when the semantic features should have also activated all the words of subsets or superordinate words of the same category.

II. Lexical Selection: Here a lemma is retrieved from the mental lexicon, based on the lexical concept that needs to be expressed. This is a relatively robust process which takes place at high speed.

III. Morphophonological Encoding and Syllabification: This is the first stage in the preparation of the articulatory gesture. The tip-of-the-tongue phenomenon has been interpreted as providing evidence about the momentary inability to retrieve the word form, given a selected lemma (e.g. grammatical gender). Three kinds of information are available in the process of word form activation: the word's morphological makeup, its metrical shape (i.e. iambic, di-syllabic and stress final), and its segmental makeup. As already briefly introduced the domain of syllabification is the "phonological" or "prosodic" word (note: this is particularly important for the perspective in this thesis and PWs are discussed at more length in chapter four).

IV. Phonetic Encoding: Phonetic encoding puts more detail onto a phonological word's gestural score. According to Levelt et al. (1999), a syllabary is used from which speakers can access a repository of gestural scores for the frequently-used syllables of the language (presumably this means infrequently-used syllables are generated in some other way). Coarticulatory properties of a word that are syllable-internal are stored in the syllabary (though it is not clear how coarticualtion effects across syllables would arise).

V. Articulation: This is where the word's gestural score is executed by the articulatory system.

VI. Self-monitoring: Internal speech can be checked during speech encoding.

To sum up Levelt et al's (1999) model, it can be stated that speech production takes place in distinctive separate stages. It is a feedforward (i.e. with the information from one stage moving strictly in one direction to the next stage), staged process, leading from conceptual preparation to the initiation of articulation with distinctive output representations for each of the stages. It has recently been labelled as one of the *minimalist* theories of speech production (Vigliocco & Hartsuiker, 2002). This is due to the fact that information flow between the modular stages is unidirectional and contains only the minimal information that is necessary for further processing. The model is very elaborate on the processing of all the stages leading to articulation. In comparison to the descriptive detail on these stages the articulatory process is covered relatively superficially. In conjunction with this theory one of these researchers (Roelofs, 1996a, 1996b, 1997) has also developed a computational model – the Weaver + + model. This model covers all the stages from lexical selection to phonological encoding, using data from studies of word production latencies of fluent speakers.

1.9.2 Challenges and Other Models

MacKay's (1981; 1987) Node Structure Theory (NST), one of the connectionist models (for other connectionist models see for instance Dell, 1988; McClelland & Elman, 1986; Stemberger, 1985), offers a different perspective to that of Levelt and co-workers on the issue of phonological retrieval effects on speech production. In this theory, processing units known as nodes are used to represent linguistic entities such as words and syllables. There are two interacting networks of nodes: a content network (content is used here in the sense of a semantic process rather than as a property of a sub-class of words) and a sequence network. Content nodes represent units such as syllables, syllable onsets, rhymes, vowel nuclei, codas, and individual segments or speech sounds. They are called content nodes because they represent the content of what is perceived or produced. They are organised hierarchically. The nodes in the sequence network control the sequencing of behaviour by imposing serial order on the activity of content nodes. Santiago, MacKay, Palma and Rho (2000) recently provided data in support of this model and that is difficult to explain by Levelt et al's (1999) theory. The two models make different predictions when it comes to onset cluster complexity. Levelt et al's (1999) theory assumes that phonological units are retrieved in parallel and associated in a sequential left-to-right manner to the word's structural frame. After segment-to-frame association, phonetic syllable programs are accessed and stored in an output buffer as a linear string. Articulation can only begin when the buffer contains one or more phonological words. Consequently the more segments a word contains the more time is needed before the word can be produced. This means, however, that how segments are structured (i.e. complexity of segments) in words of identical length should have no effects on picture naming latencies. To spell it out more clearly if an entire PW is buffered before output (as in Levelt et al's, 1999 theory) words with identical length should be pronounced equally fast regardless of their phonological structure. Santiago et al (2000) provided evidence that both consonant clusters in syllable onset position and higher number of syllables delayed the initiation of the naming response. They conclude that their results fit NST but pose problems for models that do not contain syllable and consonant cluster units, that do not propose a hierarchical organisation within the syllable, or assume that an entire phonological word must be buffered before articulation can start. Moreover, they interpret these results as evidence for the psychological reality of syllable nodes that represent fundamental phonological units.

The two level distinction (common to Levelt et al., 1999 and MacKay's, 1981 theories) is also embodied in a large number of subsequent connectionist models of speech production (Dell, 1986; 1988; Harley & MacAndrews, 1995; Hartley & Houghton, 1996). For instance, in Dell's (1986) model activation flows from lemma to lexeme, but it can also flow in the reverse direction (i.e. the same functional architecture as Levelt et al's (1999) model, but using an interactive network). In this connectionist model syntactic and phonological frames are generated. The syntactic frame specifies the noun and verb-phrase structure of the sentence, i.e. a grammatically labelled slot for each word. Words are then retrieved from the mental lexicon and entered into the appropriate slots. A similar process takes place on the phonological level. Again frames are generated that specify the syllabic structure for the word, and phonological segments are subsequently inserted into the slots in the phonological frame. It is then assumed that the articulatory program for speech can be 'read off' from left to right across the phonological slots (see Garrett, 1980).

Speech theories originating in the field of phonology have questioned whether dual-stage processing is necessary. These theories argue that speech output is organised in terms of coordinative structures that internally organise all actions including those identified as performed in separate stages in other models. Central to these is the concept of the articulatory gesture that defines a motor sequence and is therefore the basis of speech control (Browman & Goldstein, 1986; 1989; 1992; Fowler, 1986; 1995; Kroger, 1993; Lofqvist, 1990; Saltzman & Munhall, 1989). These have been described as action theories of language formulation and speech production (Strand, 1992). According to Browman and Goldstein (1997) gestures are abstract, combinatorial units of phonological contrast (that means they are distinctive), that are defined with respect to their physical properties (co-ordinating multiple articulators and muscles), i.e. completely different to what is usually referred to as a hand/facial gestures in language research. They emphasise that speech can be decomposed into these gestures that are simultaneously units of information and units of action (available for perception as well as production). Browman & Goldstein (1992) highlight research showing that these articulatory gestures vary in a constrained quantitative and gradient way according to wider prosodic variables such as position and linguistic stress. This latter point has particular relevance in light of positional and stress factors associated with stuttering (see section 0).

1.10 Bilingual language acquisition

The principle reason to study bilingual children in this thesis is to study processes that have been hypothesised to cause problems for children who stutter at the age of onset, for instance one aim is to establish certain syntactic factors that might operated differently in bilinguals as compared to monolinguals (and are known to cause older children who stutter difficulty – e.g. long content words). To aid this a brief introduction to bilingual language processing is now given. De Houwer (1995) defined bilingual language acquisition as referring to 'the result of the very early simultaneous, regular, and continued exposure to more than one language'. In the bilingual research literature it is emphasised that a differentiation should be made between bilingual first (simultaneous) and bilingual second (sequential) language acquisition (BFLA and

BSLA, terms introduced by Meisel, 1989). Researchers using these terms only refer to bilingual first language acquisition when two or more languages are introduced from birth or, at most, a month after birth. De Houwer (1995) highlighted the theoretical importance of this distinction and urged researchers in the field to always specify exactly when their bilingual subjects were first regularly exposed to more than one language. This is also the approach taken in the latter part of this thesis.

1.10.1 Processing issues in bilingual language acquisition and the relationship to research into stuttering

The issue of whether bilingual children's syntactic performance shows a delay compared to their monolingual counterparts has been a controversial issue in the bilingual research literature. Differences in syntactic and lexical acquisition have been used as evidence on both sides of the unitary / separate language debate. Both Genesee (1989) and Meisel (1989) have pointed out that for instance code switching, i.e. switching between the languages, as well as morphological and syntactic mixing is frequent in very young bilingual infants (before age 2;0) but becomes then less common. This has been taken as evidence by researchers who advocate that it is generally one overall language processing system with two subcomponents and that the child will only later in development notice the linguistic differences (see for instance the three-stage model by Volterra & Taeschner, 1978). In the other camp there are those who argue that children differentiate the two languages from birth and that early mixed utterances can be attributed to limited lexical resources rather than a single initial system (e.g. Deuchar & Quay, 2000; Genesee, 1989) . Bialystok (in her recent book bilingualism in development, 2001) reconciles both these positions by drawing on Green's (1998) inhibitory control model. She speculates that it is the processing (i.e. how the languages are activated) rather than the representation (i.e. how the two languages are stored) that is important. In the inhibitory control model (Green, 1998) it is theorised that both languages are always active but that one is inhibited when the other is needed. It is these inhibitory processes that Bialystok (2001) believes are developing in early bilingual children. One point, raised by researchers favouring the separate language hypothesis, is that differences in performance are due to language internal factors and not to language external factors, such as language dominance (Hulk & Müller, 2000). In respect of the syntactic development, Hulk (2000) suggested that the acquisition of grammatical rules for each language may be a much harder task for bilingual as compared to monolingual children. Other researchers conclude that there is absolutely no delay in the acquisition of language in bilingual children (Granda-Rodriguez, 1998). Some researchers even theorise that bilingual children may acquire some syntactical features faster and with fewer errors than many or most of their monolingual counterparts (Meisel, 1990).

Code switching has been taken in the early bilingualism literature as further evidence for the unitary system theory in the three stage model (Volterra & Taeschner, 1978). The content and function word issue that has become so central in the stuttering literature is also apparent in the code switching in early bilingual language acquisition. It has been suggested that closed class elements play a special role in the code switching between languages, meaning that there is a tendency to preserve functional elements in the dominant language when switching to the second language (Muysken, 2000). For bilingual children this has previously been observed in the case of an English – Norwegian two-year-old (Lanza, 1992). The authors specifically highlighted a large proportion of mixing of functional elements when the child used grammatical (i.e. closed class) items from the dominant language and incorporated them into the weaker language. Deucher (1999) analysed the context of mixed utterances with function words further. She hypothesised from her results that function words are not treated as language specific by young bilingual infants, whereas content words are. This again affirms the differentiation between the content (i.e. the vocabulary increase by children) and the special status of the functional elements (which are learned at the same time during early syntax acquisition). For children who are brought up in bilingual environments – it is estimated that nearly half of the world's population is functionally bilingual (as pointed out by de Houwer, 1995) – the task of syntax acquisition is doubled for the two languages acquired. If, as Bernstein Ratner (1997) suggested, the interaction of the increases in vocabulary and syntax could lead to retrieval problems, this should be even more likely in bilingual children since they concurrently have to learn two different lexicons and grammars. Bernstein Ratner (1997) links the combined effect of these two factors to the onset of stuttering in childhood. As mentioned earlier in this chapter (section 1.8.1), some speech therapists have even taken the issue of bilingualism further and have warned that mixed utterances by parents can contribute to the development of a stutter in children with a predisposition to this disorder (Lebrun & Paradis, 1984). With respect to lexical speech errors, Wijnen (1990; 1992) reported results showing that increases of repetition and substitution errors coincided with the time of the main 'vocabulary spurt'. The time period between the age of two and three years was also indicated as the time when a child would have a relatively sudden preoccupation with closed class words and there is a general increase in the production of this word class (Elbers & Wijnen, 1992). This second aspect of grammatical class indicates that children develop not just their vocabulary during this stage, but also that syntax acquisition takes place concurrently. In research into stuttering grammatical class and its development is particularly important, since the closed class words are those that are produced proportionally more disfluently by children at the age of stuttering onset (e.g. Bloodstein & Grossman, 1981; Howell, Au-Yeung, & Sackin, 1999).

To sum up there are several reasons for studying bilinguals in fluency development. Firstly there are parallels between bilingualism and stuttering. A second point is that bilingual work can provide insights into the processes that cause difficulties for children who stutter. Another issue is that stuttering research could provide some understanding of bilingual language processing (however this last point is not the intent of this thesis).

1.11 Speech Production Models of Fluency Failure

Outlined in the following two sections are models of fluency failure that have been applied to stuttering and that both highlight the importance of speech planning aspects. Both of these theories are influenced by fluent speakers' speech errors but also have specific views about the processes leading to disfluencies, such as the repetitions and hesitations in stuttering. The first, the Covert Repair Hypothesis, owes its approach initially to Levelt's work. The second (Howell's EXPLAN theory) emphasizes the importance of how long representations take to prepare - linking it with connectionist models (such as MacKay's node structure theory) and articulatory phonology.

1.11.1 The Covert Repair Hypothesis (CRH, Kolk & Postma, 1997; Postma & Kolk, 1993)

CRH explains phenomena such as those observed in stuttering (retraces and hesitations, for instance) by drawing parallels between the events seen when speakers make a speech error which they then repair. It draws heavily on Levelt's (1989) perceptual loop theory and focuses particularly on feedback monitoring. The perceptual loop theory is a previous version of the model described in section 1.9.1 which relied more centrally on feedback monitoring mechanisms in speech production.

Postma (2000) gave an example of the speech error 'left of purple is – uh – of white is purple' which is a case of an overt repair. This lexical speech error highlights three issues of self-repair (after Levelt 1983). There is the inconsistency that needs a repair, which is called the reparandum. Then there is the so-called editing phase that is the interruption followed by a delay. Finally there is the repair proper. In disfluencies such as in stuttering, i.e. an interruption to the progress of speech, both the error and the repair are assumed to be absent. This is why these cases are seen as covert repairs where corrections are made to anticipated errors. These assumed errors are supposed to be detected by the 'inner loop' which is a pre-articulatory monitor. According to this theory the inner loop gives speakers the opportunity to detect an error before it has become an overt response. With respect to stuttering Kolk and Postma (1997) speculate that people who stutter have a slow phonological system. To explain how this can lead to speech errors they use the spreading activation model of lexical selection (Dell & O'Seaghdha, 1991). According to this theory when looking for a specific lexical item phonologically related, competing words are also activated. The competing items have similar levels of activation early on in the search but their activation levels off earlier.

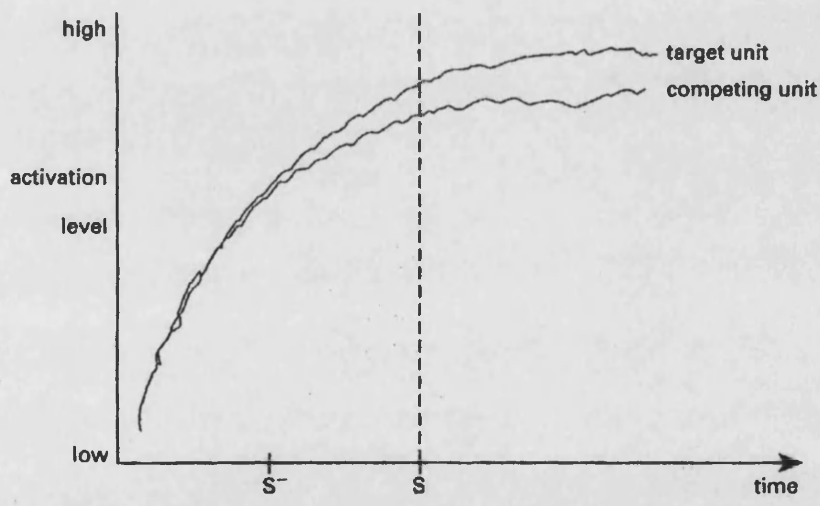


Figure 2: Spreading activation for a target and competing word used in the Covert Repair Hypothesis (Kolk & Postma, 1997). This diagram indicates normal activation rate with two selection points: normal (S) and early (S-).

This means that later in the selection process only the target item has the highest level of activation and is thus selected. A speaker with a slow phonological system would have to select a word when the activation is still being built up and the target item as well as the competitors have similar levels of activation. This could lead to a competitor item having momentarily a higher level of activation which could trigger selection, i.e. start a speech error. This wrong selection is then detected by the inner loop and a stutter is then a covert response to this erroneous selection. This would also apply to fluent speakers speaking in conditions of time pressure.

1.11.2 The EXPLAN Model of Fluency Failure (Howell & Au-Yeung, 2001; Howell & Au-Yeung, 2002)

Howell and Au-Yeung (2002) formalised and extended Clark and Clark's (1977) suggestion that word repetitions occur because the plan for a later word is not ready for execution. These word repetitions are most frequent with pronouns and conjunctions (Clark and Clark, 1977), which belong to the class of function words as described earlier. Clark and Clark (1977) suggested that pronouns and conjunctions (types of function words) are repeated because the plan for a later word is not ready for execution. It should be emphasised that, compared to CRH, it does not rely on an error monitoring mechanism to explain these phenomena. One of the underlying assumptions of the EXPLAN model is that planning and execution are independent processes that take place in parallel. The authors focused on cases where a word, that was already produced correctly, is repeated, or where the correct initial phoneme or phonemes are spoken (such as blocks and prolongations) but the rest of the word is not pronounced. They highlight that none of the sounds produced in these cases are in error, but rather they constitute instances where the forward flow of speech is interrupted. One of the components of the model is a planning unit (PLAN). As to the question whether there are one or two stages involved in the planning process (the lemma – lexeme debate for more detail on this see section 1.13), the EXPLAN model remains neutral. It does, however, use the notion that the output of the planning stage is filled in, phoneme by phoneme, in left to right order (cf MacKay's NST). This means that even if the planning of a word is not fully completed, the first phoneme may be available for execution. The execution unit (EX) produces the output of a word or as much of the phonetic representation as is currently available. This unit is assumed to be independent of the planning process. This independence of EX and PLAN allows the two processes to take place in parallel, which would allow planning to carry on without stopping the execution of previous words. A further component of their model is a general purpose timekeeper (GPT) whose principal role is to regulate speech rate.

In relation to the fluency failure model, Howell and Au-Yeung (2002) highlighted that the time at which a previously planned word is executed will determine when the plan for the next word (or first part of the word) is required. This would mean that execution time of previous words would put pressure onto the planning process by accelerating when the next plan is needed. In the case of a fluency failure, the problematic point is the relationship between the word that has just been executed and the plan of the next word. This can be problematic when the forward planning for the next word is not complete. This can lead to fluency failure when a speaker has finished executing the first word, but the plan for the next word is not ready. This is the particular case Howell and Au-Yeung referred to in their theory, a so-called EXPLAN state of fluency failure. It is noteworthy that in their theory fluency failure is not viewed as an error in the production system, but rather a case where planning is not ahead of execution any more (the two processes have run out of synchrony). This is where the function/content word distinction becomes significant again. Planning can be very rapid when words comprise a well-learned small set of everyday words (see Sternberg, Knoll, Monsell, & Wright, 1988) or, as regarded here, simple words (function) in comparison with complex ones (content). Note that although this thesis concentrates on planning processes from syntactic levels down. All levels (such as those involved in generating syntactic forms) can affect fluency if planning processes require a long time and lead to the plan not being ready.

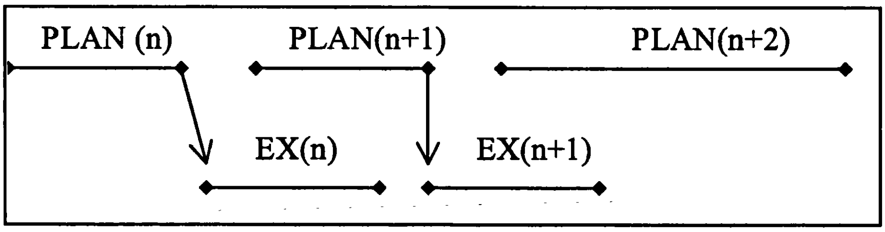


Figure 3: Representation of the EXPLAN model of fluency failure. The horizontal axis indicates time. Rows represent time for planning (top) and execution processes (second row) – see text.

Figure 3 gives a schematic representation of the EXPLAN model and highlights the relationship between the planning and execution in fluent speech. The horizontal axis represents time and the different bars along the top represent words (...n, n +1, n+2...) and the times it takes to plan each word. The bars in the next row represent the times and extents over which words are executed. Speech is proceeding fluently in the figure, because the planning for word n is ready by the time that execution of word n-1 has finished.

There are a number of reasons why the plan for the next word might not be ready in time, i.e. once the previous word has been executed. Howell and Au-Yeung outline four particular circumstances when this is most likely to happen. It can occur when there is little prior planning, with a function word (short, fewer initial consonants, unstressed), when the first word is executed rapidly, and when a word is followed by another word in need of a longer planning process (either at the lemma or phonemic level). In these cases one of the options available to a speaker is the repetition of the word that has just been executed – this would be an example of a so-called stalling disfluency as described at the outset. In this way the planning for the next word can continue until a complete plan is available. The focus on timing aspects of planning and execution (inherent in EXPLAN) is also consistent with a recent study investigating the neuronal basis of developmental stuttering (Sommer, Koch, Paulus, Weiller, & Büchel, 2002). Their study involved a comparison between speakers with persistent developmental stuttering and a control group using diffusion tensor imaging. One of their conclusions was that persistent developmental stuttering results from disturbed timing of activation in speech-relevant brain areas.

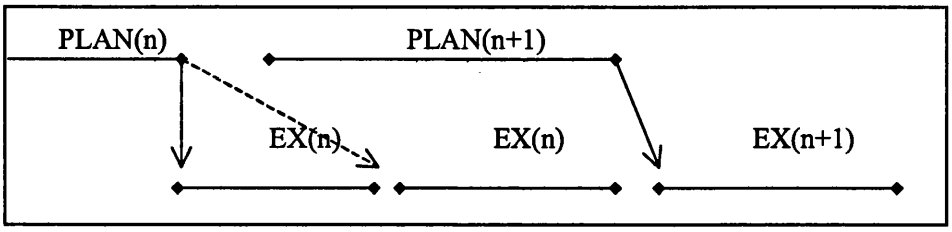


Figure 4: Schematic representation of fluency breakdown within the EXPLAN model. By the time word n is executed the plan is not ready for word n+1. In the figure the previous word is repeated and by the time the repeated word has been executed the plan for the next word is ready (an example of a stalling or delaying disfluency).

One of the main features of the model is that fluency failure on function words is viewed as a process that delays the execution of the content word, so it can be produced fluently. As a basis for this theory a method was developed (Howell, Au-Yeung, & Pilgrim, 1999) that would make it possible to divide speech in a way which would determine the position of a function word relative to adjacent content words. This method using the linguistic concept of the phonological word is outlined in a later chapter.

1.11.3 <u>Cross-linguistic / bilingual findings in relation to the two models of fluency failure</u>

It was described in section 1.8.1 that the predominant pattern for bilingual people who stutter (PWS) is either the person stutters in both languages or stutters just in the dominant language. A pattern whereby a higher frequency of stuttering occurs in the dominant language than the less proficient language is difficult to explain with the CRH (Postma & Kolk, 1993, Kolk & Postma, 1997). Recall Figure 2 and the fact that in CRH responses positioned early lead to a high proportion of disfluencies. How could this predict stuttering only in the dominant language? Would activation rate of words in L1 be expected to be slower than L2? There seems to be no reason (if anything would seem to be the opposite), that the phonetic system should be slower in L2 that leads to activation build-up. Another possibility is slower monitoring – so would there be different monitors in L1 and L2? Or is it even possible to position the threshold in different places for L1 and L2? However, Kolk and Postma's (1997) own description implies that thresholds (for the activation of a target as compared to its competitors – see Figure 2 are imposed at a fixed time across PWS and fluent speakers. This seems to indicate that there is no reason to suppose that the same speaker (with two languages) should have the threshold positioned at different points in time.

In the EXPLAN (Howell & Au-Yeung, 2002) model however, certain predictions can be made about disfluency patterns of bilingual speakers that appear consistent with previous findings. An analogy is drawn between these speakers and children acquiring a language. Recall that young children show more stalling (hesitations, and whole word repetitions) disfluency. This would be the pattern predicted for a bilingual in their second / or less dominant language. These speakers should show more advancing (prolongations, part-word repetitions and blocks) disfluencies in their dominant language (evidence of this was reported for a Spanish - English bilingual, Howell et al., 2003). A balanced adult bilingual is assumed to show advancing types of disfluencies in both languages. Structural differences between languages could then be used to predict the loci of disfluencies, such as the positions of reflexives and adjectives in Spanish and English, as described above. These could be seen as disfluencies where an upcoming difficulty is anticipated. This would then also have implications for possible treatments of bilingual speakers. Therapists ought to monitor the distribution and types of disfluencies in each language since the language that shows more advancing disfluencies is considered to be the problematic language.

Although the current investigation does not look at bilingual speakers who stutter it can pinpoint structural and distributional properties of other languages, in this case German, that could affect the foci of difficulty (as outlined in the following sections). The language development of fluent bilingual English – German children are studied in the latter part of the thesis to investigate aspects that might influence fluency development. This can highlight linguistic difficulties at the time of stuttering onset in children (the age between 2 and 3 years old - according to Andrews & Harris, 1964). The work with fluent speakers ties in with Bloodstein's (1987, 1995) continuity hypothesis, as described earlier.

The next section summarised parts of the earlier described different approaches to speech production research into stuttering and clarifies the terminology used in this thesis.

1.12 Terminology

The introduction provided so far an overview of the relevanttheories and approaches that have been taken in the linguistic research of stuttering. This current section takes a step back to examine the terminology that have been used in these various approaches and highlight the particular stance taken in this thesis. Since the relevant theories needed to be introduced to have an understanding of these issues it is provided here even though some of this terminology has been used in prior sections.

Wingate (2002, see particularly chapter seven) discussed the terminology used to refer to different stuttering events and presented some terms of his own. While he argues strongly that his terms are the appropriate ones, the matter is far from settled (see correspondences between Dayalu, Kalinowski, & Stuart, 2003; and Wingate, 2003; and an earlier comment by Yairi, Watkins, Ambrose, & Paden, 2001). Table 1 below gives a brief overview of the terms used, and subdivisions made, by different research groups.

Term	*Main Protagonist*	*Interpretation*	*Implication*
Dysfluency (rather than disfluency)	Wingate (1988, 2002)	Stuttering events are prolongations, blocks, monosyllabic whole- and part-word repetitions minus monosyllabic whole word repetitions	The predominant pattern of childhood 'stuttering' is not considered to be stuttering at all, since it involves function word repetitions
Stutter-Like Disfluencies (SLDs)	Yairi and co-workers (1992)	Part-word repetitions, monosyllabic word repetitions and dysrhythmic phonation	Based on their work with children they do interpret monosyllabic word repetitions as stuttering events and draw on similarities with normal speakers' speech errors.
Within-word vs. between word disfluencies	Conture (1990)	These two categories are considered to be subclasses of overall stuttering events. *Between word* refers to the type of errors normal speakers would make, whereas *within word* refers to part-word repetitions and blocks.	Central to these terms is the idea of the continuity hypothesis, i.e. the link between normal speakers' speech errors and stuttering events.
Stalling and advancing disfluencies	Howell and colleagues (2002)	Similarly to Conture, both of these terms are used to refer to subclasses of stuttering events: Stalling would refer to monosyllabic whole word repetitions, whereas prolongations, part-word repetitions and blocks would be classified as advancing disfluencies.	Draws on similarities between normal speakers' speech errors and highlights particularly the childhood pattern. However, stalling type disfluencies are not quite as expansive as in Conture's *between word* disfluencies.
Reperandum, Editing Phase and Repair	Kolk and Postma (1997)	Revisions, pauses (filled and unfilled), part- and whole-word repetitions, blocks and interjections.	Stuttering events are interpreted as errors that are detected before output

Table 1: Classification of stuttering events used by different research groups.

Definition of all disfluencies. Here, the class of all disfluencies is defined first (these basically correspond to Johnson's, 1961, eight features as described in a previous section - 1.3), which is then followed by a description of the way they have been divided up by different authors. Since usage here differs from Wingate's (and, indeed, from that of other authors) it is essential to be explicit about the use of terms in this thesis. In the thesis the term disfluencies is used interchangeably with stuttering to refer to the events of prolongations, blocks, monosyllabic whole- and part-word repetitions.

Taxonomy. The contentious issue involves grouping these events into classes. Wingate (2002) for example, groups prolongations, part-word repetitions and blocks on words as "stutterings" and classifies the remainder as "disfluency". Other classifications (Yairi, 1992; Conture, 1990; Howell, 2002; Kolk & Postma, 1997) are given in Table 1.

Taxonomy and Process. Whole word repetition is common in children diagnosed by clinicians as stuttering and such repetitions occur mainly on function words (Howell, Au-Yeung & Sackin, 1999). Older speakers who stutter have problems on the first part of content words, which is the true sign of stuttering according to Wingate (Howell et al., 1999). The EXPLAN (Howell & Au-Yeung, 2002, see section 1.11.2) theory maintains that both these are indications of an underlying problem in the vicinity of function and content words that are grouped together phonologically. In an utterance like "I stroked it" for example (i.e. a content word surrounded by function words), the problem is considered to be the time taken to generate the str- cluster. More time is needed to generate the rest of the plan ("-oked" in this case), and stuttering does not arise out of some error-prone planning process. According to EXPLAN, children who repeat function words are dealing with the unavailability of the last part of the plan by repetition and hesitation that stalls the time of onset of the following word, allowing more time for the generation of the last part. Adults who stutter advance from the function word to the first part of "stroked", trusting that the remaining part will be generated in the time taken for execution of the first part. If this fails, stuttering involving onsets on content words ensues. Howell et al. (1999) have shown that the two ways of dealing with this problem are reciprocally related over age groups. Thus, young children who stutter show a high proportion of word repetition on function words but a low rate on content words, whereas adults who stutter show the opposite. This exchange between stuttering event types would not be captured if Wingate's way of identifying stutterings were adopted. It should be noted that it also gives a different interpretation to the role of pauses (Howell & Au-Yeung, 2002). It suggests that the notions behind Conture's (1990) ideas about between- and within-stutterings (associated with function and content words, respectively) and Yairi & Ambrose's (1992) notion of SLDs that emphasize the importance of context, may provide important information about the way stuttering develops. Wingate (2002) denies such demonstrated development of the disorder over ages (to restate his emphatic statement on p.372 "Stuttering does not 'develop'"). The theory also has treatment implications that are not just fitted to the data with the benefit of hindsight, but that have been tested (Howell et al., 2001). Two terms used in the framework employed in this thesis (the EXPLAN model) should be highlighted here. Stalling refers to a process and the events included within this class are whole-word repetitions and pauses. Advancing is a mechanism that predicts certain disfluencies are the result of the process of attempting the execution of a word too soon. In the CRH (Kolk & Postma, 1997) stalling would be considered to be a covert repair process and refers to whole word repetitions and pauses. Therefore these are always assumed to be the result of an underlying error. Disfluencies on word fragments could then be regarded as fragments of overt errors.

Particular comment needs to be made about Kolk and Postma's (1997) position: As described in section 1.11.1 their view derives from Levelt's view on breakdowns in fluent speech control. He considers stutterings are reflections of errors made in speech production that are detected before output. So, in an example like "turn left, er, turn right", there is an error, and an example like "turn, er right" might also be a response to the same error detected before speech output (see section 1.11.1 for details). The point to note about the second example is that word repetition and pauses (filled and unfilled) can all occur which means that all these events in their view represent stutterings.

Other terms used in research into stuttering are *normal nonfluency* which refers to the events that would be classified as fluent speakers' speech errors, such as pauses, interjections, whole word repetitions, certain types of sound exchanges and revisions. Furthermore, *fluency failure* and *fluency breakdown* are wider, more inclusive terms referring to both normal nonfluencies and stuttering events.

Figure 5 gives a breakdown of positions in terms of disfluency type, classification, word type, process and what age groups are addressed.

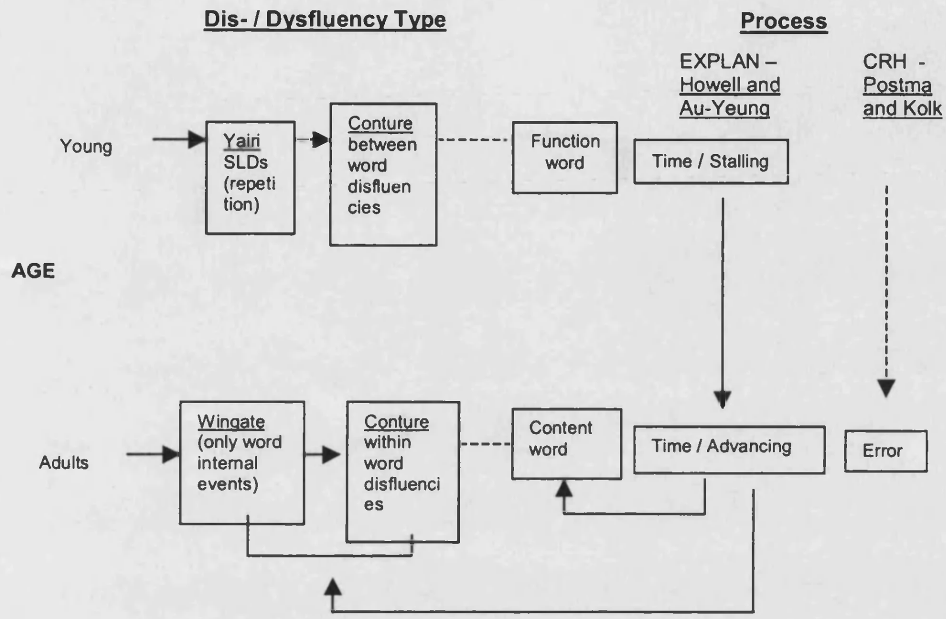


Figure 5: Relationship of positions with respect to disfluency type, classification, word type, process and what age groups are addressed.

Figure 5 highlights the fact that some researchers focus only on one particular age group and their particular pattern of disfluencies (such as Yairi, 1992, on children and Wingate, 1988, 2002, on adults respectively). Conture's (1990) within and between word disfluencies could be seen to reflect the age dependent pattern of function and content word disfluencies. However, he does not describe any processes that might lead to these different types. The two models of fluency failure on the right side of the figure describe processes leading to the two different types of disfluencies (stalling and advancing in EXPLAN and monitoring in CRH). The focus of CRH however, is on error whereas the central point to EXPLAN is timing. In the next section lexical access is related to models of fluency breakdown.

1.13 Stuttering - Difficulties in Lexical Access?

1.13.1 Error and TOT Data as Evidence for the Assessment of Theories

The main finding from speech error research is that two types of error occur that reflect the two processes (lemma and phonetic levels). Whole word exchange errors take place at the lemma or syntactic level, and speech errors that involve phonemic or syllabic breakdowns and transmutations take place at the phonemic level (Dell, 1989; Garrett 1982) – such as for instance the lemma level error of saying *left* for *right* (wrong but semantically closely related lemma), or *expedition* for *exhibition* (wrong but phonetically closely related), or *arrent curgument* instead of *current argument* (Spoonerism, i.e. exchange of onsets). These types of speech error are used in support of models in which the two levels operate discretely and are serially organised (as in the Levelt et al 1999 model) and where the two levels interact over recurrent neural networks (such as in the Dell 1986 model). The vast literature on speech errors is not reviewed in detail here (interested readers may refer to Fromkin, 1971; Garrett, 1982; 1993).

The notion of two types of speech error has been looked at experimentally using techniques that probe the tip-of-the-tongue (TOT) phenomenon. The work on TOTs is reviewed in some detail below as it is pertinent to the assessment of the theories of speech control. As such it is employed in a study with fluent children in two age groups (since it is relevant in the EXPLAN framework, i.e. as a state of an acute unavailability of plan). Parallels between the tip-of-the-tongue phenomenon and stuttering are also drawn.

1.13.2 TOTs in Relation to Models of Fluency Breakdown

Many of the studies that assess predictions derived from EXPLAN have been based on low-level (phonetic, phonological) analyses of words, or gross characterizations of words (e.g. into function and word classes). Though there has been this focus to date, higher level factors still affect the EX-PLAN processes and can affect stuttering. Thus, material that is syntactically difficult that takes a long time to prepare or words that are difficult to retrieve and, therefore, take extra time, should affect stuttering. One way of looking at lexical retrieval (already mentioned) is TOT. TOTs are relevant in the planning and execution framework of stuttering, due to the fact that they could be considered an acute form of the lack of the availability of a plan. The findings of TOT studies also link in to stuttering research in the way they show that participants who are in the TOT state know certain aspects about the beginnings of the words they are looking for (such as the first sounds etc). This can again be likened to a situation where a person who stutters repeats or prolongs the first phoneme. Again this would point to the fact that the phonemic level is being built up left to right (i.e. knowledge of the first sounds of words in TOTs and exchanges of first word parts in spoonerisms).

1.13.3 Lemma Information in TOT States

Experimental studies of TOTs were pioneered by Brown and McNeill (1966). Their methodology (and variations of it) has been used since in many studies to induce TOT states. In their original paradigm, participants were provided with a definition (e.g. 'The highest point in the sky directly above the observers head') of a rare word and then asked whether they could recall it. When subjects were unable to retrieve the word, they were asked whether they were in a TOT state. A TOT state was described to participants as being unable to say the word, being sure that they know it combined with a feeling that the word is about to "come back" to them. When subjects indicated that they were in a TOT state they were asked to provide information about the unretrieved word (such as any other words that came to mind, number of syllables, initial letter of the word etc.). Before the next definition, the target word ('Zenith' for the above example) was given and subjects, who had indicated that they were in a TOT state, were asked to indicate whether it was the word they had in mind (which they labelled *positive TOTs)* or a different word (a case referred to as a *negative TOT*). The negative TOT states can then be used as a baseline to compare positive TOTs against. Results indicating partial syntactic knowledge by the speakers in TOT states (see for instance Vigliocco, Antonini, & Garrett, 1997; Vigliocco, Vinson, Martin, & Garrett, 1999) initiated a lengthy dispute about the number of stages involved in retrieval (the *lemma – lexeme debate*). On the one end of the spectrum the main contenders in this disagreement are Caramazza and co-workers (Caramazza, 1997; 1998; Caramazza & Miozzo, 1997) who proposed that a level of lexical (lemma) representation is not needed between lexical and phonological retrieval. On the other end of the spectrum researchers argue that TOT findings suggest a unit connected to syntactic nodes (i.e. a node between abstract lexical information and representation of the phonological word form) specifying attributes, such as for example grammatical class, gender, auxiliary type and count / mass (e.g. Bock & Levelt, 1994; Jescheniak & Levelt, 1994; Roelofs, 1992; Roelofs, Meyer, & Levelt, 1998). Here the main dispute can be summed up as concerning differences between strictly hierarchical or parallel retrieval processes. For the purpose of the thesis, this debate is not outlined in greater detail (please see the above references for a detailed discussion) since it was introduced to highlight the fact that both partial phonological and syntactic (such as grammatical gender which was probed in chapter eight) information can be accessed by speakers in a TOT state.

1.13.4 Lexical Retrieval in Children

There have only been four reported studies of TOT states in children (Butterfield, Nelson, & Peck, 1988; Elbers, 1985; Faust, Dimitrovsky, & Davidi, 1997; Wellman, 1977). In these studies two kinds of measures need to be distinguished. One of these is the feeling of knowing (FOK). In FOK the speaker is aware that the word is in his/her mental lexicon, whereas in a TOT state the speaker has a strong feeling that the word's retrieval is imminent. Discussing the phenomenology of TOT states compared to FOK, Schwartz and colleagues (Schwartz, Travis, Castro, & Smith, 2000; Schwartz, 2002) emphasised that the two need to be separated. Schwartz (2002) points out that the two differ in respect of predicted outcome, i.e. a FOK state predicts recognition of a word (often used in studies concerned about meta-cognitive abilities), whereas TOTs are predictions of recall and according to Brown (1991) occur involuntarily which means that TOTs occur at a later stage in lexical retrieval. It then follows that only TOTs contain phonological information.

Of the four studies investigating children, two analyse FOK (Butterfield et al., 1988; Wellman, 1977) rather than the TOT phenomenon, and the study by Elbers (1985) is mainly anecdotal in nature. Wellman (1977), in the earliest of the four studies, aimed to investigate metacognition in children in two age groups (six and eight year olds). The focus was on an examination of recognition rates of FOKs as a measure of accuracy, rather than phonological or syntactic information available in these states. Recognition rates were higher than chance level, and moreover older children were significantly more accurate than younger children. Butterfield et al. (1988) also investigated developmental aspects of metacognitive accuracy (as measured by FOK recognition) in six, ten, eighteen and seventy year old subjects. In contrast to Wellman (1977) they reported that FOK accuracy decreased after age six, i.e. children between six and ten did not improve in their metacognitive abilities as suggested by Wellman (1977).

Elbers (1985) examined the TOT phenomenon, by recording conversations with her two-and-a-half year old son. She reported several occasions where her son used similar sounding words to the ones he could not retrieve. These events were interpreted as suggesting that children younger than three years of age can be in a TOT state and more specifically that partial phonological information such as syllable number and stress pattern is already available at such a young age.

The most recent study with children (Faust et al., 1997) investigated TOT rates and accuracy in normal and language impaired seven to eight year old children. They used pictures of animals and objects and asked the children to name them. Language-impaired children indicated a higher rate of TOT instances but percentage of resolved TOTs was much lower compared to the control children. They also recognised fewer of their TOTs and recalled more incorrect phonological information. An investigation of TOT in fluent speakers can help clarify whether any of Brown's four factors make lexical access more difficult. For instance whether words starting with a consonant would be more likely to be on the tip of the tongue than words starting with a vowel (when other attributes are controlled for – such as word frequency, length etc). Thus the factors that are known to affect disfluency can be manipulated in the target words and thus it can be tested whether these words are also more difficult to retrieve lexically.

The investigation of TOTs in children can clarify whether their planning takes place in a similar manner to that of adults. Their knowledge of the words that are on the tip of their tongues might also highlight planning differences across languages and age group. Despite some striking correspondences between aspects that make words more likely to be stuttered and to be in TOT states (such as words that are less frequent with fewer phonological neighbours and longer words - Harley & Brown, 1998), no studies have so far investigated TOT using a participant group of people who stutter. As such the work in chapter 8 can be seen as pilot work to assess TOT states in German speaking children in different age groups.

In the following section a rationale and overview is provided for the work in this thesis.

1.14 German and English comparison of fluency development and stuttering – areas of investigation

It was already highlighted in section 1.8 that cross-linguistic or comparative research is a good test case to determine whether specific linguistic factors precipitate stuttering regardless of the language they are spoken in (as pointed out by Bernstein Ratner & Benitez, 1985). The first part of the thesis analysed whether findings regarding linguistic predictors of stuttering that have been reported to operate in English also apply to German. It is outlined below where comparative linguistic differences for German might be expected to lead to differences in the stuttering patterns of the two languages. This is initially addressed in chapter two at the word level using Brown's (1938) four factors. The research question that is addressed is the extent to which the four factors operate in German. Chapter three stays at the word level by analysing phonetic characteristics in finer detail with the index of phonetic complexity (IPC, Jakielski, 1998). It was examined how phonetic factors operate in different age groups and whether the words that are stuttered are more complex than fluently-produced words. It was also analysed which specific factors have most impact on stuttering rates and whether these are the same in different age groups and across languages. Finally this was related to the frequency of use of the characteristics in spontaneous speech.

1.14.1 Exploration of Differences Between English/German Pertinent to the First Three Studies (chapters 2-4)

Although it can be stated that overall the two languages are close in their origin (both stem from the West-Germanic branch of the Indo-European language tree) and have many similarities, there are a number of differences that are important for studying fluency issues. Stress has been highlighted as a potential determinant of stuttering in general (Wingate, 1988, 2002). For instance the principle of weak forms, i.e. the existence of certain grammatical words which have a reduced form of pronunciation when unstressed in normal connected speech, is the same in English and German. There is, however, an important difference in their use in the two languages (see Hall, 1992).

As such it is the case in English that weak forms are entirely dictated by the stress and rhythm of the sentence and are completely unconnected with differences in style. This means that weak forms would be used in English even in formal speech.

In German the use of weak forms is highly dependent on the pronunciation style. Thus the frequency of weak forms and the degree of reduction varies greatly between a more formal style and relaxed conversational pronunciation. Generally in German formal pronunciation weak forms are less frequent and usually only have the first stage of vowel reduction (shortening), although if spoken very fast the second stage can also occur (vowel centralisation). Weak forms with centralisation and reduction of vowel to /ə/ are typical of conversational pronunciation. Almost exclusively in fast, relaxed, conversational style does the maximally reduced form (i.e. the elision of vowels) appear – see Wiese (1996).

Another important difference is that the process of syllabification of the words within a phonological word (PW, a concept which has been briefly introduced in Levelt et al.'s, 1999, theory and is covered in greater depth in chapter 4) would be slightly different in the two languages, and might affect stuttering. These are considered to be the smallest units of articulation.

There is an ongoing discussion in the literature concerning whether or not German permits 'word-internal' phonological words (see Booij, 1985; Hall, 1992; Raffelsiefen, 1999; Wiese, 1996; Yu, 1992). One of the issues that is connected to word internal PWs is the large number of compound words found in German. This also has the effect that there may be more than one PW within the same orthographic word, which generally is not the case in English. A more detailed background to the PW is given at the start of chapter four.

Another issue is that in German syllabification cannot cross over PW boundaries before affixes such as *+lich* (i.e. that means that according to Wiese, 1996, such affixes are phonological words in themselves whereas a similar affix *+ig* is not a phonological word – his own examples were '*täglich*' as compared to '*eklig*').

When considering Wingate's emphasis on the syllable initial position as the locus of the stutter (see Wingate, 1988, p. 207), the different syllable structures for English and German can also be compared and investigated. As such one of the main differences in syllable structure is the syllable onset. To illustrate this point more clearly a little more background information into syllable phonology is given in the next section.

1.14.2 Syllabic Theory as Applied to German

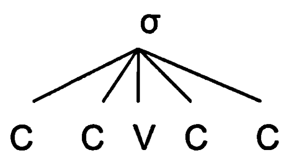
In contrast to the earliest forms of generative linear phonology, where the concept of the syllable played no role (such as in Chomsky & Halle, 1968 The Sound Pattern of English - SPE), syllabic theory has become a more central issue in phonetics and is considered a type of *non-linear phonology* (see Wiese, 1996, for a detailed discussion of this approach as applied to German phonology). The non-linear aspect refers to the view that syllables are hierarchical structures, which are usually represented by tree diagrams. This rejects the approach of SPE that tried to capture all the relevant phonological information about a word in a representation consisting of a string that formed a single line of structure. Spencer (1996) highlighted the importance of the concept of *phonotactic constraints* (which is covered in more detail in the background to the PW in chapter four) in syllabic theory. This term refers to rules that govern the possible sequences of sound, and many such constraints in a given language tend to apply at the level of syllable structure.

1.14.2.1 The Skeletal/CV Tier

As a basis leading up to a CV-tier representation of the syllable, Wiese highlights a major phonotactic constraint of Modern Standard German. This constraint refers to the fact that after a long vowel a certain number of consonants can occur. The same number of consonants is possible after a diphthong. After a short vowel, however, one more consonant is possible in the same monosyllabic words. Wiese simplified this constraint with the statement that a long vowel is equivalent to a diphthong while a short vowel leaves room for an additional consonant in the same syllable.

In Wiese's approach a syllable contains a certain, variable number of 'places' or 'positions' which can be filled by segments. Focusing on the examples given in (1), the phonotactic constraint given above can be expressed as follows: A position is labelled 'V' which denotes the place with which vowels are associated. Two more positions that follow the 'V' position would express the regularities outlined previously. These positions could be labelled 'C'. Furthermore Wiese proposed that there are also two such C-positions to the left of V in each syllable. These statements can be represented in a hierarchical tree diagram (2) in which the syllable node now dominates five C and V positions in a particular order.

(2)



This is the so-called skeletal tier or CV-tier. According to Wiese this is a template that represents the maximal syllables in Modern Standard German. The linguistically trained reader might wonder about cases such as the two German words *Spruch* (saying) and *Herbst* (autumn) for instance. These two cases seem to have more units than are available in the diagram. Wiese (1996) argues that affricates such as the

German 'sp' would take up just one of the C slots. The second example is an exception to the rule where the final affricate 'st' would be counted as an extrasyllabic segment. In other words 'h' fills just one C slot before the nucleus 'e' and 'r' fill the V and C units and then the only room is left for 'p'. Cases of extrasyllabicity are relatively rare and are not covered in detail here since they only occur at the end of syllables (and it is the onsets where the linguistic difference occurs – see next section) and as such are not central to the discussion that follows (the interested reader can refer to Wiese's 1996 section 3.2.3 which covers this issue). It should also be noted that whereas the 'V' would always stand for a vowel slot 'C' can also be taken up by a vowel. As such long vowels would take up both the V and the C slot which explains why these two slots make up the nucleus of the syllable – see next section.

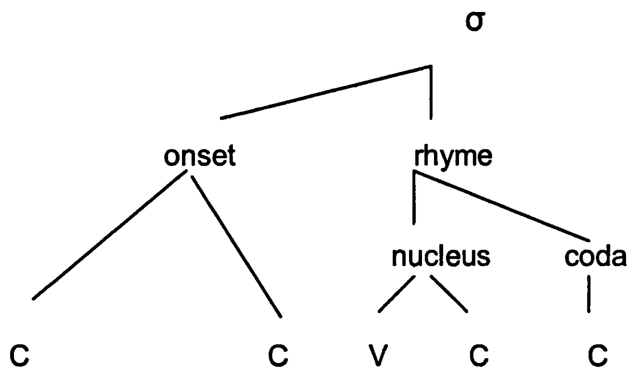
The phonotactic constraint, outlined above with respect to vowel length, is accounted for in the syllable template given in (2), because the CV-tier has only a limited number of places that can be filled. In order to represent the vowel length, a long vowel is assigned two segments of the skeleton.

When it comes to syllabic differences between English and German the so-called subsyllabic structure needs to be considered.

1.14.2.2 Subsyllabic Structure

A specification of the subsyllabic constituents is the domain of the prosodic hierarchy which is tree like. There are two constituents to a syllable. The syllable ending is referred to as the rhyme, due to the fact that this is the part that causes words to rhyme with each other. Starts of syllables are called onsets. The rhyme is generally further subdivided. Syllables have to have a sound that functions as the peak (which usually is a vowel) which is called the nucleus. The coda is the sequence of one or more final consonants. Below is an illustration of Wiese's model for German where the onset, the rhyme, the nucleus and the coda mediate between the CV-tier and the syllable node.

3)



All prevocalic consonants, together with all the phonotactic constraints governing the different combinatorial combinations of these, belong to the onset. This is where the main difference lies. In German the syllable onset is an obligatory constituent. As such syllables that start with a vowel (which would be the nucleus) have to have an added onset. In these cases a glottal stop is inserted as an onset to that syllable. This glottal stop insertion is also mediated by the stress pattern in German and the function/content word distinction. As such the glottal reflex is more frequent in content than in function words, partly because function words are mostly unaccented. With respect to sentence stress it can be stated that most accented vowels are marked by a glottal stop and these insertions are more frequent phrase-initially than phrase-medially (for all these points see Rogers, 2000). This is quite a different pattern to the case in English where glottal stops are only used as allophonic variants of certain plosives (such as /t/ in the Cockney accent as in *bottle* – [ˈbɒʔl̩]). These glottal stops, in English, would usually only appear word medially or finally. In some regional English accents glottal stops do appear in initial positions but this is on the whole less common than in German. In the case of stuttering this might also be significant because a number of researchers have observed laryngeal hyperactivity during and before stuttering (Bar, Singer, & Feldman, 1969; Freeman & Ushijima, 1975; Metz, Conture, & Colton, 1976; Shapiro, 1980; Thürmer, Thurmfart, & Kittel, 1983)

That means that vowel onsets can be considered to be stronger in German than in English. This might be relevant to Brown's finding with regard to words that start with vowels or consonants. Previous work has shown that both in Dutch and Afrikaans (languages where there is also more articulatory tension in initial vowels) more words with initial vowels were stuttered (and Uys, 1970 as cited in Bloodstein, 1995 respectively; Vaane & Janssen, 1978). It would therefore be hypothesised that this would also be the case for German.

This difference in syllable structure and laryngeal activity is an interesting issue to investigate further, due to the fact that it is also mediated by other factors that are known to affect stuttering (such as stress and content/function distinction).

Apart from these phonological differences there are also differences in the syntax of the two languages.

1.14.3 Syntactic Differences

Many learners of German consider the language to be fairly inflexible in its word order. This is mainly due to the fact that the finite verb and the past participle have fairly rigid positions in German sentences. This supposed inflexibility is mainly due to the fact that the finite verb and the past participle have rigid positions in German sentences. Comparing the two languages there are similarities as well as differences with regard to verb placement (see Döpke, 1998, for a detailed description). Identical surface structures appear predominantly in simple sentences. All finite verbs must be in second position in independent clauses (whereas they would appear in third position in English – example '*now it works*' '*jetzt geht es*'). Non-finite verbs in German are separated by verb complements, whereas in English (with the exception of adverbs such as *'just'*) they follow auxiliaries or modals immediately (example *'I will find it'* *'Ich werde es finden'*). In German dependent (subordinate) clauses, all verbs are in the final position. This means that verb phrases are head-final in German and head-initial in English. As explained by Döpke (1998) verbs in English are right-branching contrasting to left-branching in German. This means that English is a classic SVO structure whereas German is more similar to SOV languages. Verbs being rich in morphological encoding, it could be argued that there would be on average more planning involved for words in later sentence positions in German compared with English.

There is a much larger flexibility with regards to the position of the subject, direct object, and indirect object in German noun phrases. The reason for this is that in German the case endings always indicate how the constituents fit together syntactically. The following is an illustration of this point:

1. *Er hat seinem Vater eine CD zu Weihnachten geschenkt.*

   He gave his father a cd for Christmas.

   ➔ unmarked word order in both languages.

2. *Zu Weihnachten hat er seinem Vater eine CD geschenkt.*

   For Christmas, he gave his father a cd.

   ➔ unmarked word order in both languages.

3. *Eine CD hat er seinem Vater zu Weihnachten geschenkt.*

   *A cd he gave his father for Christmas

   ➔ word order that emphasises the accusative object in German.

4. *Seinem Vater hat er eine CD zu Weihnachten geschenkt.*

    *His father he gave a cd for Christmas.

    → word order that puts the emphasis on the dative object in German.

5. *Geschenkt hat er seinem Vater eine CD zu Weihnachten.*

    *Gave he his father a cd for Christmas.

    → word order that emphasises the past participle in German.

It can be seen that there are five different ways of ordering the same constituents in German (for a similar example please see Johnson, 1998). In English only the first two sentences are grammatically correct. Sentence 5, where a direct translation into English is not even possible, is particularly noteworthy. A verb cannot be foregrounded in English. Whereas German makes use of word order to stress different elements of a sentence, in English the same would be done using intonation. Thus many German learners experience difficulties not because of the rigidity of word order, but rather because of its very flexibility in that respect.

When taken together the past participle position and the greater flexibility seem to suggest that there would be more planning involved in later sentence positions in German when compared to English. Thus it would be predicted that earlier sentence positions (one of Brown's, 1945, factors) would not be associated with as many stuttering events as they have in English. These differences so far make predictions about the loci of stuttering events, i.e. in the analysis of speech samples, the aims of the analysis of bilingual children's language development is outlined next.

1.15 Reasons to carry out fluency research with bilingual infants pertinent to chapters 6-8

There are two main issues why the study of fluency development with children between the ages of two and three, and specifically bilingual children, should be a central focus in research into stuttering. 1) Stuttering does not start simultaneously with the onset of language production in infants. 2) There is a need to develop further test material to track a wider age range of children. These two issues highlight the need to directly look at this age group. The link between the onset of stuttering and the introduction of a second language in bilingual children (as described in detail in section 1.8.1) is the reason to focus on this particular group. In respect of the first issue Eisenson (1984) highlighted evidence that stuttering does not occur at the one-word stage. Even though most of the studies establishing stuttering onset retrospectively (this method makes it difficult to exactly establish which factors were involved at the time) it is now generally accepted that stuttering first appears between the ages two and three years (see for instance Yairi, 1983; Yairi & Ambrose, 1992b; 1999). At this age most children start producing sentences and there is a rapid development of syntax acquisition (Brown, 1973; Slobin, 1970). On a more general level Zebrowski (1991; 1994; 1995) pointed out some of the difficulties in finding characteristics that might differentiate children who stutter from those who are normally fluent at this time in their language development.

If, as highlighted already several times in this introduction, cognitive overload at this time of language development causes stuttering onset (as suggested by Bernstein Ratner, 1997; Elbers & Wijnen, 1992; Karniol, 1992; Lebrun & Paradis, 1984; Wijnen, 1990) then this would be more likely in those speakers who are acquiring two lexicons and grammatical rule systems concurrently (i.e. bilingual infants). Chapters 6 and 7 can be seen as pilot work to investigate these issues further. Syntax acquisition was measured through a receptive syntax test and MLU data whereas the lexical development was analysed by parental questionnaire data and picture naming – the individual measures taken are introduced in more depth in chapter 5. In these studies lexical accuracy rather than disfluency is focused upon.

Another aim in using bilingual children in different age groups (chapter 6 investigates syntactic understanding of school aged bilingual children as compared to their monolingual counterparts) is the investigation of some of the factors that were highlighted in the first three experimental chapters of this thesis. In this chapter the issue of compound nouns and word order is revisited.

Finally none of the established tests in language development literature can be used with children from language onset up to school age level using the same methodology. Therefore tests for a wider age range are needed in this type of longitudinal research strategy and chapter 6-8 can be seen as intensive pilot work for materials created by the author for this particular purpose.

1.16 Thesis outline

This review set out to provide an understanding of how psycholinguistic research can contribute to the investigation of where disfluencies are most likely to occur and possible reasons for this. How this linked in with existing psycholinguistic models of speech production was also explained and current models for fluency failure (that derive in part from this work) were described. The rationale and parallels to studies in bilingual language development were then established.

The model of fluency failure that was singled out as the basis for study is the EXPLAN model (Howell et al., 2002). This model explains disfluencies and where they occur within a framework of the planning and execution of speech and, unlike CRH (Kolk & Postma, 1997), makes some particular predictions about early patterns of stuttering and stuttering in bilingual speakers. The review also highlighted some relevant linguistic differences between English and German. For instance the difference in syllabic structure might be interesting to investigate further, because it is also mediated by the function/content word distinction. Another area of linguistic investigation is the large number of compound words and thus the higher number of word internal PWs in German. The word order differences might also affect PW structure and therefore are also an area that were analysed.

There now follows a brief overview. The thesis is in two parts: The first is concerned with whether findings reported on stuttering in English apply to German and where, based on the above review, differences between the languages might be expected. The second concerns how some of the same factors reasoned to operate in leading to stuttering, apply in early monolingual and bilingual development in fluent children. The plan of the thesis is as follows:

The research questions that are addressed in chapter two concern the extent to which Brown's factors operate in German. Some ways these factors might operate differently are due to the onset of vowels being more difficult for German people who stutter, and the relation between this factor and the function / content word distinction. The predictions that early sentence position do not lead to as much of an increase as in English, whereas the proportion of words in later sentence positions in German might be higher are examined. In respect of Brown's other two factors that increase stuttering rate, i.e. when words are long and when they are content in type, the same predictions are made for German. It is assumed, however, that these two factors are closely interrelated in German, as they are in English. Differences in the distribution of stuttering according to linguistic structural differences are difficult to explain within the theoretic framework of CRH since this framework does not propose activation according to phonetic but rather semantic similarity. Furthermore evidence for children's function word stuttering in languages other than English would strengthen the EXPLAN theory and weaken approaches that are purely based on linguistic complexity alone (such as Wingate's, 1988, 2002 emphasis on linguistic word stress).

In chapter three phonetic characteristics are analysed in finer detail with the index of phonetic complexity (IPC, Jakielski, 1998). It is examined how phonetic factors operate in different age groups and whether the words that are stuttered are more complex than fluently-produced words. It is also analysed which specific factors have most impact on stuttering rates and whether these are the same in different age groups and across languages. Finally this is related to the frequency of use of the characteristics in spontaneous speech. This is a more in depth phonological investigation compared to chapter one which focused mainly on whole word factors only. Again the analysis can differentiate between theories that make prediction only based on semantic activation, such as CRH (Kolk & Postma, 1997; Postma & Kolk, 1993) or those that propose only the linguistic complexity (Wingate's, 1988, 2002 view) in comparison to the EXPLAN model of fluency failure.

Chapter four examines the distribution of stuttering (in respect to the stalling and advancing classification of these events) in PW across age groups. The function/content word distinction and its development into the PW as a tool in the investigation of stuttering, has helped to provide a context in which to analyse the disfluency. This constituted a major shift away from word internal factors as analysed by Brown in his series of studies. The use of the PW as a research tool helps to provide evidence for the hypothesis that stuttering on function words depends on their position relative to content words within a PW, i.e. (lexical) word external factors (as proposed by Au-Yeung and

Howell, 1998). One of their findings concerns the exchange pattern that takes place developmentally (Howell et al., 1998). This means that younger children stutter more on function words whereas adults have a higher number of content word stutters. The question is now whether this is a universal pattern of whether it is a specific case for English. One study has already suggested that there is a similar pattern to this in German (Rommel, 2000). It is aimed to replicate and extend these findings for German speakers who stutter of different age groups. In this way it would provide direct evidence for patterns of stalling and advancing disfluencies such that for instance function words prior to their content word nucleus within a PW unit would be expected to be stuttered more frequently than those that appear after a content word. A pattern such as this would be hard to explain by either Kolk & Postma's (1997; 1993). nor Wingate's (1988, 2002) viewpoint.

The fifth chapter is mainly concerned with an introduction to the various test materials used in the second part of the thesis. Language development in fluent-speaking (mainly bilingual) children and the relationship of this research to stuttering is again highlighted. The aim is to give an overview of the language development of bilingual in comparison to monolingual children and to provide evidence that suggest why these children might be more prone to disfluencies. The bilingual model of lexical access (BIMOLA, Grosjean, 1988) is described. The issue of how fluency development relates to overall language development is also introduced.

Chapter six investigates possible performance differences in a receptive syntax task using bi- and monolingual school children. This research addresses a number of questions 1) whether there is a general delay in syntax performance for bilingual children compared to their monolingual peers, 2) whether the processing of complex content words is different in bilingual compared to monolingual children and 3) whether word order processing differs across language groups. The aim in this chapter is to validate the research tool as well as to assess how some of the factors that were singled out in chapters 2 – 4 operate in bilingual children.

A longitudinal study of bilingual infants (from age 23 months) is described in chapter seven. Here a number of measures are taken, starting with picture naming and vocabulary acquisition leading on to more syntactical tasks such as the reception of syntax test and mean length of utterance recordings. A behavioural temperament questionnaire was also given. Research questions that were addressed here are whether at times when the linguistic demands in the acquisition process are high, i.e. when both the lexicon and syntax are growing at a faster rate concurrently (issues that have been linked to stuttering onset), there would also be an increase in language errors. It is also aimed to ascertain whether there are differences in the order of syntactic stages that are acquired (e.g. the compound noun category).

Chapter eight investigates TOTs in monolingual German children in two age groups. How this is related to lexical access difficulties, as investigated by the TOT phenomenon, was also described. It is highlighted that the TOT state can represent a case of the lack of available plan. Thus it also fits into the general planning and execution (EXPLAN) framework of speech disfluencies. This can give an indication what type of knowledge children possess about words in TOT states and whether this is different across age groups. It is also investigated whether factors that are known to predict stuttering, such as word length and phone the word starts with, influence whether a word is retrieved successfully by children.

The final chapter provides a summary and explores the theoretical significance of the results. It is particularly focused on the approaches that have been introduced in this current chapter and how the findings can differentiate these theories. The problems encountered in the design and research practice are highlighted and future directions for this line of investigation are suggested.

2 Predicting Stuttering from Linguistic Factors for German Speakers in Two Age Groups – Brown's factors[1]

"Why didn't Kafka stutter? ... People who stutter devise elaborate routines to avoid or to ambush and take by surprise troublesome consonants, of which K is one of the most difficult. It's a good job Kafka didn't stutter. With two Ks he might have got started on his name and never seen the end of it. As it is he docks it, curtails it, leaves its end behind much as lizards do when something gets hold of their tail."
Alan Bennett, 1987, 'Kafka at Las Vegas', in the *London Review of Books*, reprinted in *Writing Home* (1994), p. 337

2.1 Background

As emphasised in the introduction the original impetus for research into linguistic variables associated with speech disfluency was triggered by Brown (1945). In this (his final) article, he summarised what he considered were the four basic factors that determined whether words are spoken disfluently by adults who stutter. It is recalled that the factors were: (1) word class (this has subsequently been interpreted as showing that content words are more prone to stuttering than function words); (2) word length (in syllables - long words are more difficult); (3) sentence position (words that appear in early positions are more likely to be stuttered); (4) phone the word starts with (words starting with consonants are more difficult than those that start with vowels). [2]

The investigation of stuttering events in general (such as Brown's) was criticised recently by Smith (1999). She pointed out that this approach misled researchers into thinking that stuttering is a static, rather than a dynamic, process, using the analogy of

[1] A version of this chapter appeared as Dworzynski, Howell & Natke (2003) in the *Journal of Fluency Disorders*.
[2] As was also noted, Brown's factors originally included linguistic stress and word initial position (Brown, 1938b) which he did not acknowledge in his 1945 review. These two further factors are beyond the scope of the present chapter. Please refer to Wingate (Wingate, 1979; Wingate, 1988) for discussion of the relationship between word stress and stuttering in English and to Natke, Sandrieser, Pietrowsky and Kalveram (Natke, Sandrieser, Pietrowsky, & Kalveram, 2001) for syllabic stress effects in German preschool children who stutter.

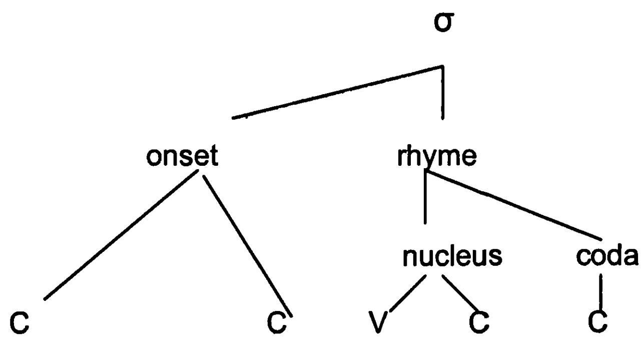
researchers investigating volcanoes by studying only the shape of the landform and type of eruptive material. The research reported here does not dispute the fact that stuttering is a multifactorial phenomenon, and Smith's criticism highlights the need to approach it from many angles, such as the study of disfluent events. Though Smith's statement which implies that surface events might be the result of unobserved processes is undoubtedly true, the point could be levelled at any aspect of behaviour that is investigated by psychologists. The view taken in this thesis is that we should seek to understand these underlying processes by formulating models (such as EXPLAN) that account for the surface behaviours (and make predictions that can be falsified).

The point was raised in the introduction, that a reason to carry out cross-linguistic, or comparative research, is to find out whether stuttering occurs in linguistic structures irrespective of their motor form; or whether difficult motor outputs lead to stuttering independent of the linguistic unit in which they occur. Some dissociation between motor and linguistic aspects can be achieved because the levels of motor complexity on different linguistic units differ between languages. Even though use of other languages allows scope for separating motor properties from the linguistic units in which they occur in English, no previous studies have made such comparisons. The main concern of the present study is why disfluency occurs on certain words and, in particular, the degree to which linguistic and motor factors affect disfluency.

Looking at the cross-linguistic work on Brown's factors in more detail than in chapter one, it should first be noted that Bloodstein (1995) pointed out that there is little work on Brown's factors in languages other than English. The exceptions are, as he notes, that the factors have been found to operate in Norwegian (Preus, Gullikstad, Grøtterød, Erlandsen & Halland, 1970 as cited in Bloodstein, 1995) and in the previously mentioned study looking at the Dravidian language Kannada (Jayaram, 1981). Although there is debate about confounding variables in Brown's factors (see for instance Bloodstein, 1995, and Wingate, 1988), and the factors are relatively crude measures (that nevertheless capture much of what determines whether a word is stuttered), the current analysis considers them as linguistic determinants of disfluencies and focuses on a comparison between languages. One of the aims of this chapter is to clarify how these factors influence stuttering rates in German. As such it provided a good starting point for the further linguistic analysis of the following chapters of part one of this thesis. Although as pointed out earlier German and English are close in origin (both stemming from the West-Germanic branch of the Indo-European language tree), there are a number of important differences relevant to the operation of Brown's factors in the two languages.

Here two aspects that were discussed at length in the first chapter (because they are applicable to the experimental work in this chapter) are briefly re-considered. One of these differences concerns syllable onset. Syllable onsets are constituents of the subsyllabic structure of German (Wiese, 1996) as shown in the diagram first shown in chapter one (σ stands for syllable and C and V stand for consonant and vowel slots):

(1)



As already highlighted in chapter 1, a syllable comprises the onset (start of the syllable) and rhyme (the syllable ending). The rhyme is generally further subdivided. Syllables have to have a phone that functions as the peak (usually a vowel), which is called the nucleus of the syllable. The coda is the sequence of one or more final consonants. All prevocalic consonants belong to the onset. The main difference between English and German is that in German an onset is obligatory. As such, syllables that start with a vowel (which would be the nucleus) have to have an onset added. In these cases a glottal stop is inserted as an onset to that syllable (Rogers, 2000). It was stated in the introduction that this should be considered a vocalic feature, but not all authorities agree. The most compelling position is that of Wiese (1996) who considers the glottal stop not as a full phoneme but a consonantal feature. He states, at the same time, that the glottal stop should not be analysed as a phoneme of Modern Standard German (p. 59). This means that a case can be made that there is no separate consonantal segment and the vowel gets the feature [+ onset] which is then realised phonetically as a glottal constriction with a tendency to be located at the beginning of the vowel. This leads to stronger vowel onsets in German than in English. Previous work has shown that in both Dutch and Afrikaans (languages where there is also more articulatory tension in initial vowels) more words with initial vowels are stuttered (Uys, 1970 as cited in Bloodstein, 1995; Vaane & Janssen, 1978). It would be predicted that this would also be the case for German.

The introduction included a discussion of syntactic differences between German and English. Most of this is not required here, where Brown's way of looking at the effects of syntax is crude (position a word occupies in a sentence). The widespread view that the German language is inflexible in its syntax (see for instance Richards, Schmidt Mackey, Mackey, & Gibson, 1960) was mentioned in the introductory chapter. This supposed inflexibility is mainly due to the fact that the finite verb and the past participle have rigid positions in German sentences. Verb positions are rule governed in all languages the point to stress here is that according to German syntactical rules they would be frequently placed sentence final. To briefly recap from the last chapter, all finite verbs must be in second position in independent clauses, and non-finite ones in final position (with the exception of cases of extraposition of adjuncts or arguments). In dependent (subordinate) clauses, all verbs are in the final position.

There is a much greater positional flexibility with regards to the position of the subject, direct object, and indirect object in German noun phrases. The reason for this is that in German the case endings / markings always indicate how the constituents fit together syntactically (see the discussion and examples given in the first chapter). In other words the issue to be highlighted here is that German sentence planning is more end loaded compared to English. With respect to grammatical encoding verbs appear on average (i.e. non-finite verbs and in subordinate clauses as described above) later within sentences and noun phrases can occur in various positions (earlier as well as later in sentences) when compared to English. Verbs being rich in morphological encoding, it could be argued that there would be more planning involved for words in later sentence positions in German compared with English. Thus it would be predicted that earlier sentence positions would not be as subject to increases in stuttering rate as they are in English. This is investigated below in German speakers and compared to Brown's original results on English. An interesting finding from German children who stutter corroborates an argument of positional differences. According to Rommel and colleagues (1997; 2000) disfluencies in a group of German stuttering children (of pre-school age) appeared increasingly in the middle of sentences and words. In Bock and Levelt's (1994) review of the planning of grammatical processes, this influence would fall into the stage of positional encoding. This stage determines the serial order of the lexical elements in an utterance.

The important issue of word class has already been mentioned in the first chapter. For function and content words, the predictions for German adults would be the same as those observed by Brown (1945) for English. Namely, content words are stuttered more than function words.

The importance of word class differences for stuttering research has already been introduced in depth in the review chapter. It should be noted here that the neurological substrates for word class processing have also been investigated. Differences in the processing of content and function words are now increasingly found in fluent speakers both for the activation of their neurological substrates, using EEG studies (Bastiaansen, Van Der Linden, ter Keurs, Dijkstra, & Hagoort, 2002; Brown, Hagoort, & ter Keurs, 1999; Osterhout, Allen, & McLaughlin, 2002), in terms of lexical access (Segalowitz & Lane, 2000) and their role in a cross-linguistic examination of speech errors (Wells-Jensen, 2000). The function/content word distinction has also been investigated in an EEG study with individuals who stutter (Weber-Fox, 2001). Weber-Fox's results indicated that the event related potentials (ERPs) of people who stutter were characterized by reduced negative amplitudes for closed-class words, open-class words, and semantic anomalies in a temporal window of approximately 200-400 ms after word onsets. Differences between adults and children who stutter have been observed with respect to the function/content word distinction. An exchange of function and content words over age ranges has been observed in English (Howell et al., 1999) and Spanish speakers who stutter (Au-Yeung, Vallejo-Gomez & Howell, 2003). In the exchange pattern, children are more disfluent on function than content words; whereas the opposite holds for adults, giving something like a cross-over pattern for disfluency rate across these word types (see chapter four for results showing this pattern in German, Dworzynski, Howell, Au-Yeung & Rommel, 2004). For German an increased rate of function word disfluencies was also reported by Rommel (2001) in pre-school children who stutter.

For the word length factor, stuttering rate would be expected to be higher on long words than short words, as Brown (1945) reported for English.

The current study investigated Brown's factors in spontaneous speech samples of German adults and children who stutter. The following research questions were addressed. In analysis one it is investigated whether in two German age groups (children and adults) stuttered words have a higher factor score (as calculated by the mean sum of Brown's factors) than fluent words. Here it was assumed that adults show a greater difference between stuttered and fluent words than the children. Analysis two was concerned with the question whether each additional factor score (as calculated by the sum of Brown's factors for each word) were having an impact on stuttering rates for the German speakers in the two age groups. The hypothesis was that there is a linear trend for the adults but not for the children. Analysis three is then concerned with the investigation of the individual factors, specified by Brown (1945) and how they operate in the two age groups. For adults it was predicted that the stuttering rate of words starting with vowels would be high in German compared with English. Early sentence positions would have a lower level of stuttering than in English. The content word, and word length, factors were expected to operate in German in a similar way to English. Furthermore, in line with previous research, it was predicted here that children who stutter would not be affected in the same way as adults by the factors associated with articulatory difficulty (Bloodstein and Gantwerk, 1967; Bloodstein and Grossman, 1981). Specifically with respect to the function/content word distinction, it was expected that for children there would be less of a difference in disfluency rates on these word types compared with adults.

2.2 Method

2.2.1 Participants

Participants were all native speakers of German. All had been diagnosed as exhibiting stuttering behaviour. Of the 15 adults, five were female and ten were male. Their ages ranged from 16 years 3 months to 47 years 1 month, with a mean age of 29 years and 8 months. The two sixteen year old subjects were included in the adult group for the following reason. Even though ages of stuttering onset vary widely, Bloodstein (1987) summed up the available research with the conclusion that stuttering is essentially a disorder of childhood. Moreover, a study by Dickson (1971) suggested that the peak age of spontaneous recovery was three and a half years, with the great majority of former children who stutter having recovered by age six. Although recovery can take place at any age, it is altogether more likely to occur pre- rather than post teens (Bloodstein, 1987) – obviously also depending on the age of onset. It can therefore be argued that the subjects in the adult group exhibit an established stuttering pattern. A conclusion that is also supported by Howell et al.'s (1999) study, which reported an established stuttering pattern by age eight. They were all voluntary participants. Recordings of seventeen, school-aged children (six girls, 11 boys) were analysed. All of the children were attending a speech therapy centre in Bad Salzdetfurth. Their ages ranged from 7years 4 months to 11 years 11 months, with a mean age of 10 years 1 month. For a detailed summary of all the speakers please refer to Table 3 in the next chapter.

2.2.2 Apparatus

Recordings were made in a relaxed and quiet atmosphere. For the adults, recordings were made in a quiet room at the University of Düsseldorf. Subjects were video/audio recorded using a VHS Movie Camera, Panasonic NV-SX30EG. The audio track of the videotape was transferred onto DAT tape. The children were recorded on tape in quiet surroundings at the speech therapy centre. All material was then transferred to the computer hard disc at a 48 kHz sampling rate, 16-bit samples. The recordings were later down-sampled to 10 kHz. A Beyerdynamic (professional DT 770) headset was used during transcription of the samples to listen to speech output played over speech filing system (SFS) software.

2.2.3 Speech Material

Spontaneous speech samples, a minimum of two minutes in duration, were analysed. The samples for the two age groups were obtained using similar procedures. The younger speakers either produced a spontaneous monologue on a topic of their choice, or a sample was elicited by prompts given by a therapist. Prior to the monologue, the subjects were given suggestions as to topics, such as family, friends, hobbies, films etc. The recordings of the adults consisted of stretches of spontaneous monologue elicited by a researcher/speech therapist. The topics the adults used were current employment, schooling or hobbies. The average length of recording for the two groups was similar (about 3 minutes of speaking time).

2.2.4 Transcription Procedure

Orthographic transcriptions were carried out using the Talkscribe transcription software. For the phonetic transcriptions a machine readable transcription alphabet was used (the Joint Speech Research Unit alphabet – JSRU). This is an alphabet originally developed for text-to-speech synthesis (for full details please refer to the UCL speech group's web site: http://www.speech.psychol.ucl.ac.uk/). The alphabet contains a full set of consonants and vowels for the transcription of English speech, and was extended to include suitable symbols for the transcription of the German samples. SFS was used in the phonetic transcription process to allow accurate location of events on an oscillographic display of the waveform. SFS also facilitated the transcription process by having two mouse-operated cursors, superimposed on the waveform, to specify a given area for replay. The German data were transcribed in the method described by Kadi-Hanifi and Howell (1992) (i.e. a broad transcription for the fluent regions and a narrow system in the region of disfluencies). The person responsible for the transcriptions (KD) measured the duration of pauses and prolonged segments to the nearest 50 ms. Syllabification was entered on the transcriptions and all words were categorised as content or function. The stuttering episodes that were marked were monosyllabic word and part-word repetitions, blocks and segmental and syllabic prolongations.

2.2.5 Coding of the Speech Samples

All words were coded according to the four factors Brown had investigated. A word was scored for every factor Brown had identified as being associated with a higher rate of stuttering. To do this, each factor was examined separately and a word was given a score of 0 or 1 for each factor, according to the following contingencies: 0 was given for the respective factor when a word started with a vowel, when it was a function word, when the word was shorter than five phonemes, when the word occupied one of the first three positions in an utterance. 1 was given for the respective factor when a word started with a consonant, when it was a content word, when the word was longer than five phonemes, and when the word occupied a position beyond the first three in an utterance. Words were also coded as produced fluently or stuttered. Filled pauses (e.g. 'um', 'er' etc.) were excluded from the coding procedure.

2.2.6 Reliability Measures

The researcher was trained in transcription skills. In this respect English speech samples (not used in the current study) were transcribed and cross-checked by more experienced speech researchers until the given sections could be transcribed without error. This was done by employing first samples from fluent speakers and then samples from speakers who stutter. For the speech data in the current study a number of completed samples were also re-checked by more experienced staff. For the transcription of the German speech samples the researcher with a background in psychology rather than linguistics took a German phonology course in University College London, Linguistics Department which involved training in transcription of German. However, due to the unavailability of experienced German transcribers, rather than using inter–judge reliability values, a measure of consistency rather than reliability was taken. Even though this procedure has obvious shortcomings the combined work of Cordes and Ingham (e.g. Cordes & Ingham, 1999; Ingham & Cordes, 1997) suggests that judges who show a high intra-rater agreement also have the tendency to have the highest inter-rater agreement. This meant that the same researcher re-transcribed and coded 20% of the samples, which were then analysed using Cronbach's alpha as a measure of consistency. A higher alpha coefficient signifies better consistency. Nunnaly (1978) indicated 0.7 to be an acceptable reliability/consistency coefficient, though lower thresholds are used in some of the literature. For the re-transcribed and re-coded samples alpha values ranged from 0.88 to 0.98, indicating a high level of consistency.

2.3 Results

2.3.1 Analysis 1: Are Factor Scores for Stuttered Words Higher than for Fluent Words in Both Age Groups?

Individual stuttering rates were calculated by dividing the number of stuttered words by the total number of words (i.e. stuttered and non-stuttered). Stuttering rates for the individual word factors were calculated in the same way. Thus, for content word stuttering rate, for instance, the number of stuttered content words was divided by the total number of content words both stuttered and non-stuttered. For the adults, stuttering rates over all words ranged from 1.95% to 39.30% with a mean of 15.73%. The respective values for the children ranged from 6% to 42.47% with a mean of 19.08%. Brown's factor scores were obtained by totalling the number of factors on each word (between 0 and 4, called the factor value). So a score of 0 would be obtained for a short, function word, beginning with a vowel, in a late sentence position. Stuttering rates for words with different numbers of factors were then computed by dividing the number of words in each of the five factor value groups by the total number of words in that group. Following Brown, speakers who stutter were subdivided into a severe and a moderate group. Brown also had a mild group but this was not appropriate for the current data set (only one of the samples could have been classified as such; this sample is included in the moderate group). The division into groups was carried out using individual stuttering rates. Rates higher than 10% for adults and higher than 12% for children were classified as severe and samples below that were moderate –different percentages were used because of the higher mean average stuttering rate in the children's group (i.e. to get roughly equal group sizes in both age groups). There were seven moderate and eight severe adults and seven moderate and ten severe children.

If articulatory complexity is a determinant of stuttering it would be expected that stuttered words would have on average a higher combined factor score than words that are not stuttered. It would also be expected that stuttered words of adults would have higher factor scores than those of children. Figure 6 shows factor scores of stuttered and fluent words separately for the two age groups.

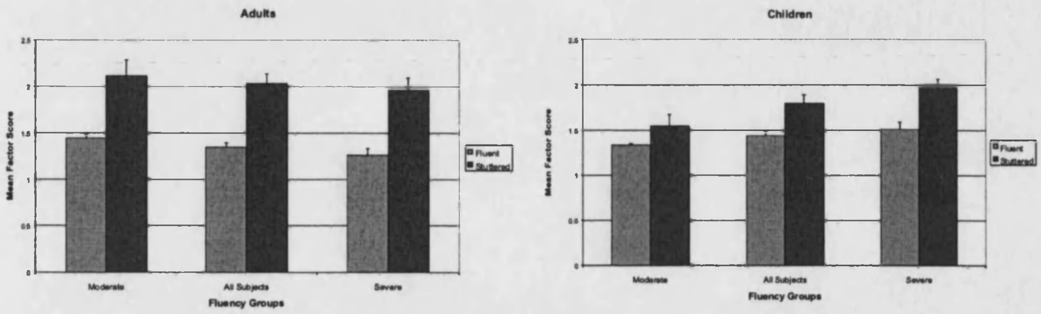


Figure 6 - a and b: Average factor values (with their associated standard error bars) as indicated on the vertical axes. The differently shaded bars represent stuttered and fluent words. Both age groups (adults on the left and children on the right figure) are divided into two subgroups (moderate on the left and severe on the right) and the overall set (in the middle) which are indicated on the abscissa – see text. All differences in the left are significant p<0.01, whereas only the combined and severe subgroups are significant for the children's age group (graph on the right).

The adult data are presented in the left-hand section of Figure 6. This shows that stuttered words had a higher average factor score than the fluent words. This was significant for both subgroups and for the combined data (severe $\underline{t}(7) = -3.76$, $\underline{p}<0.001$ moderate $\underline{t}(6) = -3.73$, $\underline{p} = 0.01$, all $\underline{t}(14) =-5.47$, $\underline{p}<0.0001$ respectively). The children on the whole showed the same trend. The combined data for the whole group showed a significant difference, with stuttered words receiving higher factor scores than non-stuttered words ($\underline{t}(16)=-4.30$, $\underline{p}<0.001$). This was also the case for the more severely disfluent children shown on the right of Figure 6 ($\underline{t}(9)=-4.28$, $\underline{p}<0.01$). However, the difference in factor scores in the moderate subgroup was not significant. When compared with the adults, articulatory difficulty was not as powerful a determinant for words stuttered by children. The two parts of Figure 6 indicate that the difference in factor scores for children was not as great as that for adults in all cases. A mixed model ANOVA was performed on these data with the repeated measures factor Fluency (stuttered or fluent words) and two independent groups factors, Age (children or adults) and Fluency subgroup (moderate or severe). This revealed a significant interaction between Age and Fluency ($F(1, 28)=5.43$, $p<0.05$), showing children have higher factor scores for fluent words and lower factor scores for stuttered words, again reflecting a smaller difference between word types for the children. An Age by Fluency subgroup interaction also occurred, which indicated that there was a bigger difference in stuttering rate between moderate and severe adults than between moderate and severe children ($F(1, 28)=12.02$, $p<0.01$).

2.3.2 Analysis 2: Is There a Linear Increase in Stuttering Rate with Each Additional Score in Both Age Groups?

If linguistic difficulty is a good predictor for disfluencies, then the more complex a word, the greater the likelihood of stuttering. Average stuttering rates were computed for words with different degrees of difficulty according to the specified factors in the way described by Brown (1945). The words were categorised according to factor value and stuttering rate for each of the resulting five factor-value groups was calculated. A linear trend between factor-value and stuttering rate was predicted for adults (as Brown found with English adults). If linguistic difficulty is not as predictive in childhood as it is with adults who stutter then this trend should not show up in the data of the children. It was apparent from examination of the data that the number of words with a factor value of 4 was negligible so these words were excluded from this analysis. Figure 2 shows the level of word difficulty and the associated stuttering rate for both groups (again subdivided into the two fluency severity subgroups).

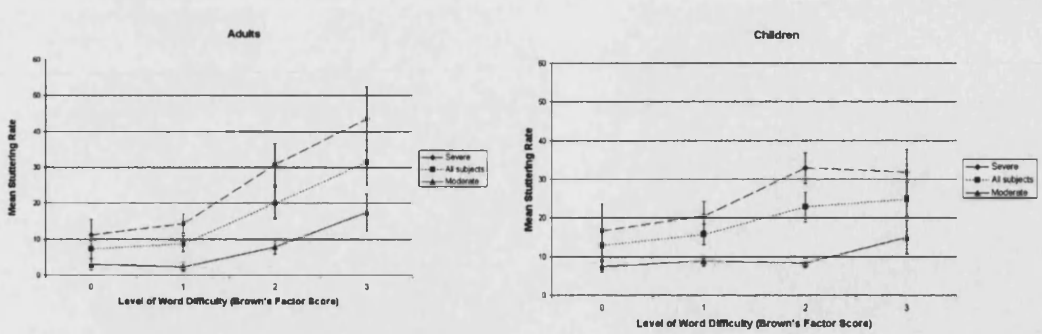


Figure 7 - a and b: Sum of Brown's factor scores are given on the abscissa and mean stuttering rate is represented on the vertical axes. Again the left figure shows the adult and the right graph the children's data. Different Lines indicate the two subgroups (moderate – bottom line and severe – top line) and the combined data (middle line) for each age group.

For the adults (left-hand section of Figure 7), increasing the level of word difficulty (as indexed by factor-value) increased the stuttering rate. A mixed model ANOVA was performed on these data with Factor-value as a repeated measures factor (with four levels: 0, 1, 2 and 3) and Fluency subgroup (moderate and severe) as an independent measures factor. The main effects of Fluency subgroup and Factor-value were significant ($F(1,13)=57.066$, $p<0.0001$ and $F(3, 39)=12.422$, $p<0.0001$ respectively). The interaction was not significant. Stuttering rate over Factor-values increased linearly ($F(1, 13)=14.234$, $p=0.002$).

A different pattern was observed in the younger age group (right-hand section of Figure 2). It is particularly striking that for the moderate subgroup, there is very little increase over factor-values. This would explain why the data in Figure 6 do not show a significant difference between stuttered and non-stuttered words for the moderate subgroup. Inspection of the combined data revealed that there was an increase, although this trend was not significant. For the severe subgroup, there was a main effect of Factor-value, although none of the follow-up tests with Bonferroni correction showed any significant differences between pairs with different factor-values.

A mixed model ANOVA was performed that compared children and adults. A significant interaction was found between Age and Factor-value when individual stuttering rate (i.e. overall individual percentage of stuttered words) was controlled as a covariate ($F(3, 87)=2.742$, $p<0.05$). The origin of this can be seen by inspecting the middle line in each section of Figure 2. 'Easier' words (those with a factor-value of 0 or 1) were stuttered at a higher rate by children – children's factor-value score 0 had a stuttering rate of 12.91 and factor-value score 1 had a stuttering rate of 15.66 (in adults these were 7.31 and 8.71 respectively). Only the difference between 15.66 and 8.71 between the two age groups, i.e. the difference on factor score 1, approached statistical significance ($p=0.055$ two-tailed). The opposite effect over age groups occurred for words with a factor score of three, where adults showed a higher stuttering rate than children (31.40 and 24.86 respectively, although this was not statistically significant).

2.3.3 Analysis 3: Which Factors Have the Highest Impact on Stuttering Rates?

The question next arises, as to which of the factors has the biggest influence on stuttering rate when considered on its own. To investigate this, the mean stuttering rates were computed for all of the different word characteristics (function/content, starting with a vowel or consonant, short/long, and positioned early or late in a sentence). These were derived from the stuttering rates for each of the individual speakers on these aspects. For this analysis the data were not subdivided into different levels of severity within each age group. Adjusted stuttering rates for each word characteristic are shown for adults (left) and children (right) in Figure 8 (the adjusted mean stuttering rate refers to the adjustment made when individual stuttering rate was treated as a covariate).

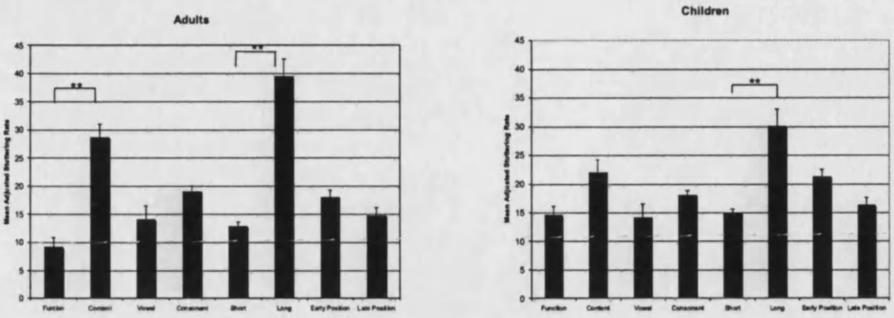


Figure 8 - a and b: The a and b figures represent adults and children respectively. Along the horizontal axes individual word characteristics with regards to Brown's factors are given. Bars represent the mean adjusted stuttering rate with their associated standard error bars. ** signifies $p < 0.01$.

First, within each of the age groups comparisons were carried out between individual factors only (the significance level was adjusted using the Bonferroni correction). Starting off with the adults it can be seen that two of the factors were associated with a dramatic increase in stuttering rate. This was clearly the case with content words and with long words. As such, the mean stuttering rate for both these factors increased nearly three times. For function words the rate was 9.14%, whereas 28.67% of content words were stuttered ($t(14)=-4.38$, $p=0.001$); the increase from short to long words was from 12.93% to 39.45% ($t(14)=-5.11$, $p<0.0001$). It can be assumed that these factors are interrelated. German function words are usually shorter than content words, as is the case in English. The two remaining factors, (whether words began with vowels or consonants, and word position in a sentence) were not significant. Trends for both these factors, however, increased in the direction that was outlined for English by Brown (i.e. words beginning with consonants and those with early sentence position were associated with a higher stuttering rate). However for English the differences in these two factors have been shown to be significantly different. This means that the hypotheses concerning the vowel onsets and sentence positions cannot be definitely ruled out by the current data set.

It can be observed from the section of Figure 8 representing the children's data, that none of the factors produced as dramatic an increase in stuttering rate as the respective factor in adults. A large significant increase occurred only for the difference between stuttering rates for short and long words ($t(16)=-3.66$, $p<0.01$) where mean stuttering rate doubled from 15% to 30.03%.

A comparative analysis was carried out between adult and child data in respect of the individual factors, using a mixed model ANOVA (with individual stuttering rates treated as a covariate). The two repeated measures factors were Brown's factors (starting phone, grammatical class, length, utterance position) and the level of difficulty (easier/harder, i.e. vowel-consonant, function-content, short-long, early-late). While the independent groups factor was age group, the dependent variable was the individual stuttering rate for each of the measure. This analysis showed a significant interaction between Brown's four factors, their individual levels and age ($F(3, 87)=2.861$, $p<0.05$). This interaction was due to the smaller increases observed in the child data. For instance when considering the function/content word distinction, although both age groups showed the expected increase from function to content word disfluency, children can be seen to have a higher function word disfluency rate, whereas their content word disfluency is lower than that of adults. The content/function factor showed a significant interaction with Age on its own (individual rate again used as covariate, $F(1, 29)=4.963$, $p<0.05$). Follow-up tests showed that function word disfluency was significantly higher in children ($F(1, 29)=5.697$, $p<0.05$) whereas the opposite pattern was found for content words ($F(1, 29)=4.205$, $p<0.05$). An interaction was also observed for the factor of word length with age where individual rate was again used as covariate ($F(1, 29)=4.975$, $p<0.05$). As can be seen from Figure 8, this is again due to a cross-over effect similar to the function/content word factor: Children stutter significantly more frequently on shorter words ($F(1, 29)=5.195$, $p<0.05$) but adults more frequently on longer words

($F(1, 29)=4.812$, $p<0.05$), in other words the difference between the stuttering rates of long and short words is larger for adults than for children.

In summary, the results confirmed the hypothesis that linguistic factors do not affect children in the same way as adults. Comparisons with the adults showed that stuttered words were not associated with as great an increase in linguistic difficulty as were adults. Furthermore, with each additional level of word difficulty stuttering rate did not increase as markedly as it did for the adult group - where a clear trend was found. It was also found that words with a lower factor-value (a value of 0 and 1) had a higher stuttering rate for children than for adults. Analysis of the specific word characteristics of children revealed that none of the factors were associated with as much of an increase as with adults. Specifically function/content word disfluency and word length analyses revealed a change in pattern with a cross-over effect. Children stuttered significantly more on the 'easier' level (function and short words) of these two factors, but significantly less on the 'harder' level (content and long words respectively).

2.4 Discussion

Overall, for the adults it was shown that increasing word difficulty was associated with an increase in disfluencies. As such, linguistic complexity is a good predictor for stuttering in adulthood. When considering the comparison between adults and children in respect of Brown's four factors, the striking and counter-intuitive finding is that children stutter more on linguistically 'easier' words. The debate about whether or not repetitions of whole function words can be classified as stuttering was already introduced in chapter one and it is relevant to the current chapter. Wingate's (2002) position is that these forms of disfluencies do not constitute stuttering whereas Yairi and co-workers (see for instance Yairi et al., 2001) oppose his view. The German data presented here gives an indication that the children (all of whom had been diagnosed as children who stutter) showed a high proportion of disfluencies on function words. In that way it would give support for Yairi et al.'s (2001), who classify monosyllabic whole word repetitions as stuttering, rather than Wingate's (2002) view, who claims that these disfluencies are not stuttering events.

For instance, an interesting finding is that the German data showed a trend that with increasing age, content word stuttering increased and function word (which can be seen as the 'easier' word type) disfluencies decreased. This exchange pattern has previously been found by Howell et al., (1999) for English, by Au-Yeung et al. (2003) for Spanish, and by Rommel (2001) for German. It needs to be pointed out that the majority of the children were over ten years old. Although recovery can take place at any age, it is more likely to occur pre-, rather than post-, teens (Bloodstein, 1987). It can therefore be argued that the subjects in the current investigation exhibit a relatively established stuttering pattern. That means that more of a developmental change would be expected if even younger children had been available for test. This is an area that is investigated with an additional pre-school age group in chapter three.

The results highlight the fact that the effect of phonetic difficulty is linked to the content word class. This corroborates with the idea of the upcoming difficulty as suggested by the EXPLAN model. However Brown's factors do not provide a detailed analysis of phonetic characteristics. These were investigated, separately for function and content words, in the next chapter using Jakielski's (1998) index of phonetic complexity (IPC).

Howell and Au-Yeung's (2002) EXPLAN theory was explicitly developed to account for the increased incidence of disfluencies on function words in early development. In their model they assume that the cause of all stuttering is on the content word. Content words are phonetically more complex than function words in English, as they are also in German. Fluency failure is then viewed as a sign that planning and execution processes are out of synchrony. The pattern of children's disfluencies, observed in the current investigation, is then interpreted as a way of gaining time to complete the plan for the following content word.

When considering the hypotheses set out in the introduction with respect to sentence position and words starting with vowels and consonants, note that it is difficult to draw any firm conclusions from null results. The comparisons on these factors were not significant and, therefore, compared with Brown's (1945) English data, the German speakers show less of an increase in stuttering rate. However, the data do show a similar trend to the English analyses, albeit much less pronounced and non-significant – however both words starting with vowels and positioned earlier in sentences were associated with a higher stuttering rate. The non-significance of these differences means that the hypotheses concerning articulatory tension in vowel onsets and syntactical sentence position (set out in the introduction to the present study) cannot be definitely ruled out.

One of the differences between the current methodology and that used by Brown (1945) is that in his case the reading of a pre-prepared text was analysed; whereas the present work used spontaneous speech samples, which had the disadvantage that children quite possibly used words that were linguistically and syntactically less difficult. This may lead to a general effect of reduced stuttering, but here children obviously showed a higher level of stuttering than adults. The main advantage of spontaneous speech samples is their ecological validity. The child is allowed to use speech at a level appropriate to his or her stage of phonological development, so they minimise other potentially confounding factors, such as the construction of a text which might be too easy for adult readers/too difficult for children.

In respect of a differentiation between linguistic and motor factors in stuttering research, which is one of the issues cross-linguistic research helps to address, some concluding remarks can be made concerning the current results. It seems to be the case that a motor aspect such as laryngeal activity in vowel onsets might play a part in differences in the observed stuttering pattern. However, these results have to be considered with caution due to the slightly inconclusive findings (null result in the consonant/vowel difference). The detailed analysis by Rogers (2000) showed that glottal stop insertion is mediated by a number of factors such as word stress, position, spontaneous or read speech, and also the function/content word distinction. This makes interpretation of the current findings more difficult and suggests that further work needs to be carried out.

It is difficult to say the same about the pattern reported for the younger group. On the whole they show that linguistic factors do not affect their disfluencies in the same way as with adults, i.e. linguistically simpler words were stuttered at a higher rate. This leads to the conclusion that both linguistic and motor complexity are not as strong a determinant in childhood, as in adulthood stuttering.

3 Predicting stuttering from phonetic complexity in German[3]

"Articulation is the tongue-tied's fighting."
Tony Harrison, 1978, 'On not being Milton'

3.1 Introduction

The Brown factors investigated in chapter two provided a simple and effective way of determining which words are more likely to be stuttered. The attraction of this approach is that (given the fact that most work on linguistic factors stems from Brown, Wingate, in press) it provides a benchmark so that the findings with German can be compared with other languages. On the other hand, it is a fairly crude system. For instance, "sentence position" can be regarded as a syntactic factor, but then using position in a word as a measure of syntax misses many nuances of syntax. Also, "starting with consonant/starting with vowel" is an elementary phonetic measure (and, it might be added, as it is based on orthography in Brown's own work, possibly misleading). Though Brown's work is the appropriate benchmark, improved techniques for investigating these and other factors are needed. In this chapter, an improved method for investigating the phonetic factor is examined.

As highlighted in chapter one and further investigated in chapter two, stuttered events do not occur at random points in utterances. In particular, their position is constrained in part by the linguistic properties of the segments that make up the utterance. There is some variation across ages both in terms of the type of stuttering

[3] A version of this chapter is published in the Journal of Fluency Disorders (Dworzynski & Howell, 2004).

events that occur and where the different types of stuttering events are positioned relative to the linguistic units that give rise to them. The description of stuttered events, developed in work on adults, needs to be refined to reveal the relation of these events to the points where linguistic difficulty within an utterance is high and how this varies over age groups. Johnson and associates' (1959) list, for characterising stuttered events in adult speech, is used as a starting point. The events on this list are; 1) interjections of sounds, syllables, words and phrases, 2) word repetitions, 3) phrase repetitions, 4) part-word repetitions, 5) prolonged sounds, 6) broken words, 7) revisions, and 8) incomplete phrases.

It is difficult to specify what linguistic characteristic led to event types 7 and 8 and how many words these events affect, so they are often not included in stuttering assessments. Howell, Au-Yeung, Sackin, Glenn and Rustin (1997) developed a parser for stuttered speech to remove them, leaving events of types 1-6. Howell (in press) has advocated that the first three categories should be grouped together (all involve hesitation or repetition of whole words, which are termed generically 'stalling disfluencies'). Howell (in press) also suggests that the remaining three categories should be grouped together as they involve breakdown within a word ('within-word stutterings').

One thing associated with the change in stuttering events over ages, is that whereas within-word stutterings occur most often on content words, stalling disfluencies occur on or around the phonetically simpler function words (Howell et al., 1999). As described before, content words are nouns, main verbs, adjectives and adverbs that constitute an open class of words that expands as new words are added to a speaker's lexicon (see Hartmann & Stork, 1972; Quirk et al., 1985, for basic introductions). Function words are the remaining words (articles, pronouns, prepositions, conjunctions and auxiliary verbs) that are a closed class or words that is not added to once the grammar of a language has been established. A second observation, made by Bloodstein (2002), was that function words that are repeated by children, are often produced fluently as in the utterance 'his ... his... his... strawberry'. These two facts suggest that it is unlikely that there is anything inherently wrong with preparation of the function word, so repetition or hesitation around them is not determined by difficulty in preparing these words for output. Howell (2002; in press) suggests stalling delays the attempt at a subsequent content word (content words would be difficult to prepare as they are more likely to include phonetic structures with complex characteristics). Stuttering on content words, unlike stalling that involves whole function words, affects the initial parts of these words alone. This is consistent with the view that the content words are not completely prepared – as already described in the introduction to EXPLAN (Howell, 2002; Howell & Au-Yeung, 2002).

The difference between Howell's (2002) and Wingate's (2002) views on stuttering have already been outlined in detail in chapter 1. Here some aspects of this debate pertinent to the aims of the current chapter are briefly recapped. Howell's taxonomy that separates stalling and within-word stutterings differentiates what Wingate (2002) considers to be true signs of stuttering (the within-word stutterings) from those events he would not consider to be characteristics of stuttering (the stalling disfluencies). Wingate's point of view reflects the fact that he has worked almost exclusively with adults who stutter who show a preponderance of within-word stutterings (conversely stallings predominate in the speech of children who stutter). Wingate (2002) argues that metrical influences specify the locus of complexity in utterances that have within-word stutterings. In particular, he suggests that the stuttering problem arises at the onset–rhyme transition, and the difficulty at this point is compounded when the word carries stress (Wingate, 1984; 1988; 2002).

Howell (2002; in press), in contrast, has focussed on how phonetic structure within syllables and words determines how difficult a word is. As indicated earlier, phonetic complexity in English and many other languages depends on word type to some extent. Consequently, these word classes should be examined separately to establish any relation to stuttering (Howell & Au-Yeung, 2002; Howell, Au-Yeung, & Sackin, 2000; Howell et al., 1999).

Phonetic influences are not ruled out by Wingate, and metrical influences are not ruled out by Howell, as factors leading to within-word stutterings even though these authors place the emphasis on different factors. Thus Wingate (2002) when commenting on the notion of whether particular phones cause difficulties for speakers who stutter, indicates that such influences can be subsumed under syllable constituency at word onset. Wingate's notion would include phonetic variables like whether a word includes a consonant string at onset and the manner of the consonants in these strings (both of which are factors that Howell et al., 2000, have examined). Wingate (2002) also emphasised that his studies that indicate that stress is an important determinant of stuttering, used somewhat contrived utterances involving English samples in which function words were stressed. It is unusual for function words to be stressed in English (whereas English content words are frequently stressed). Use of such artificial material, may prohibit generalisation of the results to more typical speech.

Howell (submitted), on the other hand, has conducted a study to determine whether metrical, as well as phonetic, factors affect stuttering in natural (spontaneously produced) material. English was not used as it is difficult using this language to dissociate the influence of syllabic and metrical factors, given that content words tend to weight highly on indexes of phonetic complexity and stress is also carried almost exclusively on these word types. Spanish, on the other hand, has stressed function words, so stress can be dissociated to some extent from phonetic factors associated with lexical word class. In an analysis of Spanish speech, Howell (submitted) found both phonetic and metrical factors are important, and independent, determinants of stuttering in adults who stutter: The importance of phonetic factors was indicated by the fact that non-stressed content words had higher stuttering rates than non-stressed function words. The importance of stress was indicated by the fact that stressed function words had higher stuttering rates than non-stressed function words.

Though phonetic or metrical difficulty can precipitate within-word stuttering, on content words, these sources of difficulty do not account for why stalling occurs on function words. Wingate's position where he dismisses the latter class of events as characteristics of stuttering avoids the issue of specifying whether and how linguistic properties lead to stalling and why stalling changes to within-word stuttering with age. This stance is reasonable if the speech of adults who stutter alone is considered, as within-word stuttering predominates in these speakers (Wingate, 2002). However, this is more problematic for those researching into childhood stuttering as children who stutter show high incidences of stalling (Conture, 1990). Understanding the change from stalling to within-word stuttering over age groups might provide important clues about why stuttering persists in some speakers.

The current study examined whether phonetic properties determine which words are stuttered by German speakers who stutter. Metrical factors are regarded as affecting stuttering on content words in the same way as they do in English. They are not examined in this study. They are treated as a random-effect factor within each word class that is not linked to words with particular types of phonetic difficulty. Ages of the speakers examined ranged from 2 years to adulthood, and function and content words were analysed separately for all age groups. In the following, 1) details of different phonetic metrics that have been used recently and justification for the one selected for this study are given, 2) reasons why an analysis of German is informative are presented, and 3) the patterns expected for German speakers who stutter of different ages are outlined.

There have been two recent attempts to quantify the phonetic difficulty of words for English that can be used to investigate how well these indices of complexity predict stuttering. The first index was derived by (Throneburg et al., 1994). One factor they included was whether consonants appeared early or late in development. They selected nine consonants that Sander (1972) showed are acquired late in development (the late emerging consonants, LEC). They also assessed whether a word had a consonant cluster (CS) and whether words had more than one syllable (MS). All factors were examined for all words (function and content words together) and CS and LEC factors were scored when they appeared at any position in a word. Throneburg et al. (1994) reported that for all words these three characteristics had no effect on stuttering rates of pre-school children who stutter.

However, Howell et al. (2000) analysed function and content words separately, given that stuttering may have different roles for each of these word classes. They also investigated Throneburg et al's. (1994) factors when they appeared in initial position in an utterance (the position that leads to the majority of instances of stuttering). They reported, for adults, that when phonetically complex material occurred in initial position in content words (but not function words), there was an increased incidence of stuttering compared with the function words that were phonetically simpler.

The second metric to specify the level of phonetic difficulty of words was developed by Jakielski (1998). Her index of phonetic complexity (IPC) was motivated by MacNeilage and Davis' (1990) frame/content hypothesis about how speech is acquired. 'Frames' are the rhythmic oscillations of the mandible in early babbling. When the infant develops control over the articulators the 'content' emerges, that is more variation in production of segments as children change manner, height and place of their articulations. Jakielski (1998) developed measures for eight phonetic factors (indicated in Table 1) using MacNeilage and Davis' framework and specified how these can be combined to yield a metric of difficulty. Phonetic properties that occur in early development are deemed to be easy and receive no score for difficulty (zero) whereas those properties that do not occur in the babbling stage are considered difficult and whenever one of these phonetic attributes occurs, it is given a difficulty score of one point. An overall IPC score is then calculated by adding up the scores on the eight separate factors. The attributes that score a point and those that do not for the eight individual factors are given in Table 2.

IPC scoring scheme		
Factor	No Score	One point each
1. Consonant by Place	Labials Coronals Glottals	Dorsals
2. Consonant by Manner	Stops Nasals Glides	Fricatives, Affricates, Liquids
3. Singleton Consonants by Place	Reduplicated	Variegated
4. Vowel by Class	Monophthongs, Diphthongs	Rhotics
5. Word shape	Ends with a vowel	Ends with a consonant
6. Word Length (Syllables)	Monosyllables, Disyllables	>=3 syllables
7. Contiguous Consonants	No Clusters	Consonant Clusters
8. Cluster by Place	Homorganic	Heterorganic

Table 2: IPC metric and scoring scheme. Along the left hand column are the eight phonetic aspects that are included in the scheme. Words receive a point when they show one of the characteristics in the right column (no point when they have the related characteristic in the middle column).

There are similarities between the factors involved in Jakielski's (1998) metric and that of Throneburg et al. (1994). So for instance IPC factors 2 and 6 are closely related to the LEC and MS factors of Throneburg et al. (1994) and factor 7 is the same as CS. However, there are more factors that make up the IPC metric than Throneburg et al. (1994) include. This makes it particularly useful for cross-language comparisons as a wider range of factors affords more scope for examining differences in occurrence of the factors across languages and how this impacts on stuttering. For example, the below analysis shows dorsal consonants occur more frequently in content words in German, than they do in English. There are two ways such differences in frequency of usage between the languages could affect stuttering: Early babbling experience might affect speech control in the long term. The influence of a factor on stuttering should then be immune from differences in frequency of usage across languages. This would predict stuttering rates would be comparable irrespective of differences in frequency of usage. Alternatively, the more experience a speaker has of a particular factor, the easier it would be to produce. The influence of a factor on stuttering should then be affected by differences in frequency of usage across languages so stuttering rates should be less in a language where a factor occurs frequently compared with one where the factor occurs infrequently.

The contribution of individual phonetic factors on stuttering rate was determined and compared with English. IPC factors were examined in English and German to see whether frequency of occurrence affects stuttering rate or not. The analyses are done separately for function and content words as the work on English indicates only the latter are affected by phonetic difficulty and the same would be expected to apply to German. All analyses were also conducted over age groups as the data reviewed earlier indicate that speakers of different ages deal with phonetic complexity of content words in different ways. In particular, young speakers avoid the difficulty by repeating or hesitating on function words whereas older speakers attempt the content word itself. This would predict that only older speakers should be affected by phonetic complexity.

All analyses examined function and content words separately for speakers who stutter in different age groups. The analyses are reported in five sections. Analysis one examines the overall phonetic complexity as indicated by the IPC score. IPC scores of content words would be expected to be higher than the IPC scores of function words. Analysis two examines overall phonetic complexity across stuttered and fluent words using IPC scores. Stuttered words would be expected to have a higher IPC score than non-stuttered words. Analysis three compares overall phonetic complexity between German and English. German content words would be expected to have higher IPC scores than English ones because there are a lot of compound words that include the different factors in German. Analysis four attempts to identify which individual IPC factors affect stuttering rate in German. Analysis five establishes the frequency of occurrence of the individual IPC factors in German. In this analysis, frequency of occurrence differences between German and English are inspected to see whether or not they affect stuttering rate.

3.2 Method

There is a considerable overlap between the methods used in chapter one and those used in the current chapter. However, since extra child data were available at the time this study was conducted, age group partitioning differs between the studies and the transcription procedure for the extra child data (youngest age group) differs; the details are given here again.

3.2.1 Participants

All speakers who participated had been diagnosed by their speech therapists as people who stutter. None of the speakers displayed unusual phonological processes affecting syllable structure and all stuttering was developmental as reported by their speech therapists. Consent to be recorded was given by each speaker. For child speakers the parent's consent was also obtained. Fifty monolingual German and twenty-six monolingual English speakers participated in the study.

3.2.2 Procedures

3.2.2.1 German speech samples

Speech recordings were made in a variety of settings in Germany. The sources were private, school-based and university-based clinics. There were eleven female and thirty-nine male speakers and ages ranged from two to forty-seven years old. Details are given in Table 3 about each individual German speaker. Information is provided about the location where the recording was made, gender, age, number of words in the sample and stuttering rate (see below for details how this was calculated). The German speakers were partitioned into age groups in different ways for analyses involving German samples alone (analyses one and two) versus comparison between German and English samples (the remaining analyses). The division in all analyses distinguished children under eleven from speakers above eleven. This was based on previous research that indicated that age eleven divides children who stutter predominantly on function words from those who stutter on content words (Howell et a. 1999).

Subject Information					
Subject	Gender	Age	Number of Words	Stuttering Rate %	Recording Place**
Adult Group					
1	F	31:8	210	10.48	
2	F	21:6	151	5.30	
3	F	29:11	135	10.37	
4	F	30:1	210	20.48	
5	M	18:4	153	9.80	
6	M	16:11	173	39.31	
7	M	17:9	208	18.75	
8	M	16:3	256	1.95*[Δ]	
9	M	23:7	330	8.79	
10	M	28:5	85	28.24[Δ]	
11	M	46:5	185	22.70	
12	M	47:1	125	9.60	
13	M	33:10	143	7.69	
14	F	47:3	200	33.50	
15	M	31.6	290	3.81	
Children					
Youngest group (2 years 10 months – 6 years 5 months)					
1	F	3:9	456	10.53	
2	F	5:8	458	1.53*	
3	M	3:5	236	8.90	
4	M	2:10	378	14.55	
5	M	4:10	553	3.98	
6	M	6:2	369	6.23	
7	M	4:4	217	5.53	
8	M	6:0	177	3.95	
9	M	5:3	275	6.91	
10	M	4:7	453	6.62	
11	M	4:7	563	7.64	
All of the youngest age group were recorded in Ulm by Rommel and colleagues					
Age group 2 (6 years and 6 months – 8 years and 11 months)					
1	M	6:6	1288	4.82	U
2	M	7:4	252	10.32	BS
3	M	7:8	120	20	BS
4	F	8:4	252	9.10	BS
5	F	7:6	325	8.92	BS
6	M	7:5	623	7.54	U
7	M	7:9	438	4.34	U
8	M	6:8	751	13.98	W
9	M	7:11	906	9.15	W
10	M	6:7	291	34.83	A
Age Group 3 (9 years – 11 years and 11 months):					
1	M	10:3	461	7.99	A
2	M	9:9	125	15.20[Δ]	BS
3	M	10:8	189	20.11	BS
4	M	10:9	85	42.35[Δ]	BS

5	F	11:11	73	42.47[Δ]	BS
6	M	10:8	89	24.72[Δ]	BS
7	M	9:2	314	26.75	BS
8	M	10:0	116	18.97[Δ]	BS
9	M	11:9	170	15.29	BS
10	F	11:1	70	32.86[Δ]	BS
11	M	11:2	263	10.98	BS
12	M	11:11	506	8.30	BS
13	F	11:5	484	6.00	BS
14	F	10:7	283	12.01	BS
** Recording Places were U=Ulm (Rommel and colleagues), BS=Sprachheilzentrum Bad Salzdetfurth, W=Sprachheilzentrum Werscherberg, A=Aachen University Clinic. Note: The two cases indicated by '*' had quite low stuttering rates. They were excluded from the rate analyses, but their values were used in the analyses referring to structural overall IPC scores and language differences. 'Δ' Indicates those subjects who were excluded from analysis four, see text					

Table 3: Details of speakers' gender (column 2), age (column 3), number of words in sample (column 4), stuttering rate (column 5) and recording location (column 6). Recording location is coded (see bottom of table). See also note regarding two speakers with low stuttering rates and those that were excluded from analysis four (see result section).

3.2.2.2 Grouping by age for analyses involving German speakers alone

For the first two analyses the German speakers were divided into four age groups (one adult and three children groups). The three child groups were 2 years – 6 years 5 months (G1); 6 years 7 months – 8 years 11 months (G2); 9 years 2 months – 11 years 11 months (G3). Eleven children, two girls and nine boys, belonged to G1 (with a mean age of 4 years and 7 months, standard deviation 1 year and 1 month). There were ten children in G2 with two girls and eight boys (mean age of 7 years and 4 months and a standard deviation of 7 months). G3 consisted of fourteen children with four girls and ten boys (mean age of 10 years and 9 months, standard deviation 8 months). There was also an adult group (G4) that consisted of four female and eleven male speakers with a mean age of 29 years and 3 months, standard deviation of 10 years and 9 months.

For these four German age groups, a one-way ANOVA indicated that there was a significant difference in overall stuttering rate (see Figure 9) between the groups ($F(3, 46) = 3.998$, $p<0.05$). Tukey HSD post-hoc follow-up tests indicated that the youngest age group had a significantly lower rate than the 9-11 year old children ($p<0.01$). None of the other differences reached significance. ANCOVAS were used to partial out the effects of different stuttering rates between groups in analyses of the German samples alone in analysis 2.

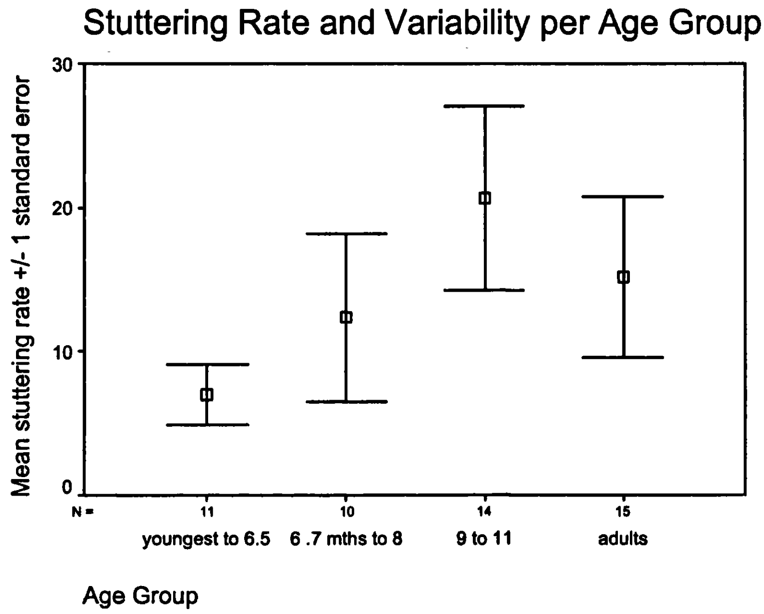


Figure 9: The mean stuttering rate (plus/minus one standard error) per age group. Age groups and number of subjects are indicated along the abscissa. Mean values and standard errors are represented by individual bars.

3.2.2.3 *Grouping by age for analyses involving German and English samples*

The groups were divided differently for cross-language analyses. These groupings allowed comparison with similar studies using English (Howell, Au-Yeung, Yaruss & Eldridge, submitted). There were two age groups - children between the ages of 6 and 11 and adults (the latter was the same as that described in the previous section). In the child group, there were twenty-six children (six girls and twenty boys, mean age of 9 years with a standard deviation of 2 years).

A two-way ANOVA with factors age group (two levels) and language (two levels) showed that there were no significant differences in stuttering rates between these age groups and the two languages. ANOVAS are used when the two languages are compared (as there are no differences in stuttering rates between age groups that need to be partialled out by using stuttering rate as covariate in these analyses).

3.2.2.4 *English sample (for comparison with German)*

The English data were taken from groups 1 and 3 from Howell et al. (submitted). These correspond approximately in age with the child and adult groups indicated in the previous section. There were sixteen participants in the child group aged between 6 and 11 (mean age 8.0 years, standard deviation of 1.0). There were ten adults aged 18+ plus (mean age 26.9 years, standard deviation of 6.2). The analyses in Howell et al. (submitted) indicate that there were no significant differences in stuttering rate between age groups.

3.2.2.5 *Speech Material and Transcription*

For all age groups recordings of speakers in spontaneous speech were used for assessment and analysis. These were a minimum of two minutes in duration and were made in a quiet, relaxed environment. A number of the recordings of German children were taken in sessions with their speech therapist, in speech therapy centres in Bad Salzdetfurth and Werscherberg. Topics that were suggested to the adult speakers and older children were family, friends, favourite films, sports and such like. The speech samples of the fourteen youngest children were collected by Rommel and colleagues in Ulm (Germany) (indicated in Table 3). This consisted of spontaneous speech which was videotaped in individual standardized play situations. These children were taped with their mothers whilst jointly playing with a toy farm. These recordings were on average thirty minutes long. Transcriptions were carried out according to the guidelines of

MacWhinney (1995) using the CLAN/CHILDES analysis system by researchers in Ulm. They were originally orthographically transcribed using a non-standard form of German (the local Swabian dialect). These files were then adapted to conform to the transcriptions of the other age groups by the author. The same events as in chapter two were marked as stuttering, i.e. monosyllabic whole word repetitions, prolongations, repetitions and blocks. All words were also classified as function or content words in type. For the intra-rater consistency analysis, 20% of samples were randomly selected, retranscribed and compared with the original transcription and Cronbach's alpha measure as a measure of consistency. A higher alpha value signifies better consistency. A threshold alpha value of 0.7 indicates an acceptable reliability coefficient (Nunnaly, 1978), though lower values are sometimes used in the literature. Alpha values for both fluency judgments as well as content/function word classification and IPC coding ranged from 0.88 to 0.98 which indicated a high level of consistency. For the English data, a second transcriber re-transcribed eight recordings selected at random to obtain inter-judge reliability measures. Alpha values for fluency judgements and content/function word classification were between .96 and .98, indicating high levels of agreement again.

3.2.2.6 Stuttering rates

Stuttering rates were calculated as the number of words that involved part word and monosyllabic whole word repetitions, prolongations and blocks divided by the total number of words (both stuttered and fluent) and converted to percentages.

3.2.2.7 *Phonetic measures (IPC scores)*

Table 2 in the introduction to this chapter gave a breakdown of the eight IPC factors. In respect to factor 1, 'consonant by place', every dorsal consonant (/k, g, χ, ç, j, ŋ, ʁ, ʔ/) in a word was given one point whereas other consonantal articulation places received no points. For the IPC factor 2, 'consonant by manner', every fricative, affricate and liquid received one point (/f, v, s, z, h, χ, ç, j, ʁ, ʒ, ʃ, tʃ, pf, ts, l/) and no point was given to other consonants. Factor 3, 'singleton consonant by place', assessed inter-syllabic relationships. A point was given to a consonant pair only when, in a '..VC-CV..' structure the syllable coda consonant had a different place classification to the onset consonant of the following syllable(V and C stand for vowel and consonant slots). In factor 4, only rhotic vowels received a point. Whether the word ends in a consonant or vowel was indicated by factor 5 ('word shape'). A word was considered long and received a point when it had three or more syllables (factor 6). For 'contiguous consonants', consonant clusters were given one point irrespective of the syllable or word position they occupied (factor 7). Finally for factor 8, 'cluster by place', one additional point was given if individual consonants in a cluster had different places of articulation.

3.3 Results

3.3.1 Analysis 1: Average IPC scores for function and content words per age group (German)

This analysis tested whether the IPC scores of content words was higher than those of function words. A comparison was made between the IPC scores for function and content words for each of the German age groups. The IPC scores are shown (mean with bars indicating plus/minus one standard error) in Figure 10 for the different age groups (age groups are indicated on the horizontal axis). Mean IPC score are shown separately for function (dotted line) and content (solid line) words for each age group. Analysis of these data was performed using a mixed model ANOVA with word type as a repeated measures factor and age group as an independent groups factor.

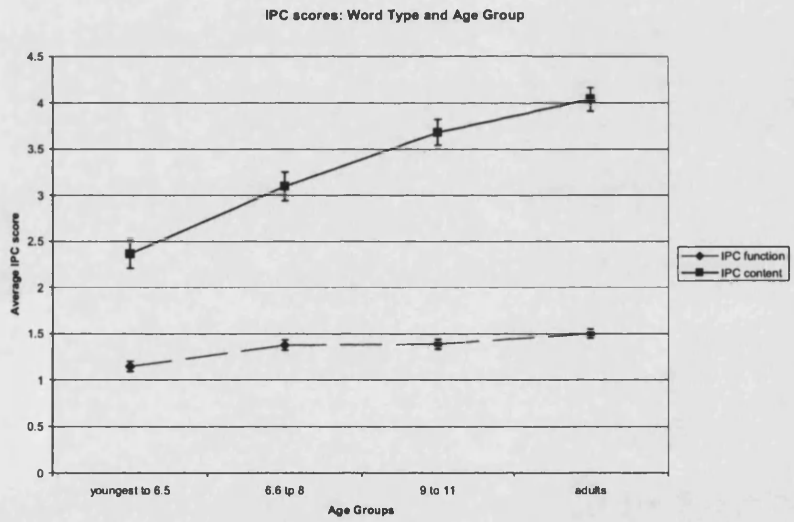


Figure 10: The mean IPC score (plus/minus one standard error) for each age group. Age groups are indicated along the horizontal axis. The solid line refers to content and the dotted line to function words (as indicated in the legend).

Consistent with the prediction for this analysis, the ANOVA showed that there was a main effect for word type with content words having a higher IPC score in all of the age groups ($F(1, 44) = 333.13$, $p<0.001$). There was also a significant interaction between word type and age group ($F(3, 44) = 9.02$, $p<0.001$). This indicated that with increasing age the difference between content and function word IPC score increases. Follow-up paired t-tests on word type showed that in all age groups the content words had a significantly higher IPC score than function words indicating their more complex phonetic structure: (G1, $t(10) = -15.75$, $p<0.001$; G2, $t(9)=-10.56$, $p<0.001$; G3, $t(13)=-15.08$, $p< 0.001$; G4, $t(14) = -16.79$, $p<0.001$). There was also a significant main effect for age group ($F(3, 45) = 28.65$, $p<0.001$). Tukey HSD post hoc follow-up tests on age groups revealed that the youngest age group used words with significantly less complex phonetic structure compared to all other groups (all $p<0.01$) as indicated by lower IPC scores. IPC scores of G2 were significantly lower than the adult group, G4 ($p< 0.001$), whereas the IPC scores of G3 and the adult age group did not differ significantly. The interaction effect between age group and word type was looked at by establishing IPC scores of content words over age groups and IPC scores of function words over age groups. The difference between G1 and all other groups was significant for content words (Tukey tests, all $p<0.01$), as well as between G2 and the two older age groups (Tukey tests, both $p<0.05$). G3 and G4 did not differ significantly. In all cases, the younger age group had lower IPC scores than the older age group. There was also an effect for function words (which is not very apparent from inspection of Figure 1) with the youngest age group using significantly 'easier' function words compared to all other ages (Tukey tests, all $p < 0.05$), but no differences between the other age groups for function words. This is a particularly interesting point since this age group showed comparatively high occurrence of stuttering events on function words compared to the other age groups. To summarise this analysis, the main prediction, that the IPC scores of content words should be higher than those of function words, was confirmed.

3.3.2 Analysis 2: Mean IPC scores for stuttered and fluent words (German)

This analysis examined whether stuttered word have higher IPC scores than non-stuttered words for each word class. IPC scores for stuttered and fluent words for each word type and each age group are shown as histograms in Figure 11. Plus/minus one standard error is shown around the mean. A three-way mixed model ANCOVA was carried out with fluency (two levels, fluent and stuttered) and word type (two levels again, function and content) as within groups factors and age group as the independent groups factor. Overall individual stuttering rate was used as a covariate to partial out he significant group differences as outlined in the method section.

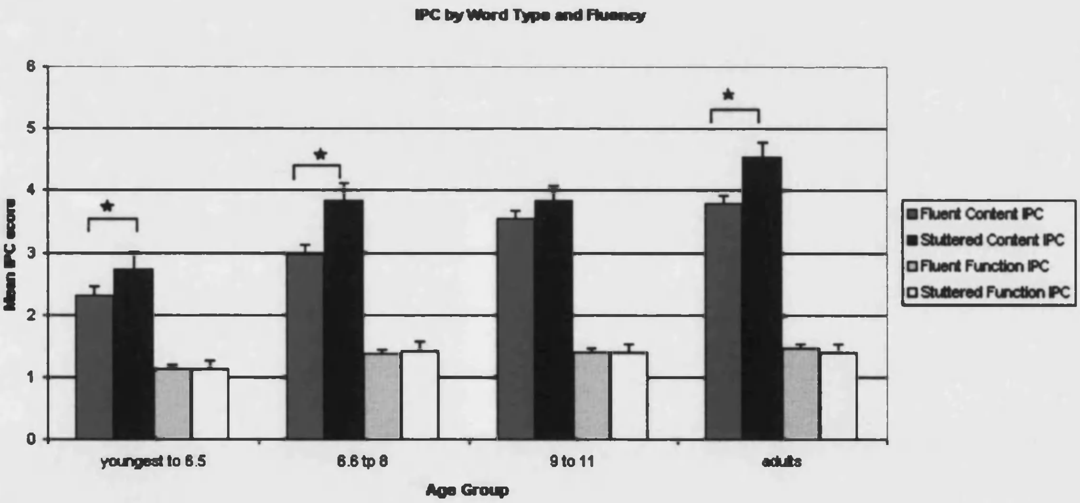


Figure 11: Histogram showing the mean IPC score (plus/minus one standard error) according to age group and grammatical word type. Age groups are indicated along the abscissa. Bars that are shaded differently refer to fluent content, stuttered content, fluent function and stuttered function words as indicated in the legend.

The two within groups main effects were significant (word type, $\underline{F}(1, 44) = 414.17$, $\underline{p} < 0.001$; fluency, $\underline{F}(1, 44) = 13.23$, $\underline{p} < 0.01$) as well as their interaction ($\underline{F}(1, 44) = 414.17$ $\underline{p} < 0.001$). Word type by age group ($\underline{F}(3, 44) = 7.60$, $\underline{p} < 0.001$) was also significant. The word type by age group interaction indicated that the gap between the phonetic difficulty of function and content words widened with age, i.e. content words became phonetically more complex with age whereas function words stayed roughly the same (the pattern seen in Figure 10). The word type by fluency interaction indicated that the IPC scores for stuttered words were higher but only when the words were content in type. In an equivalent analysis of IPC scores in English (Howell et al., 2004), similar non-significant results for function words were reported. In summary, content words that were stuttered had higher IPC scores than content words that were not stuttered for adults but there were no effects for function words and for either word type in younger age groups.

3.3.3 Analysis 3: Phonetic complexity of German compared to English

This analysis examined whether IPC scores of German content words were higher than their English counterparts (hypothesised to be the case because German has a lot of compound words). The second set of age group categories (see Method) were used for this and subsequent analyses so that they can be compared against the corresponding age-group data from English speakers. For each language, only two age groups were used. The child group for ages from 6 to 11 years and the adult group for speakers over 18 years. The results are summarised in Figure 12. Age group and fluency are indicated along the abscissa. Grammatical word class and language are represented by different symbols (the codes for these are given in the legend). A two within-subjects factors and two between-subjects factors mixed model ANOVA was carried out with word type and fluency of word as the repeated measures factors and language of the speaker and age group as the independent groups factors.

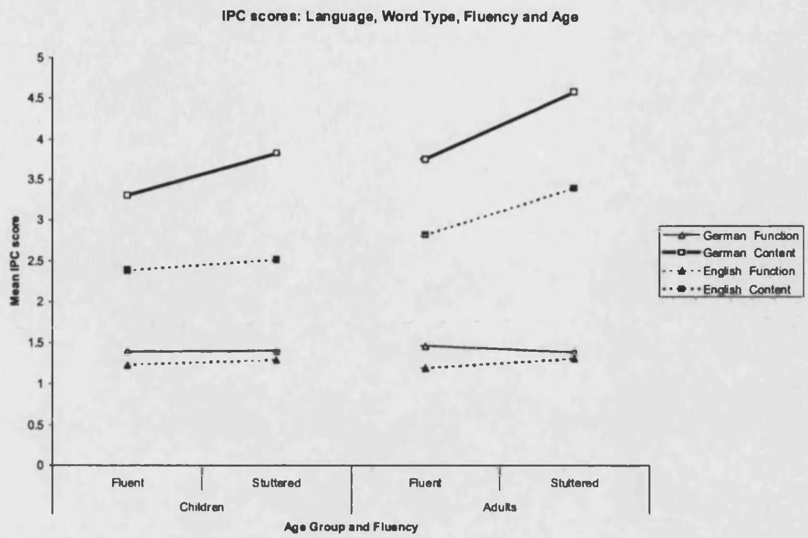


Figure 12: Line graph representing mean IPC scores according to age group and fluency (as indicated along the x-axis) separated into grammatical word class and language (as described in the legend). Unfilled markers and solid lines always refer to German cases whereas dotted lines and filled markers refer to English cases. Note that the ages are divided along different lines compared to previous figure.

All of the main effects were significant (word type: $F(1, 61)=506.90$, $p<0.001$; fluency: $F(1, 61)=18.20$, $p<0.001$; age group: $F(1, 61)=14.17$, $p<0.001$; and language: $F(1, 61)=52.25$, $p<0.001$). Ignoring the interactions for the moment, the main effects indicate that content words had significantly higher IPC scores than function words, stuttered words were associated with higher IPC scores than function words, adults used phonetically more complex words than children and the German words had higher IPC scores than the English words. The significant two-way interaction between age group and word type ($F(1, 61)=12.34$, $p=0.001$) highlighted the fact that function words had similar IPC scores in the two age groups whereas the content words of adults were significantly higher than the children. Word type also interacted with language ($F(1, 61)=27.42$, $p<0.001$). Inspection of the data showed that in both languages content words had higher IPC scores compared to function words, but the difference between function and content word IPC scores was larger in German than English. This interaction supports the prediction made at the outset of this analysis. The two-way interaction between word type and fluency was significant ($F(1, 61)=19.90$, $p<0.001$). This indicated that English and German stuttered words had higher IPC scores only when the words were content in type. The same two factors just described were involved in a three-way interaction with language ($F(1, 61)= 3.997$, $p=0.050$) indicating that the size of the difference between IPC scores of fluent and stuttered words differed between the languages. This was due to the larger difference in IPC scores between fluent and stuttered content words in German as compared to English. In summary, German content words have higher phonetic difficulty compared to their English counterparts.

3.3.4 Analysis 4: What IPC factors affect stuttering in German?

This analysis attempted to identify which individual IPC factors affected stuttering rate in German. Initially all IPC factors were included. Then, an attempt was made to identify which of the eight IPC factors were successful at predicting stuttering rate followed by an assessment of the order of importance of the successful factors.

3.3.4.1 Relationship between IPC scores and stuttering rate when all IPC factors were included

The steps in the analysis procedure to establish how IPC factors operate in German is the same as that employed by Howell et al. (submitted) for English. The three steps in the analysis (referred to above) are:

a) Words were sorted into different categories according to their IPC score.

b) Stuttering rate was calculated by dividing the number of stuttered words in each IPC category by the total number of words (both stuttered and fluent) in the same category.

c) An analysis of variance (ANOVA) was then calculated to find differences in stuttering rate over IPC scores. This analysis treated the stuttering rate for each category as the dependent variable and the IPC score category as the independent variable. If IPC score is a good predictor of stuttering frequency then there should also be an increasing trend in the data. Thus linear trend analysis was also carried out following the ANOVA. Analyses were carried out separately on content and function words and age groups.

Speakers excluded. A number of the speakers had to be excluded from this analysis since preliminary inspection revealed that some of the shorter samples had too few words to calculate a stuttering rate estimate for all the selected IPC categories. Six children and two adults were dropped for this reason. This left twenty children and thirteen adults that were analysed as indicated above (these are indicated in Table 3).

3.3.4.2 Analysis of children aged 6-11 and adults

Function words

Step one: Sorting words by IPC score. The analysis on the function word class revealed that for both age groups there were not enough instances of function words with scores higher than two and these were dropped. This reflects (again) the fact that function words have a simple phonetic structure.

Step two: Stuttering rate for different IPC scores. Stuttering rates for the three selected IPC-score categories were obtained: These were function words with an IPC score of 0, function words with an IPC score of 1, and function words with scores greater than 1.

Step three: ANOVA and linear trend analysis of IPC scores against stuttering rate. Figure 13 shows mean stuttering rate for function words over IPC score categories (the corresponding data for content words are also shown in this figure). Solid lines are from adults and dotted lines from children as indicated in the legend. Unfilled circles indicate results for function words. There appears to be no increase between IPC-score category and stuttering rate for the function words for either age group. The ANOVA confirmed that there was no difference in stuttering rate for the different IPC scores for either age group. Thus, this analysis indicated that IPC score was not a good predictor of stuttering rate for function words as found previously for English.

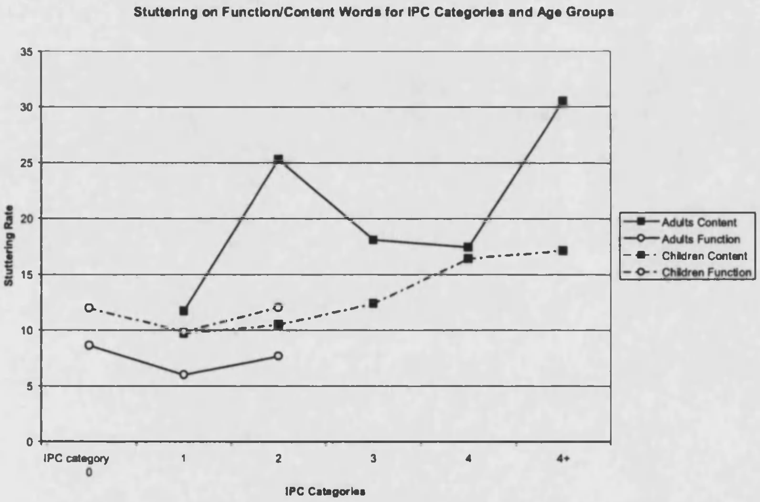


Figure 13: Mean stuttering rate for function and content words with IPC score sums 0-4+ as indicated along the abscissa. For content words too few cases had IPC scores of 0 which was the reason for dropping this category from the analysis. Solid lines refer to adults and dotted lines to children as indicated in the legend. Function words are given unfilled circles as markers whereas content words have filled square markers.

Content words

<u>Step one: Sorting words by IPC score</u>. There were very few cases where content words had an IPC score of zero. There was sufficient data for analysis of IPC categories 1 to 4 and 4+, i.e. five categories, for both age groups.

<u>Step two: Stuttering rate for different IPC scores.</u> The small number of cases where IPC score was zero were rarely stuttered (overall only 1% of the content words with IPC scores of zero were stuttered). This meant that for the IPC zero category on content words, stuttering rates could not be calculated and, thus, this category was dropped from the analysis.

<u>Step three: ANOVA and linear trend analysis of IPC scores against stuttering rate.</u> The content words are indicated by the filled symbols in Figure 4 (otherwise the data are represented in the manner indicated for function words). Initially an overall mixed model ANOVA was carried out with age as an independent groups factor and IPC category as a repeated measure. The results indicated that there was a significant main effect for IPC category ($F(4, 124)=7.31$, $p<0.001$) for the combined data for the two age groups. There was also a significant linear trend for this factor $(1, 31)=17.51$, $p<0.001$). Figure 14 is a graphical representation of this trend. This shows that as words become phonetically more difficult, stuttering rate initially increased, then plateaued and there was another sharp increase when words have a number of different characteristics combined (i.e. the increase in stuttering frequency of words with more than 4 points on the IPC measure).

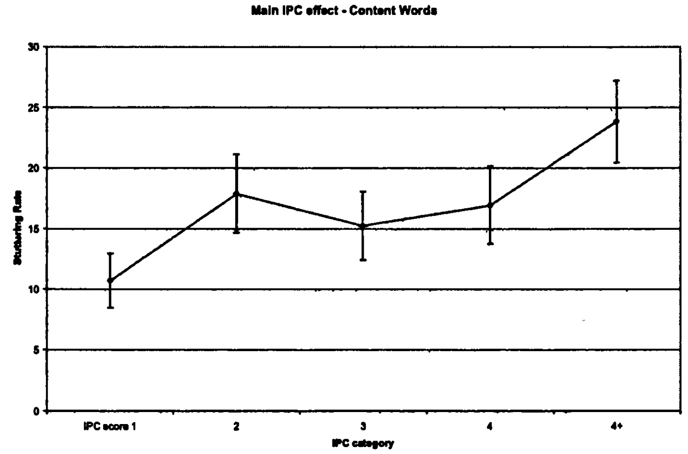


Figure 14: Mean stuttering rate (plus/minus one standard error) for content words with IPC score sums 1, 2, 3, 4 and 4+ (see legend along the horizontal axis).

It was expected that the adult age group would show a clearer example of this pattern whereas the children would exhibit a less steep increase in stuttering rate. This, however, was not the case as revealed by inspection of Figure 13. However, a significant interaction occurred between IPC score category and age group ($F(4, 124)=3.30$, $p=0.013$). For the children (Figure 13) the increase appeared to be less steep than the overall effect when both age groups were combined (Figure 14) whereas this steady increase was not as clearly the case for the adults. The adult pattern (Figure 13) showed a rise from 1 to 2 IPC points but then a fall to 4 points with a sharp increase in stuttering on words with more than 4 IPC points. This is an exaggerated version of the pattern observed for the combined groups in Figure 14, i.e. an initial rise and medium difficult words being then stuttered at more or less the same rate with an additional rise when words become phonetically a lot more complex. This may be due to some IPC factors not predicting stuttering rate as found in English (Howell et al., submitted). Later analyses examine this possibility.

Relationship between IPC scores and stuttering rate with selected IPC factors

When considering the impact of individual IPC factors on stuttering rate, each of the age groups was analysed individually. The aim of these further analyses was to find out whether individual factors increased or decreased this relationship between stuttering rate and sum of IPC factors. Since some words score on more than one IPC factor in spontaneously produced speech samples, it was not possible to examine words with selected individual factors. Thus individual factors were dropped when obtaining the IPC score and then steps a-c were performed and the results compared to the analysis containing all factors. The rationale for the analysis is that if individual factors predict stuttering rate, then when these are dropped there should be less of a relationship (indicated by fewer significant differences between IPC score categories) between stuttering rate and IPC category. Since factor 4 never occurs in German it was dropped from the analysis.

For the children, the factors that showed a decrease in the number of significant pairwise differences when they were dropped were IPC factors 1, 2, 3, 5 and 6, so these are important. When an ANOVA was run on IPC scores obtained with these factors alone, a significant effect was found for words with different IPC summed scores ($F(4, 76)=4.98$, $p<0.01$). Words with IPC scores of one and two had a significantly lower stuttering rate than the rate on words with scores of four, or more than four, IPC points. Further analysis showed that the linear trend across summed IPC scores was significant ($F(1, 19)=14.47$, $p<0.01$).

The analysis for the adult age group gave similar results though the linear relationship was not as strong. For this reason, not only the number of significant differences was taken into consideration but also whether the relationship that results was monotonically increasing, as well as whether the F value in the analysis increased.

It was assumed that a factor with a big impact would, when taken out, would tend to flatten the IPC-score stuttering rate relationship and decrease the F value.

This analysis indicated that IPC factor 1, 2, 5, 6, 7 and 8 ought to be retained. An ANOVA with just these factors showed a significant linear trend ($F(1, 12)=6.64$, $p<0.05$). The relationship between the selected factors for each age group and stuttering rate using the selected IPC factors are shown in Figure 15 (adults solid line, using 1, 2, 5, 6, 7 and 8; children dotted line, using 1,2,3,5 and 6).

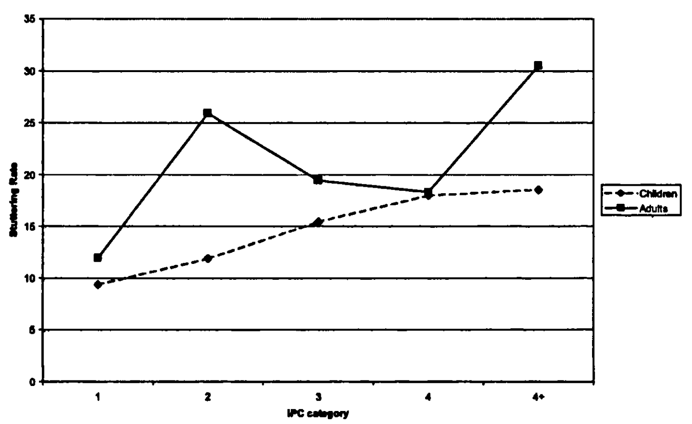


Figure 15: Mean IPC stuttering rates for words with IPC score sums 1-4+ according to each age group (children – dotted line, adults – solid line). Here IPC sums are based on the factors that were analysed as having a high impact on stuttering rates (analysis 4: adults IPC factors 1,2,5,6,7 and 8; children 1,2,3,5 and 6).

<u>Order of importance of selected IPC scores and stuttering rate</u>

To analyse the rank order of impact for the selected factors in each group, individual ANOVAs were carried out which started with the set of IPC factors that the previous analyses indicated were important and dropped one IPC factor from this set at a time. The results were then examined to see how adversely the relationship of stuttering rate with IPC score was affected (an indication of the relative importance of that factor). For the group of children over the age of 6 years old the rank ordering this indicated was (from lowest to highest impact): 3, 5, 6, 1 and 2. In the adult group the order was 8, 7, 5, 1, 6 and 2.

3.3.5 Analysis 5: Frequency of IPC factors and relationship to their impact on stuttering rate

This analysis established the frequency of occurrence of the individual IPC factors in German. These were then compared with English to see whether or not frequency of occurrence of the IPC factors in the languages determines which of them affects stuttering rate. For each age group, frequency of occurrence of each IPC factor was calculated (as the percentage of words showing that particular characteristic). This was done separately for function, content and all words for each age group. Table 4 gives the means and standard deviations of frequency of occurrence in percentages for each IPC factor for each language and age group.

Table 4a:

German									
Age	Word	Factor1	Factor2	Factor3	Factor4	Factor5	Factor6	Factor7	Factor8
G1	Content	40.82% (4.71%)	61.15% (3.31%)	3.37% (2.60%)	0.00% (0.00%)	50.00% (6.41%)	6.34% (2.24%)	21.72% (3.32%)	11.54% (3.01%)
	Function	3.39% (2.49%)	42.28% (9.20%)	0.15% (0.41%)	0.00% (0.00%)	54.41% (10.21%)	0.00% (0.00%)	4.30% (1.91%)	0.30% (0.85%)
	All words	27.08% (3.52%)	54.22% (3.86%)	2.19% (1.62%)	0.00% (0.00%)	51.62% (6.85%)	3.98% (1.55%)	15.30% (2.13%)	7.41% (2.09%)
G2	Content	44.47% (10.29%)	71.93% (6.15%)	6.17% (4.63%)	0.00% (0.00%)	60.17% (8.57%)	11.25% (7.43%)	32.27% (13.26%)	18.98% (8.24%)
	Function	11.18% (8.12%)	49.48% (5.00%)	0.18% (0.27%)	0.00% (0.00%)	61.98% (6.07%)	0.00% (0.00%)	8.28% (2.79%)	1.30% (1.40%)
	All words	27.40% (5.22%)	58.78% (3.41%)	3.07% (1.58%)	0.00% (0.00%)	61.61% (5.68%)	5.33% (2.37%)	19.94% (4.03%)	9.97% (2.37%)
G3	Content	55.68% (9.49%)	78.15% (7.55%)	10.48% (3.84%)	0.00% (0.00%)	65.28% (7.62%)	26.11% (9.22%)	38.54% (8.18%)	19.63% (6.69%)
	Function	18.45% (6.52%)	56.19% (4.79%)	0.14% (0.38%)	0.00% (0.00%)	64.39% (6.91%)	0.00% (0.00%)	7.52% (3.96%)	1.44% (1.54%)
	All words	34.31% (7.61%)	64.86% (4.83%)	4.84% (1.63%)	0.00% (0.00%)	64.35% (7.93%)	11.65% (3.98%)	20.45% (4.97%)	9.29% (3.19%)
All	Content	46.99%	70.41%	6.67%	0.00%	58.48%	14.56%	30.84%	16.72%

		(10.25%)	(7.48%)	(4.63%)	(0.00%)	(8.79%)	(10.39%)	(10.86%)	(7.59%)
	Function	11.01% (8.19%)	49.32% (6.33%)	0.16% 0.32%)	0.00% (0.00%)	60.26% (7.08%)	0.00% (0.00%)	6.70% (3.40%)	1.01% (1.54%)
	All words	29.36% (6.47%)	59.46% (4.69%)	3.26% (1.82%)	0.00% (0.00%)	58.52% (6.170%)	6.84% (4.34%)	18.52% (4.26%)	8.78% (2.66%)

Table 4b:

English									
Age	Word	Factor1	Factor2	Factor3	Factor4	Factor5	Factor6	Factor7	Factor8
G1	Content	25.50% (5.50%)	65.09% (6.61%)	3.04% (1.62%)	0.46% (0.78%)	68.45% (8.30%)	4.92% (1.50%)	30.48% (6.37%)	16.40% (4.30%)
	Function	4.89% (4.07%)	40.74% (10.32%)	0.72% (0.86%)	0.56% (0.74%)	56.07% (4.70%)	0.71% (0.90%)	14.41% (5.85%)	1.37% (1.18%)
	All words	15.04% (3.03%)	52.60% (5.55%)	1.88% (1.09%)	0.51% (0.49%)	62.07% (4.55%)	2.78% (0.84%)	22.01% (4.29%)	8.59% (1.95%)
G2	Content	31.30% (5.42%)	66.59% (7.78%)	4.74% (2.27%)	0.25% (0.64%)	72.44% (8.69%)	7.76% (3.12%)	31.74% (7.45%)	17.00% (4.72%)
	Function	2.98% (2.11%)	40.95% (8.15%)	0.46% (0.63%)	0.52% (1.17%)	56.91% (6.74%)	0.50% (0.99%)	15.53% (4.54%)	0.99% (0.84%)
	All words	16.35% (3.52%)	53.20% (5.13%)	2.45% (1.04%)	0.42% (0.76%)	64.22% (5.23%)	3.97% (1.64%)	23.07% (4.62%)	8.54% (2.48%)
G3	Content	27.14% (3.44%)	73.00% (4.97%)	4.48% (1.58%)	1.08% (0.51%)	70.20% (3.48%)	12.62% (4.15%)	34.53% (3.89%)	16.41% (2.81%)
	Function	3.40% (0.84%)	41.90% (3.57%)	0.53% (0.42%)	1.28% (0.71%)	55.86% (3.88%)	0.72% (0.38%)	8.29% (2.64%)	1.90% (1.31%)
	All words	14.01% (1.21%)	55.93% (3.24%)	2.30% (0.68%)	1.20% (0.54%)	62.31% (2.70%)	6.11% (1.98%)	20.06% (2.69%)	8.40% (1.73%)
All	Content	28.10% (5.59%)	67.54% (7.32%)	4.03% (2.01%)	0.53% (0.73%)	70.39% (7.65%)	7.84% (4.13%)	31.92% (6.42%)	16.63% (4.14%)
	Function	3.81% (2.93%)	41.10% (8.14%)	0.58% (0.64%)	0.72% (0.96%)	56.34% (5.31%)	0.63% (0.84%)	13.38% (5.49%)	1.35% (1.13%)
	All words	15.29% (3.01%)	53.62% (5.00%)	2.20% (1.00%)	0.64% (0.68%)	62.94% (4.50%)	4.03% (1.94%)	21.95% (4.18%)	8.52% (2.08%)

Tables 4a and b: Frequency of occurrence of each IPC score as a mean percentage of words with one or more instances of the specified factors (standard deviations are given in brackets underneath). The percentages are divided into grammatical word class and age group. Tables are separate according to language (Table 4a German and 4b English).

Differences in frequency of usage across corresponding age groups for German and English were first established using t-tests (with Bonferroni correction for the number of tests carried out, i.e. results were assessed using a significance level of 0.05/8 in each age group). These were then used to see whether there was any relationship between frequency of usage and stuttering rate across the languages. The results of these analyses are summarised in Table 5. The top section of this table indicates the factors that affected stuttering rate for the two German age groups and corresponding data for English are given (the latter are taken from Howell et al., submitted). The bottom part of Table 4 indicates where there were significant differences (using the t-test scores mentioned at the beginning of this paragraph) between the frequency of occurrence of different IPC factors for all words (i.e. function and content types) between the two languages for the two age groups.

Factors	**1**	**2**	**3**	**5**	**6**	**7**	**8**	
Patterns								
German G2	1	2	3	5	6			
German G3	1	2		5	6	7	8	
English G2	1	2			6	7		
English G3	1	2			6	7		
All words	1	2	3	5	6	7	8	
Base Rates								
G2								
German	27.40	58.78	3.07	0	61.61	5.33	19.94	9.97
English	16.35	53.20	2.45	0.42	64.22	3.97	23.07	8.54
Mean diff.	11.06	5.58	0.62	-0.42	-2.60	1.36	-3.13	1.43
P value	<0.001	<0.001	0.18	<0.05	0.17	0.06	<0.05	0.09
G3								
German	34.31	64.86	4.84	0	64.35	11.65	20.45	9.29
English	14.01	55.93	2.30	1.20	62.31	6.11	20.06	8.40
Mean diff.	20.30	8.93	2.54	-1.20	2.03	5.54	0.39	0.89
P value	<0.001	<0.001	<0.001	<0.001	0.40	<0.001	0.83	0.43

Table 5: The top section indicates which factors improved the stuttering rate-IPC score relationship. All factors that appear in either language or either age group are listed in the top row and age groups in the left column. Cells where the number is included indicate whether the analysis indicates that factor is important for each language and age group. The bottom section indicates the frequency of occurrence of the difficult attribute of each factor as percentage occurrence out of all words for corresponding age groups in German and English and the mean difference (diff.) in frequency of usage. The p value for the language comparisons are given in the last line of each age group.

For the children, of the seven IPC factors that appear in German (excluding factor 4), five characteristics occurred more often in German compared to English. Two occurred significantly more frequently in German (factors 1 and 2), a further one approached significance (factor 6) and the remaining two of the five factors were not statistically significant (factors 3 and 8). Of the two factors that were more frequent in English compared to German (words ending in consonants, factor 5, and words containing consonant clusters, factor 7), only consonant clusters appeared significantly more often in English compared to German.

The adult data showed that all of the seven IPC factors were more frequent in German than in English (of these four were significantly more frequent – factors 1, 2, 3 and 6), and the remaining factors (5, 7 and 8) did not differ significantly between the languages. Comparison with the child data, indicate that factor 3 (difference in frequency of variegated consonants) and factor 6 (word length) became significantly different between the languages as age increased. Consonant clusters (factor 7) were used significantly more often by English children but there was no difference in usage for adult users of the two languages. Factor 5 (words ending in consonants) changed from being not significantly more frequent in English versus German to being non-significantly more frequent in German. Factor 8 (heterorganic clusters) remained more or less constant across languages and age groups.

Those factors that affect both German and English for both age groups irrespective of frequency of occurrence were removed. The logic for this followed that outlined in Howell et al. (2004), who argued that a case can be made that factors that affect stuttering rates across languages and age groups may indicate that those IPC factors have some general status (across two Germanic language forms here) possibly deriving from early babbling experience (hypothesis one given in the introduction).

Three factors affect both age groups in both languages (1, 2 and 6) and a third factor affects both English age groups and one of the German age groups (factor 7). These reflect consonant by place (1) consonant by manner (2), word length (6) and contiguous clusters (7). In addition, these four factors involved 6 out of 7 of the significant differences in frequency of usage between languages for the two age groups. Thus these factors appear to operate in the two Germanic languages irrespective of their occurrence in the language (once more indicating the general status of the way these factors operate in these languages).

The one remaining significant difference in frequency of usage was factor 3 - singleton consonant by place (same place easy, variegated place hard) - for German adults. This factor selectively affected German children when there was no significant difference in frequency of usage between languages. The factor was used significantly more frequently by German adults compared with English adults and could account for why this factor did not affect stuttering rate (as it had done for the German children). This factor may support hypothesis two given in the introduction (i.e. characteristics that occur with low frequency should be difficult and, therefore, more likely to affect stuttering rate in that language, German here, than the comparison language, English).

Factor 5 (word shape) refers to word endings and on the face if it, as most stuttering occurs at onset, would seem unlikely to affect stuttering. Also pointing to this conclusion, it occurs with high frequency (over 60% in all samples indicated in Table 5) and this should make this factor easy to deal with. Consistent with this, this factor does not affect children or adults who speak English. However, it has an impact on both German age groups. Inspection of the frequency data for factor 5 (word shape) shows, however, that for content words this factor is much more frequent in English than German (70.39% English and 58.48% German – see Table 4). In this respect the difference in frequency could be the characteristic in content words that causes the higher impact of this factor on stuttering rates in German, again consistent with the second hypothesis. Finally, there is no apparent reason in frequency of usage of the different factors across ages and languages either when looking at all words or function and content words separately as to why factor 8 should lead to a significant increase in stuttering rate in German adults.

3.4 Discussion

In analysis one it was shown that for German there was an interaction between word type and age group, indicating that with increasing age the difference in IPC scores between function and content word gap widened (see Figure 10).

In analysis two, word fluency analysed with IPC scores in German revealed all main effects (word type, fluency and age group) were significant demonstrating higher phonetic complexity of content words, and stuttered words being more complex than fluently produced words (see Figure 11). The main effect of age group indicated that the two youngest age groups were using less phonetically complex words compared to adults. These effects, however, were mediated by two significant interactions. There was a type by age group interaction which on closer inspection revealed that the gap between function and content word complexity widened with age, i.e. adults used phonetically more complex content words than children whereas the complexity of function words was more or less constant over age groups. The other significant interaction (word type by fluency) indicated that phonetic complexity was only a significant predictor of stuttering for content words whereas function words complexity was constant regardless of whether the word was fluent or stuttered.

Overall the comparison of phonetic complexity between the two languages (analysis three) showed a number of similarities and differences. Figure 12 highlights a number of these. The main detail to single out is that when the IPC scores were compared across languages, there was a larger gap between function and content word IPC scores in German compared to English. This was supported statistically by the three-way interaction between fluency, word class and language indicating that the gap between stuttered and fluent content words was larger in German compared to the same gap in English whereas the gaps were roughly constant for function words.

In analysis four, the impact of individual factors was determined by eliminating factors and seeing how this affected the relationship between stuttering rate and IPC score. For the older children (over 6 years old) and the adults significant linear trends were established with all 8 factors for content words (but not function words), indicating that for this word class as IPC score increased, stuttering rates increased. This relationship was used as a baseline to establish the impact of the other factors once they were taken out. The same technique was later used to rank order the remaining factors. In this way it was established that for the older children the factors that made words difficult were 3, 5, 6, 1 and 2 (when the word contained variegated consonants, ended with a consonant, was long, contained dorsal consonants and had fricatives, affricates or liquid manners) and the factors that made words difficult for adults were 8, 7, 5, 1, 6 and 2 (words with heterorganic consonants, consonant clusters, those ending in consonants, dorsal consonants, long words and words containing fricatives, affricates or liquid manners).

In the final analysis (analysis five), the impact of IPC factors was examined in relation to differences in their frequency of occurrence between German and English. First those factors that affect both languages were located irrespective of their frequency of occurrence. The factors in this set were consonant by place (1) consonant by manner (2), word length (6) and contiguous clusters (7). Interestingly 2, 6 and 7 are the factors that are closely related to Throneburg et al.'s LEC, MS and CS indices. Consonant by place (dorsals hard, other places easy) is an additional factor that operates generally in these languages that was not identified by Throneburg et al. (1994). The only remaining IPC factors that affect any age group of speakers of either language were 3, 5 and 8.

There was a significant difference in frequency of usage of factor 3, singleton consonant by place (same place easy, variegated place hard) for German adults. This factor selectively affected stuttering rate of German children (though there was no significant difference in frequency of usage with the corresponding age group of English speakers). If it is assumed that the children's data show that German children experience difficulties on words that include this factor, then the change in usage with age (older German speakers use consonants with variegated place much more frequently than any of the other speakers) would make these consonants easier for these speakers and account for why it does not then affect stuttering rate. This interpretation implies that experience affects operation of this factor.

Factor 5 (word shape – words ending in consonants) had an impact on stuttering rates for both the children and the adults in German but not in English. This is a factor which appears to be equally frequent in both English and German (around 62% - see Table 4). Closer inspection of this factor showed that in English content words this factor is very frequent (appearing in about 70% of those) whereas only about half (58%) of German content words contain this factor. Therefore the lower frequency in German might be a reason for its higher impact on stuttering rates in German than English based on the same argument as used for factor three. Once again this would suggest experience affects operation of this factor. Even though it might seem counterintuitive that word endings might effect stuttering frequency which would take place at word beginnings there is evidence in linguistic literature that might explain such a process. For instance a process called anticipatory coarticulation describes the way phonemes are shaped and articulated depending on certain sounds that appear later in a word. Since English adults have more frequent experience with this factor exposure to it might have made it easier and thus does not increase stuttering rates.

Though frequency of usage could account for the emergence of significant differences for IPC factors 3 and 5, as just indicated, no reason was identified why factor 8 should emerge as a factor having an impact on stuttering rate for adult German speakers.

To summarise the results in analysis five, there seems to be evidence for manner, consonant strings and long words causing difficulty for all ages for users of both languages. There is also a suggestion that dorsal consonants are difficult for both ages of both language groups. There is some evidence that suggests that repeating consonants in adjacent syllables, and words that end with a consonant string (hard dimensions of factors 3 and 5) have an impact on German because these hard constructions occur less frequently in this language for certain age groups between German and English. The main feature of the data is, however, that four of the factors (1, 2, 6 and 7) operate similarly between these languages. This would be expected given the linguistic similarity between these languages, and in this light the similarities in the empirical findings are reassuring. Future work ought to examine languages that have dissimilar structure (such as Spanish and English, or Spanish and German).

There are a number of limitations about the IPC scheme and its usefulness in relation to stuttering rates. First of all since stuttering occurs mainly on the onsets of words (e.g. Brown, 1938, Wingate 1988) it is difficult to see how the factor 'word shape' (words ending in consonants) could influence stuttering rates. However, this factor was shown to have an influence on stuttering rates in Spanish as well as German. It is possible that word-final factors do play a role in planning and retrieval time of words. As such delays in these two processes could lead to poor preparation of word onsets. This is not ruled out by the EXPLAN theory outlined in the introduction. Another issue regarding the general IPC scoring scheme is that the factors are not sufficiently independent of one another. There are many factors concerned with consonantal difficulties that are not mutually exclusive. This makes the interpretation of results more difficult since the impact of some of the factors would not be as high because of close correlations with other factors.

On the whole the results confirm a more complex structure of the spontaneously used German content words. This also impacted, as predicted in the introduction, on the stuttering rate with a steeper increase in IPC scores from fluent to stuttered words in German compared to English.

Whereas in chapter two the focus was on gross phonetic characterisations (e.g. starts with either vowel or consonant), the examination of phonetic factors was inspected in closer detail in this chapter. The EXPLAN model highlights the content word as the difficult nucleus cause that causes stuttering. This would explain why we see the effect of phonetic complexity on the content but not on function words.

In the next chapter the content word was analysed in the context of the surrounding function words (using the concept of the PW). This examined the relationship of stuttering on function and content words in more detail.

4 Stuttering on function and content words across age groups of German speakers who stutter[4]

"[On being asked if his stammer was a problem] No, Sir, because I have time to think before I speak, and don't ask impertinent questions."
Erasmus Darwin, late 18th century, quoted by Francis Darwin in 'Reminiscences of My Father's Everyday Life', in his edition of Charles Darwin's *Autobiography*

4.1 Introduction

The majority of previous work on stuttered speech has analysed words out of context (and the work in Chapters Two and Three is in line with this tradition). This can be seen by reconsidering Brown's (1945) work. Three of his four factors pertain to properties of isolated words – starting phone, grammatical class, word length. Even his fourth factor (sentence position) is scored by an index that uses words (an indication of position based on an indication of ordinal position in a sentence). The exception to the context-free (isolated word) analyses is in work relating stuttering to syntactic properties. Researchers have highlighted the links between grammar and stuttering (see for instance Bernstein Ratner, 1997; but also Howell & Au-Yeung, 2002) and this chapter sets out to investigate the question whether syntactic units or PWs are the appropriate units for analysis. Background to the PW as a unit in phonology is given here in some detail for the purpose of highlighting the different

---

[4] The following is a version, with additional background information on the PW, of an article that is published as Dworzynski, Howell, Au-Yeung & Rommel (2004) in the *Journal of Multilingual Communication Disorders*.

approaches regarding the controversial issue of PW length that have been taken in the literature.

4.1.1 Background to the concept of the phonological word (PW)

Au-Yeung et al. (1998) and Howell, Au-Yeung and Sackin (1999) suggested that repetition and hesitation on function words could be a delaying strategy that allows speakers to wait until the full speech plan for a later content word is available (which is also formalised in the EXPLAN model). The importance of this explanation is that the context proposed extends beyond a single word, because it postulates that the point where the disfluency occurs is in anticipation of a problem on a later word. This wider context is given by the phonological word that can be larger than a single word.

First of all an outline of the phonological word as a linguistic concept is provided. In Prosodic Phonology the rules of universal grammar specify a series of prosodic units which are arranged in a hierarchy (Booij, 1983; Nespor & Vogel, 1986; Selkirk, 1980a; 1980b). This hierarchy has six basic constituents (that are recognised by almost all phoneticians) and looks as follows:

phonological utterance (U)
↓
intonational phrase (IP)
↓
phonological phrase (ϕ)
↓
phonological word (ω)
↓
foot (F)
↓
syllable (σ)

An example to illustrate this hierarchy the first three stages (the PW being described below) from the top down in an example such as 'He loves all autumn flowers which are different from spring flowers' would look as follows:

'[He loves] $_{\phi}$ [all autumn flowers] $_{\phi}$]$_{IP}$ [which are different] $_{\phi}$ [from spring flowers] $_{\phi}$]$_{IP}$]$_{U}$'.

Although the name of the PW concept refers to the fact that it is roughly the size of a grammatical word, Hall (Hall, 1999b) pointed out that a number of authors have repeatedly stressed that the phonological word is distinct from this grammatical unit. As such they do not necessarily refer to semantic factors as suggested by Brown (1945). In the history of phonology the term *phonological word* (PW) was first coined by Dixon (1977a; 1977b). It has since been employed by numerous other authors (e.g. Booij, 1983; Hannahs, 1995a; 1995b; Nespor & Vogel, 1986) and has also been referred to as the 'prosodic word', a name that was first used by Selkirk (1978; 1980b; 1984). A full history of the PW is beyond the scope of the present review – it can be found, with a comprehensive survey of the literature, in Smith (1996).

Nespor and Vogel (1986) define the PW in terms of the domain of lexical stress assignment or other low-level phonological rules. Function words do not obey the same kinds of phonological generalisations as content words (such as the stressless 'weak' and stressed 'strong' forms). *Well-formedness conditions* have also been found that only affect content (not function) words (Hall, 1999). These are rules that govern how words can be formed in individual languages. For instance, in English only words belonging to the function category can begin with the voiced dental fricative /ð/ e. g. *their, then, they* vs. *thick, thing, throw.* Due to these differences several authors, most notably Selkirk (1984; 1996)for English and Booij

(1996) for Dutch, have concluded that function words constitute 'special cases' in that they are not independent PWs. A quote by Selkirk (1984, p.336) illustrates this point:

> "It is also characteristic of function words that they may exhibit an extremely close phonological connection, or juncture, with an adjacent word – usually to the word that follows but sometimes to one that precedes."

Shortly after she developed this argument further with regard to the role of function words (p.337):

> "We claim that these and other ways in which function words are not treated like 'real' words in the grammar are to be attributed to a single principle, the *Principle of the Categorical Invisibility of Function Words* (PCI), which says (essentially) that rules making crucial appeal to the syntactic category of the constituent to which they apply are blind to the presence of function word constituents."

More recently, Hall (1999a) and Vigário (1999)have examined the phonology of function words in German and European Portuguese respectively. Both found that the weak forms were exempt from the *bimoraic minimum* principle (a constraint that governs the minimal size of words) that holds for PWs in these languages. As such monosyllabic content words in these languages end in either a long vowel, diphthong, or a sequence of short vowel plus single consonant, i.e. a sequence of two moras – e.g. the German words 'flott' (brisk), 'rau' (rough), and 'roh' (raw). Both researchers concluded that weak forms of function words are not independent PWs. These weak forms of function words belong to the class of *clitics*. For example the words 'will it' and 'see you' could be phonologically likened to 'billet' and the name 'Mia'. The current debate in phonology concerns the question of how the clitics are represented in prosodic phonology. In that respect a content word is generally termed a *host* (also called *head*), a preceding function word *proclitic*, and a function word positioned after the host an *enclitic*. There is considerable disagreement among theorists on the prosodic representation of host + enclitic, proclitic + host or even proclitic + host + enclitic sequences. A number of authors have opted to classify at least some of these sequences as single PWs (Booij, 1996; Hall, 1999; Selkirk, 1984, 1996; Vigário, 1999).

Examples for English are:

proclitic + host as in (a fence)$_\omega$ or (can pile)$_\omega$ and the host + enclitic (give it)$_\omega$

For German:

host + enclitic as in (kommt es)$_\omega$ 'is it coming' or (spiel es)$_\omega$ 'play it'

Other authors (Nespor and Vogel, 1986) view the PW as a prosodic constituent that is 'word – internal'; i.e. cannot span two or more grammatical words. As such they would treat affixes as independent PWs with the word stem as another PW:

prefix + stem $\longrightarrow$ (prefix)$_\omega$ (stem)$_\omega$ $\longrightarrow$ (im)$_\omega$ (polite)$_\omega$

In a later published article, Selkirk (1996) argued that a function word may either be prosodised as a single phonological word, or as one of three different types of prosodic clitic, which are not phonological words in itself (free, internal or affixal clitic – meaning that they have to be attached to a host). In her earlier work function words were all assumed to be clitics.

In psycholinguistics Levelt (1989; 1992) has used the PW concept to explain phonological encoding in speech production. He asserted that, in the phonological encoding process, a speaker generates a skeleton or frame for the purpose of creating a pronounceable metrical pattern for the utterance as a whole. These frames are for PWs, which are distinct from lexical units. He views the PW as the domain of syllabification and of word stress assignment (Levelt, 1992). By his definition, in English the PW is composed of a head word with its affixes and clitics. He even suggests that there may be two or more head words in a single PW (as in some compound words). This would make it particularly interesting for German where long compound words are a common feature. With regard to function words he stated (Levelt, 289 p. 299):

> "... function words (auxiliaries, pronouns, determiners and so on) tend to be destressed as if they were affixes. Phonologically, they are not real words at all."

To illustrate the point the following example was given (Levelt, 1992): *Black Bear gave it him.* This he divided into two PWs: (*Black Bear*)$_\omega$ which is a compound with trochaic word stress, and (*gavitim*)$_\omega$ which has a head word (*gave*) and two dependent words (*it* and *him*) that are cliticised to the head word. He thus concluded that in the phonological encoding stage of speech production lexical boundaries have lost their significance. Moreover he employed the PW to explain common speech error such as Spoonerisms – *peel like flaying* instead of *feel like playing*. In accounts prior to his work the error was seen to involve an exchange of consonants across an intermediate word. He suggested that the speaker was probably just planning two PWs *(feelike)* $_{\omega}$ and *(playing)*. Due to the fact that word onsets are more vulnerable to error than other syllable onsets (Shattuck-Hufnagel, 1992), it would also explain why the former error is more likely than *leel fike playing*. The /l/ in *like* is PW internal, i.e. it is not word initial and hence less vulnerable.

The significance of this concept for research into stuttering is that, in connected speech, a single constituent (what Levelt called frame) can be wider than one lexical item, i.e. the PW. This can provide the context in which to understand fluency breakdown. To date Howell and Au-Yeung in the UCL speech group have adopted the following interpretation of the PW. Based on the notion that function words are not PWs themselves (as pointed out for the stressless forms by Levelt, 1989; Selkirk, 1984 and others), they are viewed as prefixes or suffixes to a neighboring stressed content word (Au-Yeung et al., 1998; Howell et al., 1999). They go on to define a PW on lexical criteria as consisting of a single content word (C) plus adjacent function words (F) leading to the general form $[F_nCF_m]$, where *n* and *m* are integers greater than or equal to zero. Consider for instance the utterance "I look after her cats". There are two function words between the two content words 'look' and 'cats'. Au-Yeung et al. (1998) and Howell et al. (1999) have developed Selkirk's (1984) semantic sense unit rules to establish which function words are associated with each content word. These are rules, Selkirk (1984) had proposed, that define which words are semantically related in intonational phrases (i.e. two stages up in the prosodic hierarchy):

"Two constituents C_i, C_j form a sense unit if (a) or (b) is true of the semantic interpretation of the sentence:

a. C_i modifies C_j (a head)

b. C_i is an argument of C_j (a head)."

These rules were extended by Au-Yeung et al. (1998):

c. both C_i and C_j modify C_k (a head)

d. both C_i and C_j are arguments of C_k (a head).

These are explained in detailed steps in the method section. In the above example it is important to find out whether one or both of the function words are prefixes to the content word 'cats', or suffixes of 'look'. If the former is the case, a stutter on the function word could delay production of the content word. In the latter case, if they belong to 'look', they could not be used to delay production of either 'look' – because they follow the content word – or 'cats', because they would not be part of the same PW. Applying the rules to the above example, 'after' is part of the same PW that includes 'look' (it is a phrasal verb and forms a special meaning with 'after') and, because it is a suffix, it cannot delay the production of 'look'. The function word 'her' is a prefix of 'cats' and thus can be used to delay production of the content word – as in fact can 'I' as a prefix to 'look'. More detailed examples of how spontaneous speech is divided into phonological words are given in the method section of this chapter.

In many ways the interpretation of the PW by the Howell and colleagues can be seen as a taking a 'middle of the road' approach. They have used the concept in a wider sense than the *word internal* definition by Nespor and Vogel (1986). It can be viewed, however, as somewhat narrower than Levelt's (1989,1992) interpretation, because they only allow one content word as a nucleus, whereas he suggested that in some circumstances it could have two (as in the *Black Bear* example).

4.1.2 The PW as a tool in stuttering research

There is a tradition of work on fluent speech, in contrast to the prevailing theme of work in stuttering, that examines the effects of context on fluency control. (The work on fluent speakers' development is relevant, as the reader recalls that, according to the continuity hypothesis – see chapter one section three -, research on fluent speech development may hold a clue to the speech patterns seen in young speakers who stutter). So, for instance, it has been observed that young fluent speakers often repeat the function word that precedes the content word a result of what Clark and Clark (1977) describe as delaying the attempt of the subsequent word (the difficult content word in this case). The notion of context is here to be understood as the view that the disfluency on one word might indicate an upcoming problem on a later word. Clark and Clark (1977) suggested that the hesitations and repetitions that frequently occur in fluent speakers' speech indicate that the speech plan for a later word is not ready for execution.

Au-Yeung, Howell & Pilgrim (1998) noted that word repetitions in well-known corpora like that of Maclay and Osgood (1959), occur predominantly on function words. Au-Yeung et al. (1998) took the 'delaying' explanation and hypothesized that while young children who stutter delay by repeating function words (albeit at a higher rate than fluent children) that eventually led to the EXPLAN proposal (Howell, 2002; Howell & Au-Yeung, 2002). These types of stuttering, i.e. repetitions of monosyllabic words or part words would be so-called stalling disfluencies, whereas the type referred to as advancing disfluencies are particularly prolongations and syllable repetitions – see chapter one. To account for these advancing stutterings, the EXPLAN model assumes that older speakers do not delay but attempt the difficult content word that follows. This results in disfluencies like "split, s.s.split or sssplit" because the content word is not fully prepared, and the disfluency is on the part that is ready (part of the onset).

To investigate the proposal that the locus where disfluency is observed in continuous utterances changes with age, Au-Yeung et al. (1998) used phonological words (PWs) as a unit for segmental analysis. In this chapter an empirical approach was taken (see chapter one for a detailed review of the theoretical background to the PW). To briefly recap, within a PW, a content word serves as the nucleus and function words can (optionally) precede and follow the content word. This allowed Au-Yeung et al. (1998) to examine the prediction that follows from their extension to the delaying hypothesis, that only initial function words are repeated as these are the only function words that can serve the delaying role. This prediction was confirmed. They also examined stuttering rate over different positions in PW, separately for function and content words. Function words had higher stuttering rates the earlier they occurred in PW, this being most apparent for the younger speakers. This is consistent with a delaying role of initial function words. Content words showed no differential stuttering rate across PW positions for any age group. This is also consistent with the view that content words are the locus that is difficult and that this difficulty is experienced whatever the position of the content word.

In a further study, Howell et al. (1999) examined the proposal that speakers change the way they deal with the locus of disfluency over ages. Disfluency rate was computed separately for function and content words for five different age groups. It was found that disfluency rate was highest at the youngest age on function words and that the disfluency rate decreased with age. The high rate at the youngest age indicates the widespread use of function word repetition in this age group, and the decrease over age shows that this happens less often as speakers get older. The pattern was the opposite with content words (disfluency rate on these words was low initially and increased with age). The complementary pattern appears to indicate that as young speakers repeat function words less frequently as they get older, content word disfluencies emerge (this pattern is referred to as an exchange relation to capture the reciprocity in how speakers tackle the locus of disfluency).

Spanish has stressed function words, so stress can be dissociated to some extent from lexical word class. Nevertheless, an exchange relation occurred when analysis was based on function and content words (Howell, submitted). PW can be segmented in Spanish using stressed words (either function or content) as nuclei. When PW segmentation is made according to word type and, separately, according to stress patterns, the data can be allocated into disjoint sets where the segmentation differs for the two types of PW. The stress-based segmentation then has stressed function words as the nucleus for stuttering, and the lexical-class based segmentation has unstressed content words as nuclei. If stress is paramount, the first of these methods alone should show an exchange relation, whereas if lexical status is paramount, the second of these methods alone should show an exchange relation. Howell (submitted) has shown that both segmentation methods lead to an exchange relation. This suggests that both lexical status and stress can be a focus of stuttering in PW or that some factor that is common to these segmentation methods mediates these patterns.

4.1.3 Aims of the current study

This chapter mainly addresses whether general characteristics found in English also occur in German using lexically-defined PWs. While German and English are in some ways similar languages, there are some structural differences that may be relevant to the level of stuttering. The question arises as to whether these should be treated as one locus of difficulty, or more than one, when they comprise more than one content word. There is also the issue of whether the components of compound words are treated as separate words (as argued here) or not. As compound words (as a class) occur with high frequency in German, and given the arguments for their division (as was outlined in depth in chapter the background to the PW) into their constituents, it was decided that the context in which to analyse stuttering in German PW should divide compound words into their individual components. This means that one orthographic item can appear in more than one PW in German. The effect of not segmenting compound words might make the PW longer (for German compared to English), an effect that probably disappears once the individual units are divided.

The following study examined the pattern of disfluencies within PWs in German. Comparison of PW length for two different segmentation methods (based on whether or not compound words are treated as single lexical items) is reported, and it is found that when compound words are decomposed the length is similar to English. This was carried out for the adult group only since they were the ones with the highest percentage of content words. The reason for this was to choose a method appropriate for the German PW and which would give a comparable PW size to the English segmentation method. The PW where compound words are decomposed are used to test predictions that German is producing a similar pattern to English this segmentation method is then used for all age groups. Analysis one concerns the position of the function word in respect to the content word within a PW. Similar to the English PW analysis, it is predicted that function words that precede content words would be more likely to be disfluent than those that are positioned after a content word. In analysis two it is predicted that a serial position effect should occur for function words, but not for content words. Analysis three is the exchange analysis. It is hypothesised that there should be a decrease of function word disfluency with age and the reverse for content words. The findings, and their implications for current models of fluency failure, are considered in the discussion.

4.2 Method

4.2.1 Participants

All of the subjects included in this study (43 English subjects and 50 German speakers – see subject information table in chapter three for the details of the German speakers) were diagnosed as exhibiting stuttering behaviour. Speakers of both language groups were recorded in conversational speech with a researcher or speech pathologist. Details of the English speech samples are given in Au-Yeung et al., (1998) and Howell et al., (1999). The speakers of each language group were divided into different age groups (the details of which are given below). Speakers in the youngest German age group were each recorded, in a standard play situation, with their individual care giver, in most cases their mother.

The German speakers and age groups were the same as those in the last chapter for analyses one and two. It was shown in the last chapter that there was a significant difference in stuttering rate for the youngest age group. This meant that for all the analyses in the result section the overall stuttering rate for each individual was taken out as a covariate. This means that the effect of differences in overall disfluency rate across participants (which would include group differences) is dealt with in the analyses by partialling out each participant's overall disfluency rate. The output is then giving the adjusted rates for each factor after the effects of the covariate (here the speaker's overall disfluency rate) have been removed.

These German speakers were compared to the English disfluent speakers used in the Howell et al. (1999) paper, with the exception of the teenage group since no comparable age group was included in the German data. The following groups were used. Since the English age groups and number of subjects in each group are different to those of the last chapter details are given here:

4.2.1.1 *English speakers (adults).*

These were twelve male adults aged between 20 and 40 years (the mean age being 28 years and 4 months).

4.2.1.2 *English speakers (children).*

These were divided into three different subgroups: 2-6 year-olds, 7-9 year-olds and 10-12 year olds. Four boys and two girls were in the first age group (mean age of 4 years and 2 months). Fifteen children (11 boys and four girls) made up the second group, which had a mean age of 7 years and 3 months. There were ten children aged 10-12 (8 boys and 2 girls with a mean age of 11 years and 4 months).

4.2.2 Speech Material

Speech material and transcription procedure was the same as that described in the previous chapter – please see the methods of chapter three.

4.2.3 Segmentation into Phonological Words

For the segmentation the PW used here consisted of a single content word (C) plus adjacent function words (F), leading to the general form $[F_nCF_m]$ where n and m are integers greater than or equal to zero. First the content words that constitute the nuclei of the PWs were located. It was necessary to determine the position of each individual function word in the PW. In order to gauge whether a function word was a prefix to the subsequent, or a suffix to the preceding, content word. Selkirk (1984) developed rules that define which constituents of an intonational phrase form a so-called 'sense unit', i.e. are semantically related. Au-Yeung and Howell (1998) and Au-Yeung et al. (1998) extended these in order to apply them for the segmentation of PWs. As highlighted in the introduction the following rules were used:

"Two constituents C_i, C_j form a sense unit if (a) or (b) is true of the semantic interpretation of the sentence:

a. C_i modifies C_j (a head)

b. C_i is an argument of C_j (a head)."

These rules were extended by Au-Yeung et al. (1998):

c. both C_i and C_j modify C_k (a head)

d. both C_i and C_j are arguments of C_k (a head).

The first two rules (Selkirk's original ones) can produce both segmentations into an intonational phrase and a PW. They are therefore given precedence over the two additional rules (Au-Yeung et al. 1998) that deal with cases for PW segmentations that do not conform to rules a and b. These rules are illustrated, for a selected utterance from the samples used in the current analyses. The utterance used is:

And he said to her that he would get on the next boat.

First the content words are marked as the nuclei of the PWs:

And he [said] to her that he would [get] on the [next] [boat.]

In the next step the status of the function words is determined, according to the rules given above. The first nucleus or head is the verb 'said', which has two arguments - 'he' (the subject of the sentence) and 'her' (the object). Thus the first PW segmentation is:

[And he said to her] that he would [get] on the [next] [boat.]

The next PW starts with 'that', which is the beginning of an embedded clause and is grouped as such. A further PW is the prepositional phrase 'on the next' in which 'on the' is linked indirectly to 'next' via 'boat' (the last PW). The two additional rules c. and d. govern the segmentation of such indirect links, in which a PW can be formed via a third part outside the PW. As such 'boat' is the outside part, 'on the' is an argument and 'next' a modifier:

[And he said to her] [that he would get] [on the next] [boat.]

The same segmentation process is used for the German samples, for example:

Und dann fange ich an zu schminken und zu frisieren...
'And then I start to do the make-up and the hair styles'

Step one:

Und dann [fange] ich an zu [schminken] und zu [frisieren...]

The head of the first PW is the verb 'fange' which is a word stem, having a separable prefix 'an' and subject 'ich'. Thus the first PW segmentation would look as follows:

[Und dann fange ich an] zu [schminken] und zu [frisieren...]

The other two PW are verb phrases:

[Und dann fange ich an] [zu schminken] [und zu frisieren...]

The individual PWs were then coded as to the fluency, number of words, position of content word, and in case of disfluencies, location of the disfluency and whether the disfluent word was a function or content word. Filled pauses as well as 'er' or 'um', were excluded from the word count (e.g. 'she erm she said' would be counted as two words). In the case of compound words in German these were divided into individual segments each as their own PW. An example of that would be a word like 'Obststand' (fruit stall) which would be classified as two PWs ("Obst" and "stand" respectively – i.e. phonologically /o:p/ and /ʃtɑnt/).

4.2.4 Reliability Measures

The measure was the same as described in the previous chapters. For the re-transcribed data, alpha values for both fluency judgments as well as content/function word classification ranged from 0.88 to 0.98%. Additionally for the segmentation marking of PWs the alpha values ranged from 0.78 to 0.95 %.

4.3 Results

4.3.1 Analysis 1: Structure of PW in German Compared to English

For the German adult age group all the PWs were segmented orthographically, each content word being classed as a nucleus of the PW regardless whether it was a compound noun or not – for details see the method section. In the second segmentation method compound nouns were then divided into the individual word segments. An overall percentage of the number of compound words was then calculated by dividing the number of PWs that included a compound word by the overall number of all PWs. Compound words were found in 14.2 % of all PWs. One structural aspect that would be affected by such frequent compound words would be the length of a PW. It was found that the variability in the German data was significantly higher than in the English (Levene (2)=7.24 $p<0.05$) – see error bars in the graph below Figure 16. The larger variability could have been due to the fact that there were three more adults in the German adult group than the English one. Thus two non-parametric Mann Whitney U tests were used to analyse the difference, comparing the English PW method with each of the German segmentation methods. This revealed that when compared individually to the English group, only the segmentation that treated compound words as single units was significantly longer than the English (PW with compounds treated as single units ($U(15, 12)=31$ corrected for ties $z=-2.88$, $p<0.01$; PW with compounds treated as their components $U(15,12)=57$, corrected for ties $z=1.61$, $p> 0.10$ respectively). As it is desirable to have equivalent length PW for the two languages (because word length itself affects stuttering rate), together with the other justifications for treating compounds as multiple nuclei (introduction), meant that the PW method where compounds were divided was chosen for the analyses below.

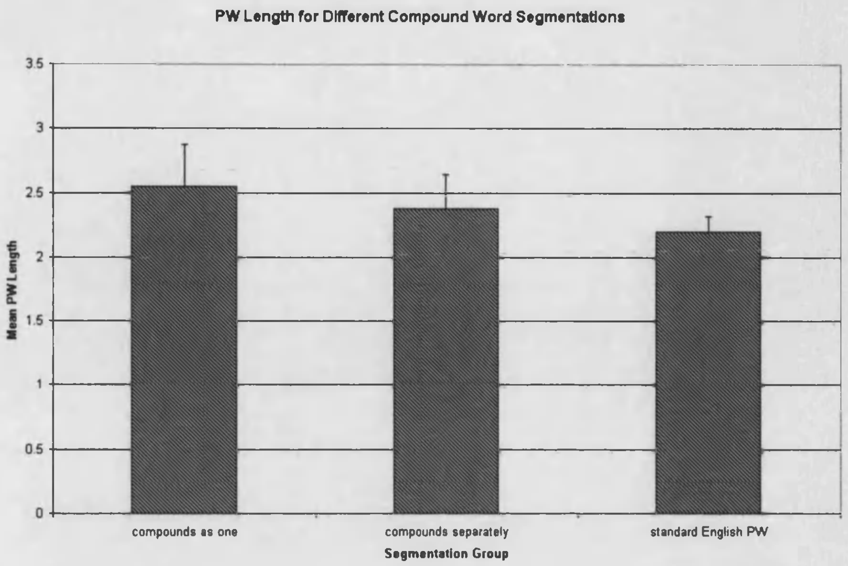


Figure 16: Histogram showing the mean length of PW for the two segmentation methods in German (the left and middle bar) and the standard English method (the bar on the right). Variability is indicated by standard error bars.

Another structural aspect that might be affected by the number of compound words is the proportion of post-content function words within each PW. It was found that this proportion was slightly higher in the two German segmentation methods. However this difference was not significant ($\underline{F}(2, 39)<1$ – in this case Levene's test for equality of variance was not significant).

4.3.2 Analysis 2: Position of Function Word in Relation to the Content Word in PWs

Function words that were positioned after the content word within a PW were significantly less likely to be spoken disfluently than those that were before the content word. The proportions of PWs with stuttering on post-content function words were below 5 percent for the four age groups (four $\underline{t}$ tests were carried out – all $\underline{p}<0.05$ with Bonferroni correction). The left-hand side of Figure 17 shows stuttering rate of function words for different age groups separately for pre-content and post-content positions.

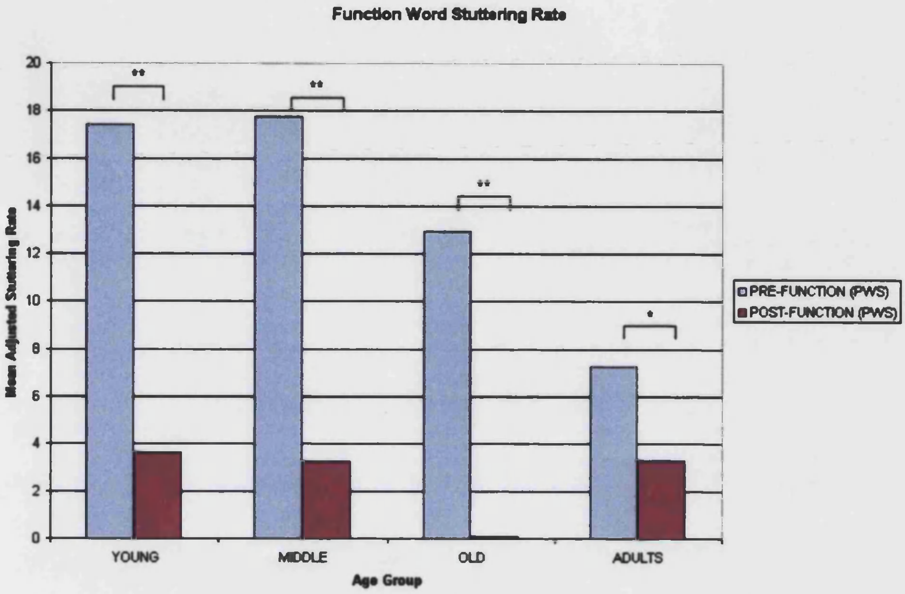


Figure 17: Figure a shows pre- and post-content, function word disfluencies in PWs across age groups. This graph and all the graphs below show an adjusted stuttering rate which is the percentage disfluencies once overall individual stuttering rate is controlled for.

4.3.3 Analysis 3: Serial Position of Function and Content Words in PWs

For the serial position analysis a mixed model ANCOVA was carried out with word type (function/content) position within the PW (first, second or third; there were few instances of stuttering in positions four onwards) as the repeated measures factors and age group (young, middle, old, adults) as the between subjects factor. The main effect for word type was not significant. The main effect for position was significant ($F(2, 70)=3.21$, $p<0.05$). The first position had the highest likelihood of being stuttered, followed by the second, and then the third, as seen in Figure 18.

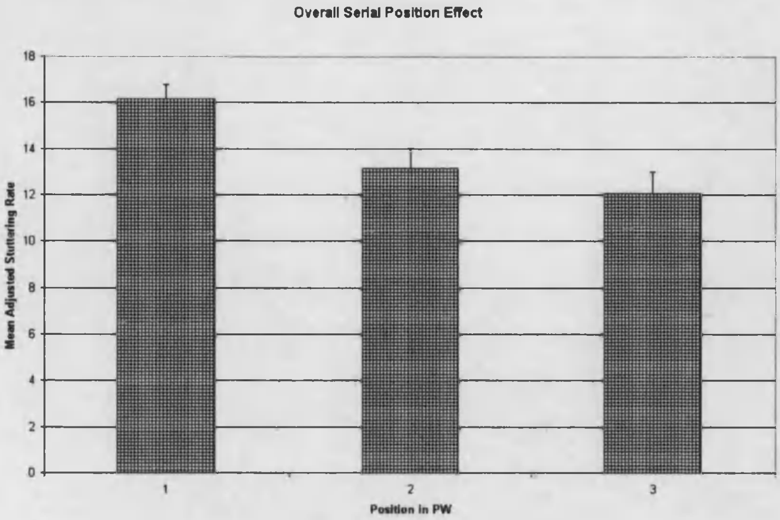


Figure 18: The histogram shows the effect of serial position on stuttering rate. Each bar represents the mean stuttering rate (variability is indicated by standard error bars) for the position as indicated on the x-axis.

In addition to the main effect of word position, there were interactions both with word type ($\underline{F}(2, 70)=5.08$, $\underline{p}<0.01$) and age group ($\underline{F}(6, 70)=3.12$, $\underline{p}<0.01$) but the three-way interaction between all these factors was not significant. (The remaining two-way interaction, the age group by word type interaction ($\underline{F}(3, 35)=8.46$ $\underline{p}<0.01$) was also significant in this analysis. (This reflects the exchange pattern and a non-positional analysis on this, is the subject of analyses three). The word type by position interaction showed that whereas the serial position effect mentioned earlier (as main effect) occurred for function words, content words were stuttered at more or less the same rate in any of the three positions. This is shown in Figure 19, and a similar effect has been observed for English (Au-Yeung et al., 1998) and Spanish (Au-Yeung et al., 2003).

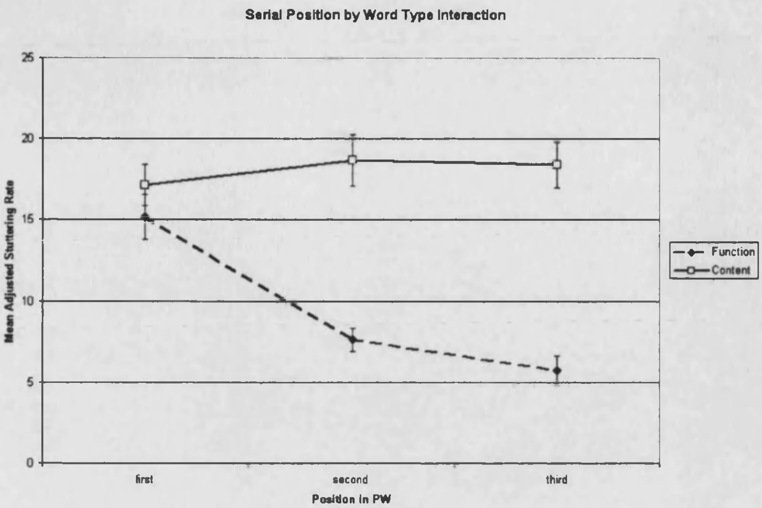


Figure 19: The position of the word within a given PW is indicated along the x-axis. Values along the y-axis represent percentage stuttering rate which has been adjusted by taking individual disfluency rates as a covariate across age groups. The two lines indicate different word types with content words being represented by the solid line and function words by the dashed line (see legend). Standard error bars indicate variability around each mean.

Figure 20 shows the word position by age group interaction. Age group is along the abscissa and word position is the parameter of the connected points (the three serial positions can be identified from the inset). The youngest children showed a clear trend of stuttering rate decreasing from first to third serial position. In middle and older child groups stuttering rate in the first serial position was highest, with virtually no difference between second and third serial positions. By adulthood, there was a non-significant reversal, with the third serial position having highest stuttering rate and first position having lowest stuttering rate. Simple interaction analysis on this two-way interaction revealed that the position effect was only present in the youngest age group. (The mean differences for both first – third ($p<0.002$) and second – third ($p<0.02$) positions were significant and the first – second difference just missed significance (p=0.063), adjusted for multiple comparisons using Bonferroni correction.) Another feature to note is that the stuttering rate in third position increased over age groups. The serial position effects for youngest children are clearest, and this is consistent with a major role of word repetition in young children. The progressive loss of differential stuttering rate between position three and other positions is the clearest sign that the serial position effect disappeared over age groups.

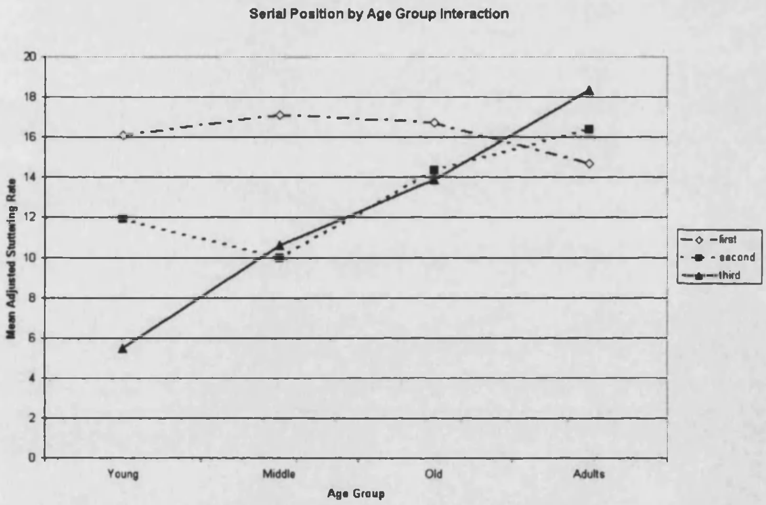


Figure 20: Along the x-axis the different age groups are indicated. The y values represent stuttering rates that have been adjusted by partialling out individual stuttering rates. Different lines indicate first, second and third position within a PW (see legend).

4.3.4 Analysis 4: Exchange Analysis

As reported previously for the English data the following pattern was observed in respect of function and content word disfluency across age groups. In the Howell et al. (1999) analysis a main effect for word type was reported which was not present when the teenage group was excluded in the current investigation (p=0.161). The interaction between word type and age group was significant across the four age groups ($F(3, 38)=2.92$ $p<0.05$). Both of these tests took overall PW stuttering rate as a covariate between the age groups. The results of this reanalysis are given in Figure 21.

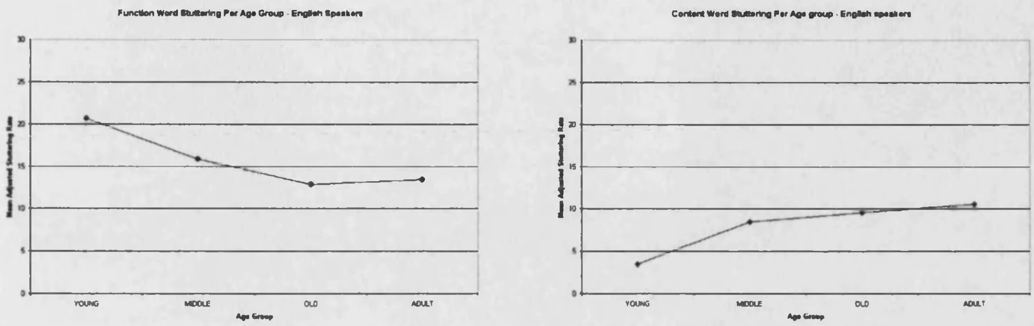


Figure 21 a and b: The mean disfluency rate of function and content words (on the left and right respectively) for English subjects in different age groups (YOUNG = 2-6 years, MIDDLE=7-9 years OLD=10-12 years ADULTS = >18 years) – this data is a subset and re-analysis of the Au-Yeung et al. (1998) sample.

The same analysis as that on English was conducted for German. The main effect of Word type was not significant. The interaction between Age Group and Word type (content and overall function words) that reflects the exchange, was significant ($F(3, 41)=4.698$, $p<0.01$). Both of the word types show a significant linear trend when analysed individually (content: $F(3, 45)=6.205$ $p<0.001$ pre-content function words: $F(3, 45)=7.374$ $p<0.001$). Figure 22 shows that the function word disfluencies across age groups decrease, whilst disfluency rate for content words (right-hand side) increases.

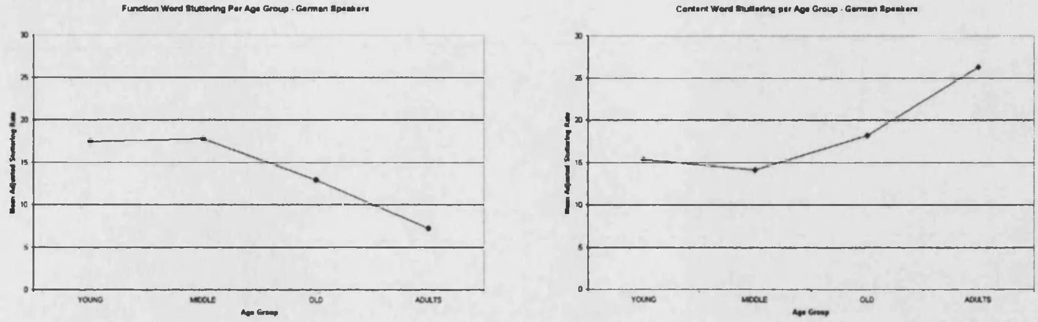


Figure 22 a and b: The mean disfluency rate of function and content words (on the left and right respectively) for German subjects in different age groups (YOUNG = 2-6 years, MIDDLE=7-8 years OLD=9-11 years ADULTS = >18 years).

4.4 Discussion

It was first shown that the segmentation method where compounds were divided yielded most appropriate and language comparable PWs for German.

There are three main findings on the pattern of stuttering in PW in German, all of which correspond with those in English (Au-Yeung et al., 1998; Howell et al., 1999) and Spanish (Au-Yeung et al., 2003) First, function (but not content) words that occupy an early position in a PW have higher rates of disfluency than those that occur later in a PW, this being most apparent for the youngest speakers. Second, function words that precede the content word in a PW have higher rates of disfluency than those that follow the content word. Third, young speakers exhibit high rates of disfluency on function words, but this drops off with age and, correspondingly, disfluency rate on content words increases. The first two positional findings refer to overall stuttering rates within any PW regardless of utterance position. It would be interesting to investigate further whether there are also positional and length effects of PWs in the wider context of utterance position.

Differences in the processing of grammatical word classes are now increasingly found both for the activation of their neural substrates using EEG studies (for example see Bastiaansen et al., 2002; Brown et al., 1999; Osterhout et al., 2002), in terms of lexical access (Segalowitz & Lane, 2000) and their role in cross linguistic examination of speech errors (Wells-Jensen, 2000). Additionally it was reported that infants prefer open-class to closed-class words in preferential listening tasks (Shi & Werker, 2001). The function/content word distinction has also been investigated in an EEG study with individuals who stutter (Weber-Fox, 2001). Her results indicated that the ERPs of people who stutter were characterized by reduced negative amplitudes for closed-class words, open-class words, and semantic anomalies in a temporal window of approximately 200-400 ms after word onsets. However, none of these studies considered the way these two word classes are affected when the words are pronounced in connected speech.

Considering the results of the current investigation in respect of current theories, the positional findings for function words can be explained either by Kolk and Postma's (1997) covert repair hypothesis (CRH) or Howell and Au-Yeung's (2002) EXPLAN theory. The difference between these models were described in detail in chapter one. CRH starts from the assumption that speakers who stutter have a phonological deficit that leads to errors in the speech plan, resulting in disfluency (Kolk & Postma, 1997; Postma & Kolk, 1993).

The EXPLAN model regards function and content word disfluencies as two different ways of dealing with situations where the content word's plan is correct as far as it is generated but not complete after the preceding word was first produced (Howell, 2002; Howell & Au-Yeung, 2002). In this sense, the problem is one of timing, not an error-prone phonological system as in CRH. Function words are repeated for the purpose of gaining more time for completing the plan of the content word. As only function words that precede the content word can serve the delaying role, EXPLAN accounts for why stuttering rate is higher on initial than final function words. The serial position effect on function words in PW would then reflect a position-dependent tendency for initial function words to be used for delaying onset of the content word. Thus, both CRH and EXPLAN can explain ordinal and serial position effects on function words on different premises (reflecting a covert repair in CRH and a way of gaining time for planning in EXPLAN).

CRH does not appear to offer an account of the exchange relationship between function and content word disfluencies (assuming that both are the result of covert repair processes). CRH would have to supply an answer to (1) why the incidence of covert repairs on function words is highest in childhood; (2) why covert repairs on function words are complementary to covert repairs on content words; (3) why the incidence of the covert repair options in (2) depends on the speaker's age. The principal problem is the age-dependent changes (3) that occur.

According to EXPLAN, only the first part of a content word is produced as this is the only part for which the plan is complete. When the plan runs out, the part that is available may be prolonged or repeated until the remainder of the plan is ready. In this account, disfluencies on function and content words are in complementary distribution (repeating function words prevents content word disfluencies, and content words disfluencies ensue when speakers do not repeat function words at points they could have). This accounts for the exchange relation originally found in English (Howell et al., 1999) and Spanish (Au-Yeung et al., 2003), and found here to apply to German. The EXPLAN model appears to account for the developmental exchange relationship while the tenets of the CRH would appear to provide no ready explanation for this exchange.

The results show that the patterns for English and German are quite similar. Both languages show that with increasing age function word disfluency decreases whereas content word stuttering increases. In German it is the case that these lines appear to be steeper, and in the adult age group content word stuttering is actually shown to be more likely than function word disfluencies. This is consistent with previous findings in respect of adult stuttering patterns (e.g. Brown, 1945; also see Bloodstein, 1995, and Wingate, 1988). One possibility for the German data is that content words are often longer and therefore more complex in German (e.g. compound words).

As to whether the PW provides a good context to analyse disfluencies in German, several points can be raised. First of all the structure of the PWs need to be analysed further. It was observed that there may be a higher incidence of post-content function words in PWs in German (such as 'spiel ich', 'geb ich' - I play, I give, respectively). However, the same patterns of function word disfluency were observed as in English, i.e. post-content function words were very unlikely to be stuttered. This would further strengthen the EXPLAN interpretation of findings. Another aspect is the phonetic difficulty of the content words (see the previous chapter).

Overall it was shown that there are three types of contextual effects that affect stuttering in German corresponding with what has been found in English and Spanish. For example in respect to function word position in German the pattern would be *ich ich ich streichel es* not *ich streichel es es* and serial position of function words - *in in dem Topf*, more likely than *in dem dem Topf*. Regarding the exchange pattern – younger German children would be more likely to say *ich ich trage es* whereas adults are more likely to say *ich t t trage es*. These contextual effects are readily explained by EXPLAN, but not so readily by theories like CRH.

5 Rationale for studying fluent children to gain insights into stuttering

'I speak Spanish to God, Italian to women, French to men, and German to my horse.'
Emperor Charles V, 16th century, attributed in Lord Chesterfield, *Letters to his Son* (1932), IV, p. 1497.

The second part of this thesis is concerned with the analysis of fluent and bilingual children's speech development. There are three main topics to be introduced in this chapter. The first is to give an overview of the language development of bilingual in comparison to monolingual children. Second, the evidence that suggests that bilingual children might be more likely to be prone to disfluencies are highlighted. A third issue to be introduced is the relationship between fluency development (including stuttering) and language development in general (both of mono- and bilingual children). Some of these issues have already been touched upon in the first chapter of the thesis and are recapped and extended here. An introduction to those measures that were employed as indicators of fluency development in the remaining experimental chapters is then following. Since these are new or adapted measures in the investigation of early bilingualism and its relationship to stuttering, those chapters should be viewed as work to assess their suitability for this line of investigation.

5.1 Bilingual compared to monolingual language development

As there is a vast literature on bilingual and monolingual language acquisition, the focus of this review is narrowed down to the preschool years. Two landmarks in acquisition occur during this period, 1) acquisition of the lexicon and 2) the onset of syntax. The reasons for the focus on these particular developmental landmarks is becoming clear in section 5.3 of this chapter since these two milestones in development have been directly linked to the onset of stuttering. Issues concerning language development of school aged children are considered in chapter six. The materials in this chapter are used to design a study that pilots the method used with school aged children with a preschool age group (chapter seven).

5.1.1 Monolingual preschool language acquisition

After the initial development of pragmatics the next step is the acquisition of the phonological system. The stage where most of this takes place is outside the age range of the children examined in this thesis. (See Kuhl & Meltzhoff, 1997; Mehler et al., 1988; and Oller, 1988 for work on the emergence of first speech sounds) After this the next stage in typical monolingual development is the acquisition of the lexicon. This obviously takes place in conjuction with advances in phonological development (for a review, see Vihman, 1996). Normative data on lexical acquisition by Bates, Dale and Thal (1995) using the parental reports in the MacArthur Inventories (see section 5.4 and chapter seven for more details) showed that by the age of sixteen months children could on average understand 191, and could produce a mean of 64, words. This is then followed by steep acceleration in vocabulary growth (the so called 'vocabulary burst') and by the age of thirty month the average productive vocabulary is 534 words. The more normative data is gathered, the more it emerges that individual variability is the predominant characteristic of lexical acquisition (Fenson et al., 1994). Other research suggests that some children do not even going through a 'spurt' but that these children's vocabulary growth rate is a gradual process (Goldfield & Reznick, 1990).

A long line of research has shown that children during the early stages of lexical acquisition do not always use the words correctly. Typically children make four types of characteristic errors: mismatch, overlap, underextensions and (most extensively studied) overextensions (for different theoretical approaches to these lexical errors see for instance Clark, 1993; Pinker, 1995; 1999; Tomasello & Abbot-Smith, 2000). According to different perspectives children either acquire syntactical rules (as in the formal theories derived from Chomsky, 1965, and later Pinker, 1995; 1999) or statistical regularities (as in the functional theories see for instance Tomasello & Abbot-Smith, 2000) from the input of the language spoken in their environment. It has also been suggested that the number of naming errors (i.e. early lexical retrieval problems) is particularly high during the time of the vocabulary spurt (Gershkoff-Stowe, 1997; Gershkoff-Stowe & Smith, 1997) – see section 5.4 for more details on this measure.

The extension of the lexicon eventually results in children combining words into utterances (this occurs at an average age of about 18 months). Most of this early grammatical and morphological development was extensively described in the pioneering work of Roger Brown (1973). At the two-word utterance stage children only used sparse syntax since these utterances usually consist of two object, or later action, words being connected together – function words and morphemes are omitted (Brown, 1973). As an indicator of later syntactic development Brown (1973) created a measurement method based on the average length of a phrase in terms of the number of morphemes that a child produces at a time (the so called mean length of utterance – MLU). Miller and Chapman (1981b) normed and classified MLU data and concluded that infants of age two to three years are, in what they termed the stages II (grammatical morphemes) or III (negations and questions) of grammatical development.. This is relevant in the current context since this is the age at which stuttering onset takes place (Andrews & Harris, 1964). Brown (1973) also described the order of morpheme acquisition during the early stages of productive language which was later replicated with a larger sample of children by de Villiers and de Villiers (1973). During early development, however, comprehension is far ahead of production. This has been assessed using the preferential looking paradigm in which infants as young as twelve months old view two video screens whilst listening to a simultaneously presented stimulus sentence. Using this method infants in their first year were shown to be able to distinguish between reversible sentences using only word order as a cue. This method was pioneered by Golinkoff, Hirsh-Pasek and colleagues famously using sentences such as 'Cookie Monster is tickling Big Bird' (Golinkoff, Hirsh-Pasek, Cauley, & Gordon, 1987; Hirsh-Pasek & Golinkoff, 1996a; 1996b).

	Semantics	**Syntax**
24 to 36 months	Average expressive vocabulary size at 24 months: 300 +/- 75 Average receptive vocabulary size at 24 months 900 words Comprehension strategies include interpreting sentences according to knowledge of probable events. (for the above see Fensen et al 1994)	Average MLU at 24 months: 1.92 +/- 0.5 Average MLU at 30 months: 2.54 +/- 0.6 Average MLU at 36 months: 3.16 +/- 0.7 Use of overgeneralised forms. Use of negation and questions. Comprehension of word order cues. (see in text references for MLU and syntax)

Table 6: Summary of milestones in semantics and syntax for typical monolingual children between the ages of 2 and 3 years.

The above table gives an overview of some of the main stages in typical monolingual development in the acquisition of sematics and syntax. The next section is describing lexical and syntactical aspects of bilingual development. The issue of whether bilingual children's language acquisition is equivalent to that of two monolinguals is highlighted. A model of bilingual lexical access is then introduced and how this model can be modified to account for bilingual language acquisition is described.

5.1.2 Language development in preschool bilingual children

The main question to be addressed in this section on preschool bilingual language development is whether each of the two languages' lexicon and syntax is acquired at the same rate and sequence experienced by monolingual children. The history of bilingual research has been riddled with methodological shortcomings. Some researchers have included in their samples children who can speak or understand a second language even to a minimal degree, others have not specified the time, context and reason of second language acquisition or the frequency of usage of the two languages (methodological issues that have been highlighted in the reviews of the literature by for instance Abudarham, 1987; de Houwer, 1995; Meisel, 1989). These methodological problems are particularly apparent when considering the acquisition and size of the lexicon. This has been mainly a measurement problem. For instance, is it a fair representation of the child's vocabulary size to compare the number of words known in one of their languages to the size of the lexicon of a monolingual child acquiring this language since one of the languages might also be less dominant than the other? Even when the entire number of words in both languages are combined the question arises as to how translation equivalents should be counted (see Aburdarham, 1987 for a discussion of the alternatives). To overcome some of these difficulties some researchers attempt to create groups that are comparable in all respects but their bilingualism. However even when careful selection criteria were applied, studies using the Peabody Picture Vocabulary Test (PPVT) have frequently shown differences in the scores between mono- and bilingual children – bilingual children scoring lower (Ben-Zeev, 1977; Bialystok, 1988; Merriman & Kutlesic, 1993; Rosenblum & Pinker, 1983; Umbel, Pearson, Fernandez, & Oller, 1992). Using a different measure, i.e. the MacArthur parental report inventories, Pearson, Fernández and Oller (1993) and Pearson and Fernández (1994) provided a comprehensive source of data from bilingual Spanish/English infants for toddlers between 8 and 30 months, based on longitudinal research in Miami. Their findings were analysed using four measures: the vocabulary size in each language, the combined vocabulary size in both languages and the conceptual range (which is the number of unique concepts labelled in either language). When group percentages were used (Pearson et al., 1993) the total productive vocabulary size was not found to be different to that of monolingual children. However when individual vocabulary size was investigated (Person & Fernández, 1994) it was reported that even though the majority of children fell within the normative range over forty percent of their sample were well below these levels placing them at the 10th percentile or lower for the monolingual norms. As highlighted above this is again consistent with the fact that individual differences are generally large in vocabulary growth which from this description seems to apply to both monolingual and bilingual children.

With respect to syntax studies in the literature have attempted to establish whether there are differences between children who learn two systems at once compared to those just learning one. Reviewing this literature Romaine (1999) reported that the majority of studies show that the developmental sequence of the bilingual child is the same 'in many respects' as that of the monolingual. Acquisition of syntax has been linked to the question whether the child has one or two linguistic systems (in relation to both the phonology and the lexicon) from the very beginning of speech. A commonly-held view in the early research was that the bilingual child goes through a unitary system stage before he or she succeeds in differentiating the two linguistic systems (the most elaborate model of such a process was proposed by Volterra & Taeschner, 1978). This hypothesis was predominantly based on the finding that bilingual children's language mixing - i.e. the usage of both languages within an utterance - decreased during the first few years of productive speech. As pointed out by Nicoladis and Genesee (1997) virtually all bilingual children code mix and there are vast individual differences in mixing and rates of mixing do not necessarily decrease. They go on to present evidence for early differentiation in respect of the lexicon and syntax (as well as phonological and pragmatic differentiation which are not covered here). In the case of lexical differentiation they cite research which highlights that even early in development, bilingual children possess a vast percentage of translation equivalents. In respect of syntax they describe research that shows that bilingual children use the appropriate word order, verb-agreement morphology, gender assignment and placement of negative markers for both of their languages from the earliest combinatorial utterances onwards (Meisel, 1989; Meisel & Müller, 1992; Müller, 1990).

On the whole there is now a general consensus that bilingual children differentiate their two languages at an early stage of language (most influentially argued by: de Houwer, 1995; Genesee, 1989; Meisel, 1989). From these two extreme positions (unitary – completely separate) research has recently emerged that further refined the 'separate language hypothesis' arguing that even though separate from early on the two languages are in contact and may have some influence on each other (see for instance Döpke, 1998; Hulk & van der Linden, 1996; Müller, 1998). In this line of research it has been suggested that some of the input could be problematic for children to cope with. Problematic input is seen as language combinations where there is generally a large overlap with some exceptions (Döpke, 1998; Hulk, 2000; Hulk & Müller, 2000). Children acquiring English and German would be a good test case for such views since both these languages are closely related, but show differences as indicated in chapter one. And in fact bilingual language research has shown that the English-German combination generates more cross-linguistically influenced structures (Döpke, 1998; 1999; 2000; Gawlitzek-Maiwald, 1997; Tracy, 1995) than French-English (Paradis & Genesee, 1996; 1997) or French-German (Meisel, 1990; 2001) combinations. (Also see the introduction of the next chapter for more details on this topic.) In sum the review of the literature, even though hotly debated, indicates that there is some delay in lexical acquisition for bilingual children. On the whole the development of the two grammatical systems seems to take the same sequence and there is no consistent evidence of delays. The consensus in the literature indicates that children differentiate the two languages from early on and that there is evidence of cross-linguistic influences. One of the measures that was employed with bilingual infants to investigate this process was picture naming. A model that accounts for lexical retrieval in bilinguals is discussed next.

5.1.2.1 Grosjean's (1988) bilingual model of lexical access (BIMOLA)

Grosjean's BIMOLA is a connectionist model inspired by TRACE (McClelland & Elman, 1986). According to TRACE language is processed via activation (either excitatory or inhibitory) of units in a network. Each of these units is continuously updating its activation levels on the basis of the input of other units to which it is connected. These units are organised into three levels: features, phonemes and words (with bidirectional connections between those levels). If activation of a unit is above a threshold it sends out excitatory signals to other levels and inhibitory signals to units that are connected to it. These basic principles also apply to BIMOLA. But where TRACE consists only of one such network BIMOLA assumes that there are two language networks (i.e. two sets of features, phonemes and words). These two language networks are independent but interconnected. The interconnection aspect accounts for the fact that bilinguals sometimes show interference between the two languages and can code switch when talking to other bilinguals. This interconnection also links in with the cross linguistic influences that were described in the previous section because the links between the two networks can activate each other. A second assumption of this model is that bilingual speakers have two different language modes. In the monolingual mode one language network is strongly activated whereas the other is at resting level whereas in the bilingual mode both language networks are activated, but one more so than the other. Both word frequency and close neighbours effects in the other language are represented in the size and inter-connections of the units (size of the units and shading in the model see Figure 23).

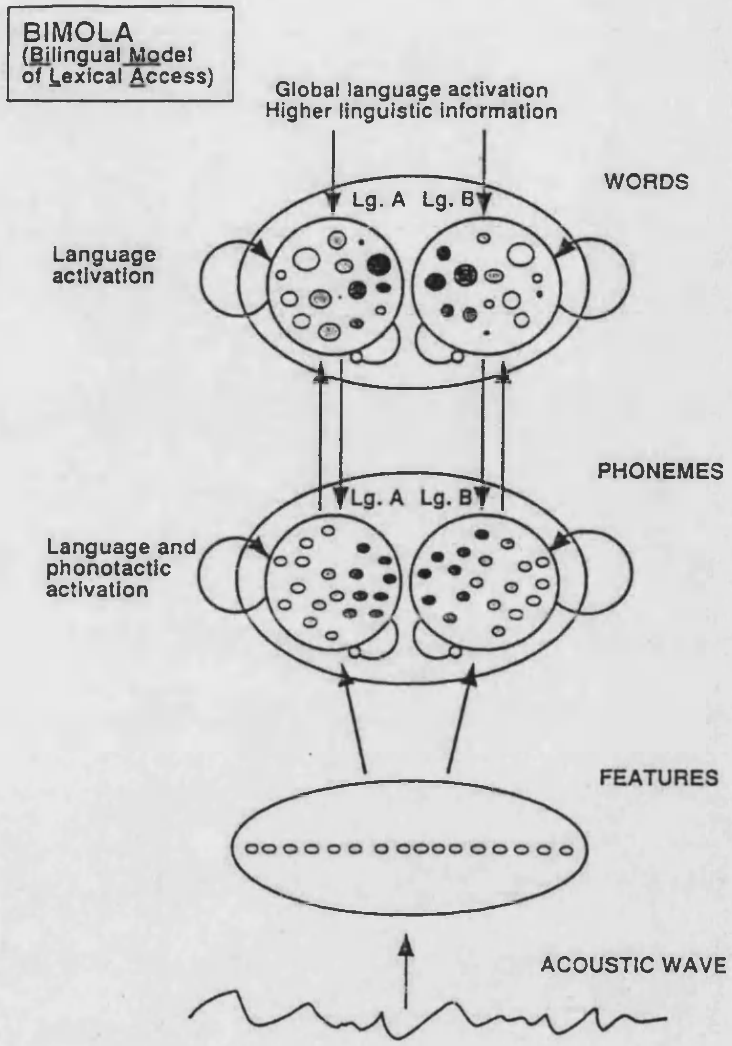


Figure 23: Graphical representation of the BIMOLA model of lexical access in bilinguals (reproduced from Grosjean, 1997).

Bimola is not specifically a developmental model but in can be presumed that in the acquisition stages the connections and activation levels, both excitatory and inhibitory, are all being developed in the child's networks. This would take place both through signal inputs from the language environment of the child and the child's own language use.

This means that the effects of cross-linguistic influences in early development can be accounted for by a BIMOLA model in which neither the excitatory nor the inhibitory connected pathways are fully matured. In this way when both languages are activated at any one time the activation can spread from one language to the other without being blocked in the process. When the child gets older these connections develop more fully and cross language transfer becomes less frequent.

5.2 Might bilingual children be more likely to be prone to disfluencies?

Recall from the first chapter, that a direct link has been made by certain authors between bilingualism and multilingualism and stuttering. One line of evidence in favour of this was the higher prevalence of stuttering in multilingual children compared to their monolingual counterparts (Stern, 1948; Travis et al., 1937). From these prevalence rates these and other authors draw a causal link between stuttering and bilingualism, for at least some speakers. The other line of evidence looked at the time of stuttering onset. In the study by Travis et al. (1937) the age of onset of stuttering coincided with the introduction of a second language for 26 % of the bilingual children who stutter. The book by Pichon and Borel-Maisonny (1964), in which the authors concluded that early bilingualism is conducive to stuttering, was based on their finding that 25 (or 14%) of the children they studied had to use more than one language at the time of stuttering onset. In a Belgium study looking at language disorders in the children of foreign mine workers, Mussafia (1967) suggested that stuttering as well as other speech and language disorders are caused by the sudden change from one language to another upon arrival in a new country. Language switching as a trigger which precipitates the onset of stuttering was also suggested by Lebrun and Paradis (1984). These authors specifically focus on the language switching of parents but also mention the difficulty of children's use of two or more languages. Both parental language mixing and the child's own code switching can, according to Lebrun and Paradis (1984) precipitate stuttering in children with a predisposition to this disorder. A further study by Karniol (1992) described the case of a child who started stuttering when a second language was introduced. The author made a direct causal link between those two incidences and described how the stuttering stops when the child was encouraged to only communicate in one language.

In contrast to this in the domain of speech therapy Abudarham (1987) has researched the effects of bilingualism on speech development in his capacity as national UK adviser to the speech therapy profession on the subject of bilingualism. In his analysis he highlighted the methodological shortcomings of the many studies that have reported adverse effects on bilingualism on speech and language development. Some criticisms he raised were uneven focus on one of the two languages only, not being specific enough about when and how the two languages were acquired and also not being clear about the relationship between the two languages (i.e. how closely the two languages are related in syntax, phonology and morphology). Because of these methodological problems he was extremely sceptical about claims of lexical and/or syntactical delay in bilingual language acquisition. For instance, in the case of Macnamara's (1966) review which concluded 'bilinguals have a weaker grasp of language than monoglots', Aburdarham (1987) specifically criticised Macnamara's interpretation of bilingualism as being too wide. In the acquisition of articulation and phonology he quoted authors who describe a period of phonological confusion in their bilingual child subjects (Leopold, 1950; Ruke-Dravina, 1967), which seems to suggest some degree of difficulty for bilinguals. With regards to fluency development he made a distinction between bilingual children of immigrant populations and those he refers to as 'non-ethnic' bilingual children (for instance English / Welsh bilinguals in Wales, or Spanish / Catalan speakers in Barcelona). Aburdarham (1987) assumed that the higher incidence of stuttering in the children of immigrants suggested a stronger influence of psychosocial factors. He therefore concluded that bilingualism in itself is not a precipitating factor in fluency failure, i.e. that in the immigrants the social factors alone lead to more disfluencies regardless of the additional language. This, however, contrasts with his own advice to parents of bilingual children who stutter, to delay the exposure to a second language until the problems of the first are resolved (Miller & Abudarham, 1984). The contradiction being that if two languages are not a problem in development of fluency why should the parent restrict the child's speech to one language alone. This stance is not as clearly causally linking stuttering to bilingualism as such it does, however, suggest that the demand on the child would be lowered when only one language were in use. This links in with the demands and capacity model of stuttering (Starkweather, 1997; Starkweather & Gottwald, 1990; Starkweather & Gottwald, 2000). This is a framework that proposes that the amount of attentional / processing load is directly linked to the onset of stuttering (more detail about this model is given in chapter seven). Aburdarham (1987) also based some of his criticisms on the assumption that children start out with a unitary language system and only later on separate the two (see for instance Volterra & Taeschner, 1978). This issue had since been re-evaluated see previous section, i.e. it is now generally accepted that the two languages are separated from early on in language acquisition.

To summarise this section there are three main lines of evidence suggesting that bilingual children might be more prone to disfluencies than their monolingual counterparts. There is the research indicating a higher proportion of speech problems within bi- or multilingual populations. Secondly age of stuttering onset seems to coincide with the age of second language usage at least in some children. Finally therapeutic advice is given to drop one of the two languages when a speech problem occurs in childhood.

5.3 Fluency development (including stuttering) in relation to overall language development in general

Another issue needing examination is the relationship between non-fluencies in normally developing children and the fluency problems by children who become children who stutter. Relatively early on it was established that young normally developing children exhibit all of the types of disfluency found in the speech of children who stutter (Haynes & Hood, 1977; Johnson & Associates., 1959; Wexler, 1982; Wexler & Mysak, 1982; Yairi, 1982). This research focused on the ages of between 2 and 6 years. Yairi's (1982) study, for instance, looking at the developing speech of thirty-three 2-year-old children for a year, found that they showed disfluencies that would be described by parents and other observers as stuttering. Johnson and his co-workers (1959) interpreted such findings as evidence that through the influence of parental behaviour striving for perfection in the child, the child develops an anticipatory avoidance reaction leading to stuttering. In contrast Bloodstein (1988, 1995) interpreted these studies as evidence for a continuum between early stuttering and certain types of normal disfluencies (see chapter one for more details on the continuity hypothesis). This has implications for the diagnosis and treatment of children with speech disfluencies in infancy (see Bloodstein 1995 for further details).

As was highlighted in chapter one, another problematic issue for theories of stuttering is the finding that the onset of stuttering does not co-occur with the first expressive communication of a child (see for instance Andrews & Harris, 1964; Yairi, 1983; Yairi & Ambrose, 1992b). One study even suggested that that children who persist in their stutter have a later onset (mean of roughly 38 months) than those children who recovered from early periods of stuttering (Yairi, Paden, Ambrose, & Throneburg, 1994).

When considering both of these issues together, i.e. the often occurring disfluencies in normally developing children and the time of onset not being identical with the time of first language production, the study of infants (both bi- and monolingual) at the time of stuttering onset is essential. This is to clarify diagnostic procedures (normal non-fluencies and those that might be indicators of stuttering) and to find out which processes occur at the time of stuttering onset and whether these are contributory, causal factors.

Recall also from the first chapter that one of the authorities on stuttering, Nan Bernstein Ratner (1997) has made a direct link between the rapid changes in the acquisition of the lexicon in conjunction with start of syntactical processing as a factor in the onset of stuttering. This was based on Wijnen's (1990) 'Development of the Formulator' hypothesis which predicts that developmental patterns of disfluencies (reasoning from disfluency patterns in normally developing language learners) are precipitated by the inclusion of function words and serial order planning in the expressive grammar of children. This is one of the reasons for the study of bilingual children in the remaining chapters since the task of acquisition of the lexicon and grammar is doubled in this population. It could thus be the case that bilinguals also show language development problems at the ages where stuttering onset is reported.

5.4 Aims and measures of the second part of the thesis

Though there are debates about how to define stuttering (see chapter one), there is unanimous agreement that stuttering is a disorder in which fluency fails. Children who are always fluent, children who stutter and will become fluent, and children who later develop a persistent stutter experience occasional points in utterances where their fluency breaks down, i.e. fluency failures (the continuity hypothesis – see previous section and chapter one). Two points covered above should be highlighted again. 1) fluent children show the whole function word disfluencies seen in stuttering, 2) you only start to see children referred to stuttering clinics some time after language onset. Two sets of questions can be raised: 1) As fluent children and children who recover show whole function word disfluencies patterns, what differences arise in the children who stutter who continue into adulthood, and at what age? 2) Is there something that happens at the age of stuttering onset that precipitates the fluency problem in the children who stutter (CWS)? Moreover if bilingualism precipitates stuttering, as suggested from the above review, then it might be expected that bilingual children at the time where stuttering first occurs would also show more language problems at this point. Specifically when drawing on both EXPLAN and Bernstein Ratner's (1997) suggestion of the combination of lexicon and syntax, it was examined whether these children show lexical planning problems (both as in naming errors and TOTs) and / or difficulties in syntax acquisition.

Researchers attempting to answer these questions have to adopt a position with respect to two major issues: On the first question, some framework is needed that a) suggests the relationship between fluent and stuttered speech (cf the continuity hypothesis as proposed by Bloodstein, 1987, 1988), and b) provides an explanation of how CWS diverge from the fluent pattern shared with the younger fluent speakers.

These issues have been addressed in chapters 2-4, where I examined some implications of the EXPLAN model as it pertains to issue b). Further work should be carried out to clarify a) – a point that is going to be taken up again in the final chapter of this thesis.

The studies in the remaining chapters are designed so that they examine language development in young German speakers who will, more than likely, become fluent adult speakers of this language (even in speakers who stutter, Andrews' work in the introduction shows, 99% of all speakers do not suffer from stuttering, so an individual who stutters is unlikely to be encountered in the comparatively small samples used below). The factors to be examined in early language development were chosen because they have relevance to stuttering (and, indeed, some connection with those factors examined earlier). If bilinguals experience more difficulties than monolingual children at the time when stuttering first occurs than this would strengthen the suggested link between bilingualism and language disfluencies.

The measures to be examined are drawn from work on fluent monolingual speech development and a reason for their inclusion follows:

1. Picture naming – Gershkoff-Stowe and Smith (1997) have shown a peak in naming errors at 50 word stage. The word usage data is obtained from the Communicative Development inventory, CDI (the reliability and validity of the test instrument have been demonstrated by Dale, Bates, Reznick, & Morisset, 1989) and a German translation equivalent, GCDI - see below. It was examined if there are children

who have a high number of picture naming errors at the time of the main vocabulary-spurt. If there are, do these relate to other aspects of fluency development (examined below) – such as later MLU and receptive syntax?

2. The Communicative Development Inventory and a German adapted version (CDI and GCDI) – measures of vocabulary development. These data are mainly lexical, but as there are classes of words that are content and that are function, there are possible ways to link the scores with syntactic development. Are rates of vocabulary development (lexical and syntactic) different between children who have picture-naming errors? Also the relationship between the two languages' vocabulary development can be assessed (i.e. a correlation of CDI – in particular the UK version of CDI [OCDI] and GCDI should have a 45 degree slope if the two languages are acquired at the same rates).

3. Mean Length of Utterance (MLU). The standard measure of productive development, in particular index of syntax. One question to ask is whether this measure relates to picture naming, and to OCDI/GCDI.

4. Reception of syntax test (ROST). This affords a perceptual measure of the infants' receptive syntax. Are children with a faster vocabulary acquisition also better at syntax perception? If so, there should be a close relationship between the child's receptive syntax and their MLU.

5. Carey's temperament scale. This is a measure of the infant's temperament. Here the relationship between certain temperament factors with the other measures can be examined.

In the following part of this chapter more details are given about the above measures and the reason for their selection. The first study introduces a German form of a syntax test for the longitudinal study. The reception of syntax test (ROST) for English and its validation is described in Howell, Davis and Au-Yeung (2004). The German form of this test was designed with the same principles as ROST but looks at some syntactic factors specific to German. The experimental work was designed, a) to validate the German form of ROST (by looking at the relationship to the validated English form) and, b) to examine performance by three groups of children on the English and German forms, where the groups were selected to reflect a factor that may be associated with fluency development (whether the children are mono- or bi-lingual). The tests were made on a monolingual English group, a monolingual German group and a bilingual English-German group. It was analysed whether the bilingual children show syntactic deficits relative to each of the monolingual groups (Hulk, 2000). More specifically the two factors that were emphasised in the first part of the thesis, i.e. the issue of complex content words (compound nouns) and word order differences were investigated in bilingual children and compared to their monolingual counterparts. The second study on normal fluency development, constitutes an intensive, longitudinal investigation of five bilingual children from age two. Five is a small number for experimental investigation, but in longitudinal work with very young children, this is an adequate (and arguably large) sized sample. The tests made on these children are a picture naming task, productive vocabulary size measures in both languages, ROST (English and German forms, validated in the previous study), and MLU data. This chapter analysed whether there is evidence of a link between bilingualism and language disfluencies at the time that has been identified as the point of stuttering onset. If there are more naming errors and/or more receptive and productive syntax problems it would specifically highlight the planning part of EXPLAN as being affected in bilingual children. Evidence such as this could provide a reason why this link between stuttering and bilingualism was established. Finally the third study looks at lexical access (the TOT paradigm) in German speaking children in two age groups. Some of the factors that were highlighted in the first part of the thesis, such word length and phone the word starts with, are analysed in this respect. Only a very limited number of studies has ever looked at children's TOT states. Investigating whether the factors that affect which words are stuttered more often also affect whether they are going to be retrieved by children, is directly looking at the planning of words. TOT states in this respect are seen as acute instances of cases where the planning of a word has not been successfully finalised.

6 Reception of syntax test in bilingual school children compared to monolingual control groups

"Learn well your grammar, / And never stammer."
Lewis Carroll, c. 1845, 'Rules and Regulations', in the Faber *Book of Useful Verse* (1981), p. 90

6.1 Introduction

It has been highlighted throughout that one factor raised in the psycholinguistic research into stuttering is the acquisition of syntax and its impact on fluency development, as well as the onset of stuttering (Bernstein Ratner, 1997). To address this particular issue this chapter sets out to develop test material that can be used to assess syntactical development of bilingual children from very early infancy onwards. One of the general aims of the current chapter is the investigation into whether bilingual children do acquire syntax in the same manner (i.e. sequence and rate) as their monolingual counterparts. The nature of cross-language influence were also investigated and if it does occur what can it tell us about bilingual language processing in general and more specifically can it shed light on issues that were investigated in the first part of this thesis. As described in the previous chapter there is evidence for bilingualism as being a precipitating factor in the onset of stuttering. It could then be hypothesised that bilingualism might tax speech production in a similar way to stuttering which would then mean that bilinguals might show a deficit on acquisition of language processes thought to be implicated in stuttering.

6.1.1 Bilingual English-German syntax acquisition

As described in previous chapters the simultaneous acquisition of two language system in bilingual children has been a contentious issue in the last twenty years. Initially the 'one system' hypothesis (Volterra & Taeschner, 1978) was influential in this line of research. Drawing on evidence from bilingual infants' mixed utterances it was assumed that these children at the outset do not differentiate the two languages. Differentiation is then argued to be a gradual process occurring during the third year of life first in the domain of the lexicon and later on a syntactic level. This view became unpopular about ten years later and was replaced by the separate language hypothesis (Genesee, 1989; Meisel, 1989) whose supporters went as far as suggesting that the two languages of bilingual children develop autonomous in the same way as monolingual acquisition. In this line of investigation researchers argued that any form of cross-linguistic influence was too unsystematic and too infrequent to indicate a difference in bilingual acquisition processes (de Houwer, 1995; Meisel, 1990).

As highlighted in the previous chapter these two positions represent extreme ends along the unitary / separate continuum and more recently various researchers, even though assuming the two languages develop largely separately, allow for the possibility of cross linguistic influence and have put this line of analysis back into focus (Döpke, 1998; 1999; 2000; Gawlitzek-Maiwald, 1997; Müller, 1998; Paradis & Genesee, 1996; 1997). This influence between the two languages is not, as was in the unitary system view, interpreted as mixing or fusion, but rather as either facilitation, delay or transfer (Paradis & Genesee, 1995; 1996). The same view, i.e. one that hypothesises cross – linguistic influence on certain syntactic structures is also adopted in this part of the thesis. This debate has strong theoretical implications. Strong proponents of the separate language hypothesis come from a formal universal grammar background (as so revolutionary proposed by Chomsky, 1965; 1995; and re-interpreted most prominently by Pinker, 1995; 1999) whereas those that do allow for cross linguistic influences have a more functional developmental language perspective (see for instance the connectionist model by Bates & MacWhinney, 1989; or the construction grammar account by Tomasello & Abbot-Smith, 2000).

For the acquisition of the syntactic structures of the English-German first language combination the research of Döpke (1998, 1999, 2000) is of particular relevance. Based on her longitudinal research with bilingual English-German children she has argued that this particular language combination generates more cross linguistically influenced structures. This, she argues, is due to the fact that there are many surface similarities between the two languages which are however mediated by underlying differences (see Döpke, 2000). She has focused on the acquisition of specific structures that are different in the two languages – one study analyzed the acquisition of verb placement and the second was concerned with the placement of negation and modal particles. Since the former is relevant with respect to studies in the first part of the thesis it is outlined here in a bit more depth. It was described in the first part of the thesis that English and German place verb phrases differently (head-initial in English and head-final in German). Döpke (1998) points out that monolingual German children differentiate the placement of finite and non-finite verbs immediately (finite appear in second whereas non-finite verbs are placed in final positions). She points out that none of the literature on monolingual syntax acquisition in German has ever reported cases where non-finite verbs precede the verb compliment at any stage in development. In her data of the bilingual children there was evidence of confusion in verb placement which indicated cross-language influence between English and German. In a later study she also found cross-linguistic influences acquisition of the placement of negation and modal particles in German – English bilingual children (Döpke, 1999) – both word order and negation are also included in the test material used in this chapter. She interpreted her findings in light of the Bates & MacWhinney's (1989; 1997) competition model. This model is based on the assumption that children learn language by weighing the probabilities of cues presented to them in the input.

The first part of the thesis established that the large number of compound words used in German are probably a reason for the higher phonetic complexity of German content words which also lead to a higher stuttering rate of German content words (see particularly the results of chapters 2 and 3). There now follows a description of the acquisition of compounding and cross-linguistic studies of compound production and comprehension at preschool age. The frequency of compound noun usage in the creation of new lexical items varies across languages. This difference across languages is one of the reasons for the ease and age of acquisition of these structures. For instance Germanic languages on the whole use the process of compounding more frequently than the Romance languages – French for instance tends to use derivation for lexical creations (Clahsen, 1995; Clark, 1998). In English for example, it has been shown that compound production and comprehension is acquired relatively early, i.e. before the children were three years old (Clark, 1981; Clark, Gelman, & Lane, 1985). In contrast, in Hebrew where compounding involves complex morphology and is infrequently used, it has been reported that children rarely produce compounds before the age of four years (Berman & Clark, 1989; Clark & Berman, 1987). For English noun-noun compounds (e.g. 'police car') the modifying word appears in the first position whereas the one to be modified always appears in the final position. The opposite is the case for French where compound nouns are left headed (Clark, 1985). In a number of studies with bilingual French-English speaking infants, Nicoladis (2002; 2003a; 2003b) has shown that preschool children do show cross-linguistic influences of one language on the other in their compound usage. Even though the structure of German and English compounds is not significantly different there is a difference in usage frequency which then might increase the strength of this cue from the input. It might therefore be argued that German speaking bilingual children hear a larger amount of compounds in their environment which then affects their ease and age of acquisition. In the following section the type of measurement task is described.

6.1.2 The reception of syntax test (ROST)

The reception of syntax test (ROST) was designed to test the perception of syntactic categories across a wide age range, from two to ten years (Au-Yeung, Howell, Davis, Sackin, & Cunniffe, 2000; Howell, Davis, & Au-Yeung, 2003).

ROST is loosely based on the most widely used such test (Bishop's 1983, test of the reception of grammar - TROG). The aim of ROST is to create a test that can be used with very young children, as well as school-aged children, and was adapted for the German language by the author. The original TROG test was created for children of four years and older, which meant that a new procedure had to be introduced which is simplified so as to be suitable for a younger age group.

In ROST a sentence is played whilst two pictures are simultaneously presenting on a laptop monitor. Of the two pictures only one corresponds with the sentence heard. The child is then asked to touch the picture that corresponds to the sentence (the child's response is detected via a touch screen). Pointing to objects the child hears about can be seen as one of the earliest responses in language development (Brown, 1973).

As ROST is a perceptual test (rather than a productive measure as in the case of MLU) and it has the advantage of being less time consuming in administration and analysis, it can test the understanding of syntactic categories that a child may or may not be using productively. For very young children the 'preferential looking' paradigm has also been used to test lexical and syntactic knowledge (with infants as young as 14 months) - (Golinkoff & Hirsh-Pasek, 1995; Golinkoff et al., 1987; Hirsh-Pasek & Golinkoff, 1996a; 1996b). However, the technical equipment (eye tracking and / or video cameras) needed for the preferential looking paradigm make it impossible to use the task in home or school settings.

As previously stated ROST is loosely based on Bishop's (1983) TROG test. The TROG test is administered in the form of a work-book and can be used with children aged 4 and over. ROST extends this type of test to a wider age group, from 2 years onwards, and uses a laptop computer to display the stimuli. The interactive nature of the computer program (it allows the child to control and time the display of sounds, images, and it collects and stores the responses) make it comparable to a type of game. As such it keeps the attention of the child. Even though the child can work at his or her own speed, time is saved on randomisation and category decisions as this is done automatically by the program. The literature was also examined to see whether there were other suitable tests for assessing all the above range of abilities. In German, a relatively new and widely used diagnostic test for two year old children is the 'Sprachentwicklungstest für zweijährige Kinder (SETK-2)'. There is also a version of the same test (SETK- 3-5) for three to five year old children (both designed by Grimm and colleagues, 2000). This tests both productive and receptive language development with picture cards, objects and cassette tapes. However, this is a language test for the diagnosis of serious developmental language deficits, such as autism, hearing and other sensory problems. As such it was not suitable for the present investigation and would not allow for comparisons between the two languages.

For children of school age it is important to design a task that is neither too difficult nor too simple (both would decrease motivation to finish the test) for a wide age range. In Bishop's TROG test this is achieved by dividing children in two age groups below and above 7 years old. Within these age groups there is also room to move to easier or more difficult categories depending on each child's individual performance. This is a feature that was also included in ROST. The details of this are described in the methods section.

One aim of the current investigation is to examine certain of the English – German cross-linguistic differences that could affect fluency particularly when a child learns both English and German (some linguistic differences such as word order and compound nouns were covered in chapter one of the thesis). Both word order and compound noun category are included in ROST testing. Details of the syntactic factors chosen for examination, and how the German differs from the English categories appear in the method.

6.1.3 Aims and Hypotheses

For children of school age (who would have already be assumed to have acquired the various different German plural forms – see next chapter) specific attention were paid to the syntactic categories of compound nouns and word order. The following hypotheses were investigated: 1) It is predicted that bilingual children pass fewer categories in their less dominant language. 2) Since compound nouns are very frequently used in the German language it is assumed that the bilingual children find this category easier than the English monolingual children. 3) Word order, however, might be more difficult for the bilingual children since the German language is both more rigid in respect of verb position but also more flexible in respect of noun word order (see introduction and previous chapters).

A further aim of this chapter, on a more general level, is the validation of the current German version of this test. A developmental pattern emerging in the German data showing that, with increasing age, receptive syntax knowledge is improving would provide internal validity of the test material, i.e. it measures what it sets out to measure. Finding similar results in the monolingual German sample compared to the monolingual English sample (which has been validated in a large scale project – Howell et al., 2003) would give the results a measure of external validity. Finally, a more detailed analysis of individual categories investigating differences between bilingual and monolingual children would indicate whether the test is also a sensitive measure of group differences.

6.2 Method

6.2.1 Participants

Forty-four bilingual children participated at the European school (23 girls and 21 boys). The breakdown of age groups was as follows: ten pupils were aged between 10 and 11 years old, eight were between 9 and 10; six within the age range 7.5 to 9.0; seven between 6.5 and 7.5, five between 5.5 and 6.5 and eight children were in the youngest age group (4.0 to 5.5 years old). Questionnaire data were obtained to assess the level of the child's bilingualism. Thirty-four children (18 girls and 16 boys) participated in a monolingual school in Germany (Kreuzschule – Münster – an RC primary school). Six were between 6.5 and 7.5, eight between 7.5 and 8, six between 8 and 9 and fourteen between 9 and 10 years old. Forty-four monolingual English speaking children were selected from a London school of comparable, socio-economic background (St. John's RC primary school). The monolingual English-speaking children were age- and sex matched to the bilingual group.

6.2.2 Test Material

The complete set of syntactic categories tested over all ages are given in Table 7 (blocks A-I are the nine simple syntactic categories that are tested in the under 5 year age group – see caption and description below).

English ROST Categories and examples	
Subject-verb-object	A8. The boy carries The girl
Article	B7. Some sheep
Plural	C1. Cats C2. cat
Pronoun (He/she)	D1. he drinks D2. she drinks
Pronoun (his/her)	E1. her cat E2. his cat
Pronoun (number)	F1. They sang F2. He sang
Preposition (In/on)	G1. in the house G2. on the house
Preposition (other)	H1. on the table H2. under the table
Compound noun	I7. a car book I8. a book car
Negative	J3. The shoes are not red.
Direct Object / Indirect Object	K1. The girl throws the boy the ball.
Passive	L1. The girl is chased by the man
Long Distance Movement	M5. The apple that the Girl thinks the boy has
Seems	N1. The girl seems to The boy to be happy
Comparative	O4. The girl is taller Than the boy
As (adjective) as	P1. The girl is as Short as the boy
Relative clause referring to subject of the sentence:	Q1. The boy chasing the girl wears a hat
Relative clause referring to the object of the sentence:	R7. The girl carries the boy who wears a hat
Relative clause referring to the subject of the sentence:	S5. the man the woman chases wears a hat
Other relative clauses:	T3. The girl the boy thinks kicks the ball.
PRO	U1. The girl wants to throw the boy a ball
Passive Negative	V7. The boy is not kissed by the girl
Not as (adjective) as	W1. The girl is not as tall as the boy
X but not Y:	X1. the ship but not the car is red
Not only X but also Y:	Y1. the girl has not only a duck but also an apple
Neither X nor Y	Z3. the bunny wears neither a hat nor shoes

German Rost Categories and examples:	
Plural Morpheme	A1 Stift A2 Stifte
Plural Article	B1. die Löffel B2. der Löffel
Plural Vowel Change	C7. Vogel C8. Vögel
Comparative	D7. Der Stift is länger als der Löffel
As (adjective) as	E6. Das Mädchen ist so groß wie der Junge
Not as (adjective) as	F5. Das Mädchen ist nicht so klein wie der Junge
Prepositions (in/on)	G1. In dem Käfig G2. Auf dem Käfig
Prepositions (other)	H7. hinein gehen H8. heraus gehen
Pronoun (he/she)	I3. er singt I4. sie singt
Pronoun (he/they)	J1. Der Hund jagt ihn J2. Der Hund jagt sie
Subject Verb Object	K1. Die Frau schiebt das Kind
Object Verb Subject	L2. Die Frau jagt der Mann
Relative Clause referring to the object of the sentence:	M1. Die Frau schiebt das Kind, dass einen Hut auf hat
Relative Clause referring to the subject of the sentence:	N8. Das Pferd, das die Katze jagt, ist braun
Direct Object / Indirect Object:	O1. Der Junge holt dem Mädchen einen Kuchen
Compound Noun	P3. ein Ballonbuch P4. ein Buchballon
Negative	Q3. Der Junge läuft nicht
Passive	R6. Die Frau wird von dem Kind gezogen
Passive Negative	S3. Der Junge wird nicht von dem Mädchen getragen
X but not Y:	T1. Der Schneemann hat einen Schal aber keine Schuhe an
Not only X but also Y:	U2. Das Mädchen hat nicht nur einen Teddybär sondern auch einen Hund
Neither X nor Y:	V3. Weder der Hund noch die Katze sind braun
Seems	W5. Der Junge scheint dem Mädchen traurig zu sein

Table 7: Syntactic categories tested in ROST in English and German. Examples of sentences used in the test are given. For the under 5 year old age group in English nine categories (A-I) are used. For German eight categories (A, B, C, G, H, I, J, K) are used in this age group. Children were tested on one language at a time, alternating between the languages on subsequent visits.

As highlighted in the introduction responses are available to ROST while the test is being performed (and testing of categories can be completed at different times, depending on the individual randomization). It is therefore possible to reduce the number of categories presented. In each syntactic category there are 8 different sentences as examples. A category is considered to be achieved if 7 from the 8 are correctly selected (sign test $p<0.05$). If a child failed two test items in a category (which meant that the child could not pass this category to meet the statistical criterion), the child failed this category and no further test items from that category would be presented. If a child passed 7 items from any one category the eighth item would not be tested because the child had already reached the statistical criterion. A more detailed description of the processes which allow participants to move to easier or harder categories is now given. The same procedure is used in both languages – throughout this description the alphabetic category letters of the English language test are used. However, the procedure for the German test is identical, which means that it refers to the same syntactic categories (even if the letters sometimes do not correspond). Decisions about omitting certain blocks (similar to the procedure in TROG) can be made automatically due to the fact that the software is programmed to check and record online responses of each individual child participant. The reduction of categories takes place in two ways: there is one contingency where syntactic categories are credited which are not tested (equivalent to those below the baseline in TROG) and another where syntactic categories are only tested if certain other blocks are passed. The first of these conditions is implemented at macro and micro levels. At a macro level, children are assigned to different forms of ROST depending on their age. Typically this means that children under five are tested on nine simple syntactic categories and the children of five and over are not tested on the first eight of these simple categories (with certain finer rules as indicated below). One micro level of adjustment made with the young children is that the next eleven blocks (J-T in Table

1) of syntactic categories are tested if a child passes the first block. The second fine level of adjustment is that the more advanced category V is tested if L and J are passed, W is tested if Q and J are passed and U is tested if K is passed. Finally, the four picture tests (X, Y, Z) are tested if J is passed. The gross level for children five years and over is that they start at block nine (compound nouns) that the younger age group was also tested on. The first fine level is that if J and K are failed, the first eight easy categories (A-H) are tested. Also L is tested if V is failed and L is credited if V is passed. P is tested if W is failed and P is credited if W is passed. U is tested if the easier K is passed, as for the young group. The tests that involve four pictures (X, Y, Z) are tested if J is passed. The test sentences were pre-recorded in WAV format and were played over an external loudspeaker connected to the computer. For the English version a male native English speaker was used and for the German version a female native German speaker (the author) was recorded. Recordings of a single speaker ensures that the test sentences are standardised (which would otherwise have to be specified in manuals – see for instance the TROG manual on stress patterns). The recordings were checked and some items had to be re-recorded due to prosodic factors such as stress of individual words or of a sentence prosody that could lead to misinterpretation of the item. Location of correct test picture was counterbalanced across all trials and test items were individually randomised.

6.2.3 Procedure

The participating pupils were assessed individually using a laptop computer in quiet surroundings. Headphones were used to eliminate any distracting noise. Instructions were given in German to the pupils who would perform the German grammar assessment and in English for the monolingual English group of children and whenever the English test was administered. Bilingual children in the European school were tested on two individual sets of visits (each language once) and choice of language was counterbalanced across pupils.

For both languages there were four items that were used in the learning phase. Once the first set of pictures appeared the researcher or parent would ask the child to touch the picture that would go with the word the child would hear. A picture of a loudspeaker appeared beneath the two pictures and the child was encouraged to touch it to start the sentence / word. Whenever the child had selected a picture by touching the computer screen the background colour of the picture changed from blue to red. In the learning phase of the experiment the child would be provided with feedback by the experimenter. Once a picture was selected a green arrow became visible underneath the pictures which led to the next trial. Trials in the learning phase could be repeated to ensure that the child fully understood the procedure. A child was asked to repeat the learning trials if he or she selected the wrong picture. When the researcher was satisfied that the child understood the procedure the experimental trials started. Experimental trials followed exactly the same pattern as the learning trials. However, rather than having two different object names, as was the case in the learning phase, the pictures presented in the experimental trials only varied on the target syntactic dimension (for the English subject-verb-object category an example would be a picture of a girl carrying a boy and vice versa). The child could not select a picture until the test sentence / word had been heard at least once, but the child was allowed to re-play the test utterance as many times as needed. If a child accidentally selected the wrong picture but it was clear he or she really favoured the other picture (the child made a statement to that effect, for instance) they were allowed to select the other picture by touching it. The children were encouraged to control the running of the trials themselves by touching the icons on the screen. No feedback was given for the test trials to allow re-testing in the other language on the second sets of visits to the school. The bilingual children were then asked to complete a questionnaire to assess level and status of bilingualism (Appendix 1). This was usually carried out in an interview format (particularly with the younger children), i.e. the researcher asking the questions from the questionnaire and noted the results.

6.3 Results

6.3.1 Analysis 1: Description of Questionnaire Assessing Level of Bilingualism

Results of the questionnaires administered to assess level of bilingualism of the pupils showed that all of the children rated themselves as bilingual and competent in both languages for reading and writing (for those children who had acquired reading / writing skills at the time). However, closer inspection of the questionnaires indicated that the group consisted mostly of German-dominant bilingual children. The majority of children had been born in Germany (63.6%), 34.1 were born in the UK and 2.3 % in other countries. Graphically this dominance can be clearly seen in the following pie chart:

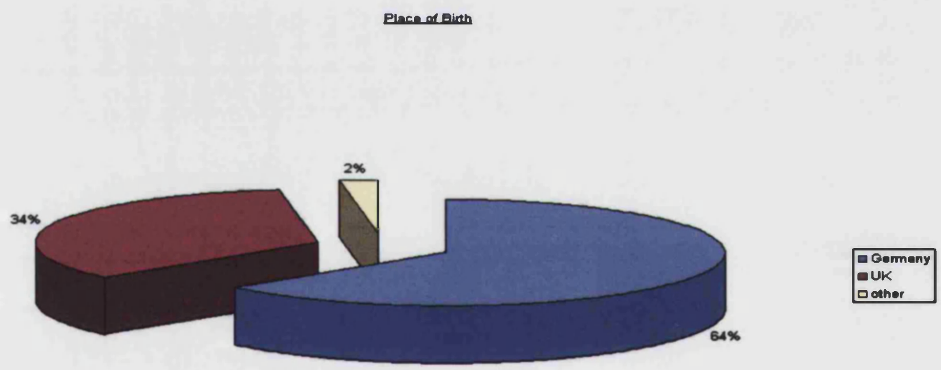


Figure 24: Pie chart indicating the percentage of bilingual pupils who were born in Germany, the UK or other countries.

This was also reflected in the native language of the parents. Pupils with both parents speaking German as their mother tongue predominated (81%). The remaining children had one speaker of each language as parents but none had both English-speaking parents. Time in months stayed in an English speaking country also confirmed German as the more dominant language. Length of stay varied from 1 month to 125 months with an average time of 54.6 months since arrival in an English speaking country. 83.7 % were right, 7.0 % left-handed and 9.3% used both hands equally. It was thus decided not to analyse the more simultaneous bilingual children separately since the number and spread across age group would be too small.

6.3.2 Analysis 2: ROST Results

Since the two language tests contained a different number of categories (23 in the German and 26 in the English version) "number of categories passed" was converted into proportions of the whole set and calculations were based on these transformed data.

The following figure represents the proportion of categories passed for each school per age group:

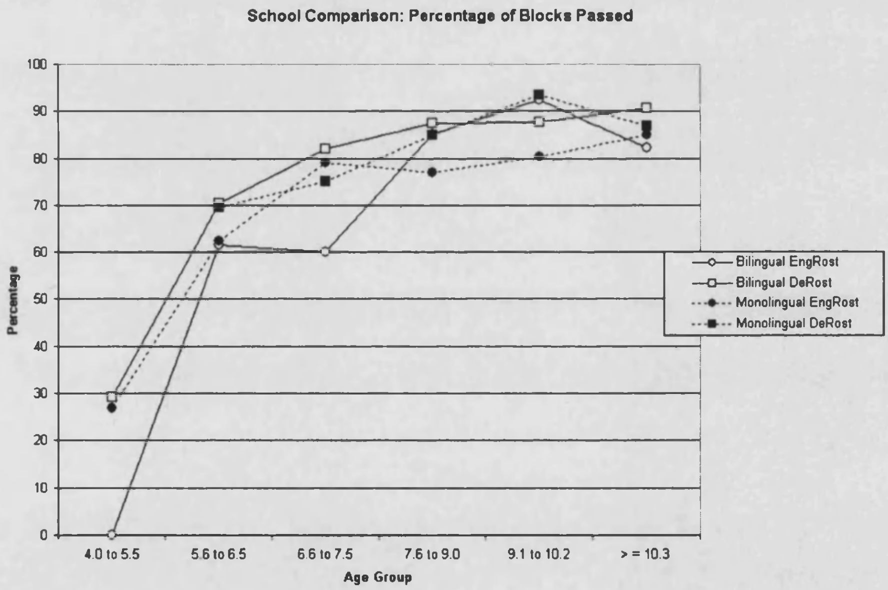


Figure 25: Mean percentage of categories passed per age group (as indicated on the abscissa). Different lines represent different language groups. The dashed lines refer to the monolingual language groups and the solid lines refer to each respective language of the bilingual age group. Round markers refer to results of German whereas square markers refer to the English ROST results (unfilled for bilingual children and filled markers for the monolingual comparison group). Note, since primary schools in

Germany start at a later age, the line referring to the monolingual German results starts with the second age group. As such comparisons with that school used age as a covariate to partial out the difference in age groups. A graph with error bars (to indicate variability is not presented here to aid comparison but can be found in Appendix 2).

This graph quite clearly indicates that the four lines showed a very similar pattern of results. The zero mark for the youngest children in the EngRost result was mainly due to the fact that most of them did not complete the task and thus were not credited with the categories. However the German version for these speakers showed a very similar result, in the figure indicated by the overlap. On closer inspection of the graph it can be seen that the bilingual children score consistenly below the monolingual comparison groups at least in the English version of the task (but not significantly so).

6.3.3 Analysis 3: Does Language Dominance Affect the Results?

To assess the bilingual performance within their school, a mixed model ANOVA was carried out with language of test as a repeated measures factor and age group as the between subjects factor. The impression of the graph that the shape and slope of the lines are quite similar was confirmed by the analysis. The language the bilingual subjects were tested on did not have a main effect ($F(1, 38)=2.227$ $p>0.15$), neither was the interaction between language and age group significant ($F(4, 38)=1.36$ $p>0.28$). The only significant main effect was for the age variable ($F(4, 38)=3.61$, $p<0.05$). This indicated an overall improvement in test performance with age for the bilingual children.

6.3.4 Analysis 4: How do Bilingual Children Compare to Monolingual Children – Overall Results

Separate univariate ANOVAs were carried out on the English and German test results comparing the bilingual with each monolingual control groups (in case of the monolingual German group an ANCOVA was calculated with age as a covariate). For the results in English it was shown that only the main effect for age group was significant ($F(5, 68)= 16.39$, $p<0.001$). Neither the school comparison nor the interaction between age group and school reach significance. With respect to the comparisons in the German language of the test, results of the ANCOVA showed that the pattern of the schools did not differ here either and only age group was a significant main effect ($F(5, 65)=3.602$, $p<0.01$). These calculations are in agreement with the graph's first impression, i.e. that the schools show similar patterns and that there is increasing percentage of passed categories with increasing age. However the previously mentioned conosistent pattern of lower scores for the bilinguals (at least in English) might suggest some syntactic delay/deficit.

6.3.5 Analysis 5: Comparison of Individual Syntactic Categories: Are Compound Nouns Processed Easier by Bilingual Children?

As indicated in the introduction, for reasons of cross-linguistic differences between the two languages it was expected that there would be differences in the acquisition pattern of compound nouns and the SVO category between bilinguals and their monolingual counterparts. As such it was hypothesised that bilinguals would find it easier than their monolingual English speaking counterparts to pass the compound noun category, whereas word order would be a more problematic category for these children.

The following figure indicates pass levels for the compound noun category in a comparison of mono and bilingual English:

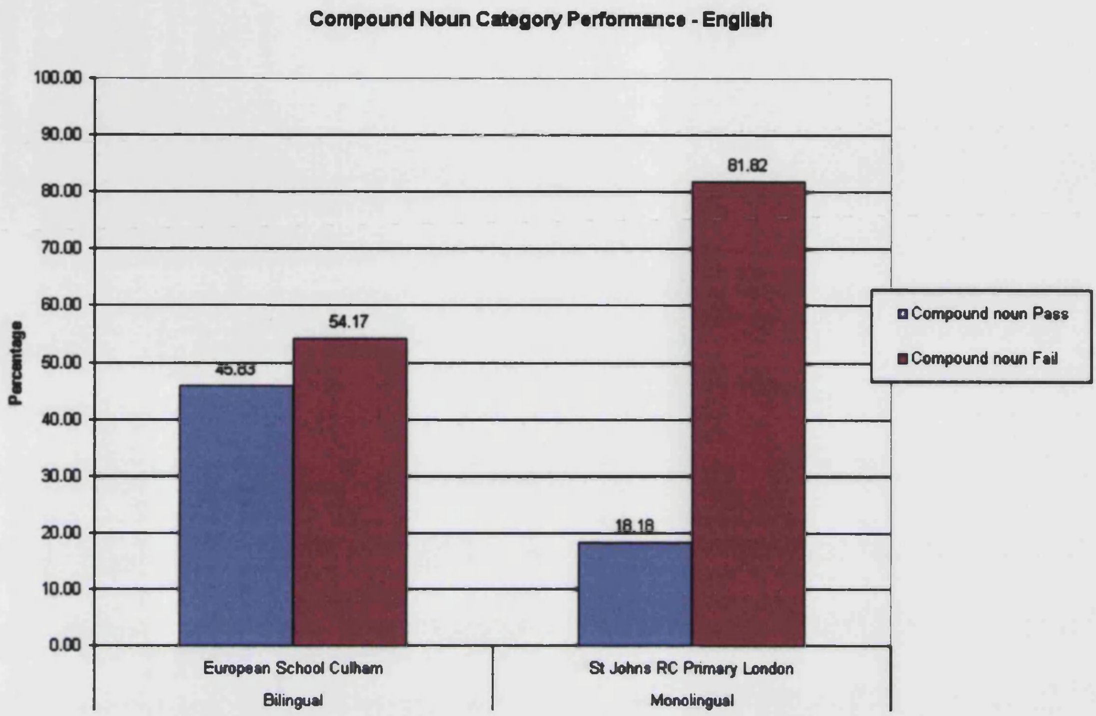


Figure 26: Percentage of bilingual and monolingual children who passed or failed the compound noun category in English.

The figure indicates clearly that this syntactical category was overall a difficult one. In both schools there was a higher failure than pass rate. Since this was categorical data a chi-square analysis was carried out which indicated a significant association between school and pass/fail of compound category (Chi-square test (1) = 5.90, p=0.015). This showed that overall the bilingual children performed better in this category. For the German comparisons the youngest age group was not included since no pupils were in that age group in the monolingual school. There was no significant association between these two categories in a German language comparison, as would be expected.

6.3.6 Analysis 6: Comparison of Individual Syntactic Categories: Is Word Order More Problematic for Bilingual Children?

For the SVO category the pattern looks quite different. Again results are given for performance on the English test first, since it was expected that there would be differences with the bilingual children finding word order more problematic than the monolingual children.

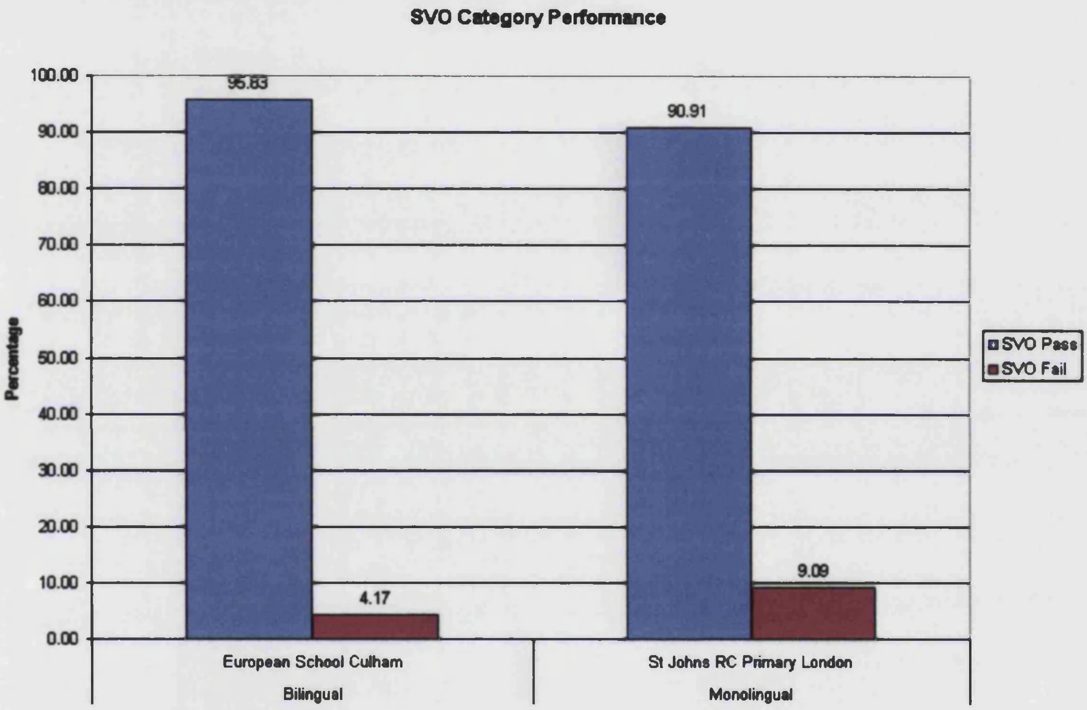


Figure 27: Percentage of bilingual and monolingual children who passed or failed the SVO category in English.

The pattern here clearly shows that the participating children did not have many problems with this category. Pass rate is more or less at ceiling level for both these language groups. A chi-square analysis confirmed this impression (chi-square test (1)= 0.553, p>0.45). The inclusion of a more unusual word order in the German version of the test (object-verb-subject see Table 7 for an example sentence) makes it possible to further analyse bilinguals' performance in word order. Again the youngest age group was excluded from this analysis. The following figure graphically represents performance in this category on the German version of the test.

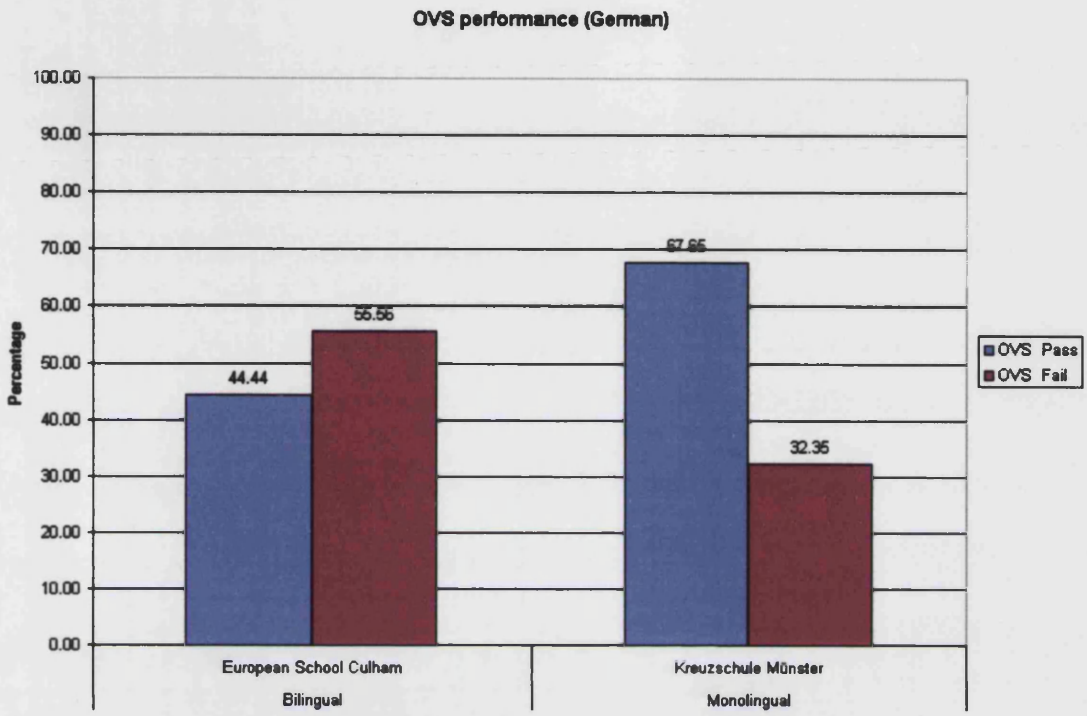


Figure 28: : Percentage of bilingual and monolingual children who passed or failed the OVS category in German.

The figure clearly shows that, whereas the bilingual children had a higher failure than pass rate, the monolingual German children had the opposite pattern with a considerably larger pass rate. A chi-square analysis confirmed this impression (chi-squared (1) = 3.815, p = 0.0508).

The analysis of the individual syntactic categories showed that the bilingual children performed more like their German counterparts for the compound noun category. Whereas in word order they showed a pattern more similar to their English monolingual peers, i.e. being possibly more likely to assign words to the order of subject–verb-object rather than allowing for unusual word order through noun case marking.

6.4 Discussion

Considering the age group results first, it can be stated quite clearly that with increasing age, receptive knowledge developed for all three schools. This means that the German version of the test does test what it set out to investigate, which indicates the internal validity of the test material. Moreover, the similar pattern of monolingual English speaking and monolingual German speaking children strengthens the external validity of the German version of the test.

In respect of the first hypothesis, the analysis of the questionnaire data showed that, overall, the bilingual children in the present study represent a group of bilingual second-language (sequential) acquirers. The majority of children had been born in Germany and had parents who were both native German speakers. As such it is surprising that they do not show evidence of being less competent in the syntax acquisition of English (both when compared within as well as between groups). Thus the first prediction was not corroborated. However, considering that the average bilingual language exposure had been over 4 years for these children and that they are educated in a multilingual environment which has possibly a higher than average focus on language teaching and live in an English speaking country (which means that they most likely have English speaking friends), might explain this result.

The results of the more detailed analysis of individual categories, however, revealed subtle difference between the bilingual and monolingual groups of children, i.e. instances of cross linguistic influences/transfer in school aged children. Regarding hypothesis two, it was shown that in the difficult compound noun category the bilingual children were associated with lower failure rates than their monolingual English counterparts. This implies that the more frequent use of compound nouns in German might have helped these children to perform better in this category. Which provides evidence for the prediction made in the introduction. Furthermore it shows that cross linguistic influence can be observed even when there are only difference in frequency but not in structure (structural influences as described by Nicoladis, 2002; 2003a; 2003b). These results are consistent with the BIMOLA (Grosjean, 1988) model – introduced in chapter five - since it could be assumed that the higher frequency of compound words in German has strengthened the links between certain word units.

When considering word order (i.e. hypothesis three) it was shown that in the SVO category bilingual children performed equally well in this category. However, in the more unusual German OVS word order category they performed less well then their monolingual German peers. This means that in this case they were performing more like their English counterparts where SVO strongly constrains word order. The results for the German OVS category thus confirm the hypothesis. This part of the current analysis seems to suggest that differences in performance are due to language internal factors, i.e. structural differences between the two languages, rather than to language external factors, such as language dominance (Hulk & Müller, 2000). This means that cross-linguistic influence/transfer, compared to the literature introduced in the introduction, can occur at later stages, in receptive rather than productive tasks and even in less balanced bilinguals. Overall two of the three hypotheses set out in the introduction to this chapter have been found to be supported. The language dominance hypothesis was not confirmed. However, in compound noun processing the bilingual children performed better than their monolingual counterparts, whereas in word order they performed worse. The difference in language performance for compound nouns as compared to word order is interesting, since it points to a difference in a lexical, as compared to a syntactic, category.

It can be argued that the analysis of the individual categories has brought up possible criticisms and ways of improving the test instrument. The very high pass rate for the SVO category, for instance, seems to suggest that this category might be too simple for the pupils tested whereas the compound noun category with a more than 50% failure rate, seems highlights the fact that this category was quite difficult. Thus these two categories possibly need to be changed in future investigations. Another concern is that some of the youngest children did not complete the English version of the test. It was felt that the age group they were in, which was falling on the juncture between the two (under- and over five-year old) gross level contingencies of the test, would find the test quite long. This was not helped by the relatively tight testing schedule the school had drawn up.

The implications of the results for the wider aims of the study, concerning the validity of the test, is now outlined. It can be argued that the results have strengthened the case of the validity of the test instrument. First a developmental pattern emerged which gives the German test internal validity. Secondly, the patterns in both monolingual groups were very similar, providing external validity of the test. Thirdly, it was also found to be sensitive to group differences between the bilingual and monolingual children.

Overall results of the current chapter show that the German version (abbreviated to DeROST) can be used over a wide age range, which makes it an appropriate test instrument to investigate aspects of early syntax acquisition. As such it can be used to study fluency development in infancy.

As shown in the first part of the thesis both complex compound words and word order were implicated as difficult processes for people who stutter. The results reported in this chapter show that these two processes also show differences in bilingual children (one of them being superior the other inferenior performance compared to monolingual counterparts). If there was link between stuttering and bilingualism then similar processes should be taxed, i.e. both the compound category and word order should have been negatively affected. This was the case in the people who stutter investigated in the first part of the thesis where more complex content words were stuttered.. However, the better performance in the compound category might suggest that it is not the process of compounding itself (since it is a receptive rather than productive task) but rather the more complex structural nature of these words that causes the speech disfluencies in stuttering.

7 Picture naming and early receptive syntax development in bilingual infants – an investigation into semantic and lexical development

"A child, when it begins to speak, learns what it is that it knows."
John Hall Wheelock, 1963, 'A true poem is a way of knowing', in *What is poetry?*

7.1 Introduction

The argument has been made earlier that understanding why stuttering onset (between ages two and three according to Andrews and Harris, 1964) does not coincide with age of language onset (Bernstein Ratner, 1997) is one of the key questions in stuttering research. Consequently, more information is needed about what aspect of fluency is developing at the age of stuttering onset and if or how the study of bilingual children could shed light on this issue. As previously highlighted throughout the thesis leading researchers in the area of childhood stuttering and onset of stuttering have made the direct causal link between the combination of the vocabulary burst and syntax acquisition as directly leading to stuttering (Bernstein Ratner, 1997; Elbers & Wijnen, 1992; 1990; Wijnen, 1992). In this chapter this issue was investigated within the relevant age group (between ages two and three years – when lexicon and syntax come together) and in a group of bilingual children, a group that has been argued throughout the thesis as being more prone to disfluencies (see chapters one and five). This study examined whether there are children being brought up as bilinguals who produce many errors when naming objects in early language development. Errors that are analysed in this chapter are simple word naming errors – i.e. producing the wrong label for a given picture (in other words not stuttering). The reason for this is as follows. Since one of the subcomponents of EXPLAN is the planning aspect of language, producing the wrong label is seen as an instance where the correct plan has not been made. The investigation of these processes (both in respect to lexicon and syntax) at the age of stuttering onset is the main aim of this part of the thesis. Another example of a failure to fully plan a lexical item is the TOT state which was introduced in chapter one and was investigated in children of different age groups in the next chapter. Subsequent analyses in the current chapter examined the relationship of these scores to other language measures (according to the plan outlined in chapter five). In the following sections, it is discussed a) picture naming and how lexical retrieval may be related to fluency failure, b) mean length of utterance, c) communication development inventory, d) reception of syntax tests (German and English forms).

7.1.1 Picture Naming Errors at Early Stages of Lexical Development

Research on fluent children shows that naming errors of children are likely to occur at the time when there are sudden changes in productive vocabulary size (Gershkoff-Stowe & Smith, 1997). With respect to speech errors, Wijnen (1990; 1992) reported similar results showing that increases of repetition and substitution errors coincided with the time of the main 'vocabulary spurt'. The time period between the age of two and three years was also indicated as the time when a child would have a relatively sudden preoccupation with closed class words and there is a general increase in the production of this word class (Elbers & Wijnen, 1992). This second aspect of grammatical class indicates that children develop not just their vocabulary during this stage, but also that syntax acquisition takes place concurrently. In research into stuttering grammatical class and its development is particularly important, since the closed class words are those that are produced proportionally more disfluently by children at the age of stuttering onset (e.g. Bloodstein &

Grossman, 1981; Howell et al., 1999). See previous chapters for a more detailed analysis of grammatical class and stuttering. Parallels can also be drawn to the demands and capacity framework of stuttering (Starkweather & Gottwald, 1990; Starkweather et al., 1990; 2000) – see section 7.1.4. for more details. In this framework the onset of disfluencies are explained by an increase in the demands onto the linguistic system of the child. This could also be the case in planning processes, i.e. when the demands are high (rapid increases in both lexical and onset of syntax acquisition) then there might be more wrong labels selected.

If the bilingual children show a larger amount of naming errors it suggests that the acquisition of two separate lexicons and the suggested process of simultaneous activity of these two lexicons at any one time in the BIMOLA model (Grosjean, 1988) directly disrupts the planning process of words. This then would have direct links to the processes suggested in EXPLAN (Howell & Au-Yeung, 2001) as leading to disfluencies.

The children in the current study are slightly older than those children in Gershkoff-Stowe and Smith's (1997) study (24 months as compared to 15 months). The reason for this was the fact that vocabulary size is often depressed in bilinguals (as found in numerous studies, such as Ben-Zeev, 1977; Bialystok, 1988; Merriman & Kutlesic, 1993; Rosenblum & Pinker, 1983; Umbel et al., 1992). Although research reveals a great amount of individual variation in lexical development (Fenson et al., 1994), it was assumed that the main vocabulary spurt of the bilingual children would be later than that of their monolingual counterparts. The current study investigated whether there are differences in naming errors as a function of the language the words had been acquired in (i.e. are there more naming errors in the dominant or less dominant language) and also whether naming errors are associated with sudden accelerations in vocabulary growth (and if so in which lexicon).

7.1.2 Oxford Communication Development Inventory

The original MacArthur Communicative Development Inventory (CDI) constitutes one of the most completely standardised measures of early vocabulary competence and has been validated with both children with 'normal' and children with atypical language development (see for instance Fenson et al., 1994; Fenson et al., 2000). Figure 29 provides an indication of the vocabulary as a function of age (in months) for typical lexical development.

For instance it has been recently used to investigate early linguistic development of children with autism spectrum disorder (Charman, Drew, Baird, & Baird, 2003).

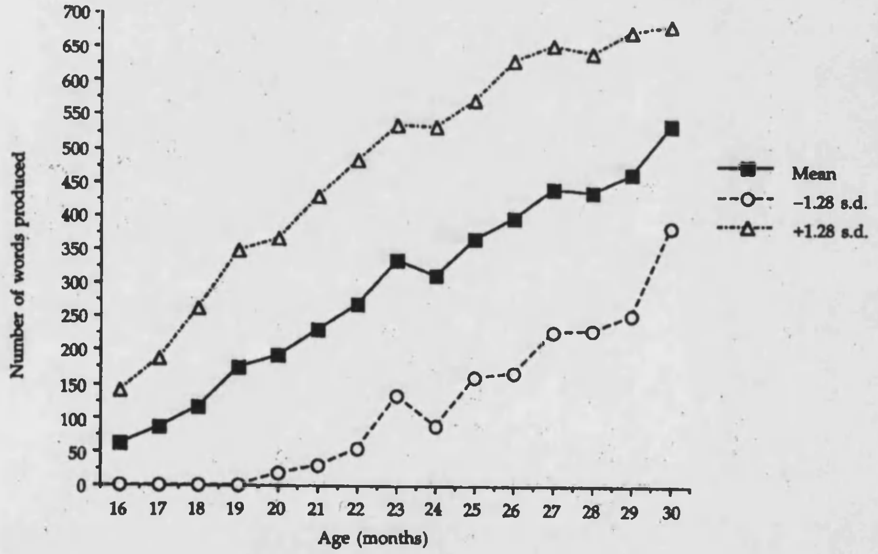


Figure 29:Word production on the MacArthur CDI Toddler Scale from normative data by Bates, Dale & Thal (1995).

Charman and colleagues (2003) found wide variability of linguistic skills in these children (134 children with an average age of 3 years 2 months) and furthermore found group differences characterised by delays in the children with autism spectrum disorder. The Oxford Communicative Development Inventory (Hamilton, Plunkett, & Schafer, 2000) is the UK version of the questionnaire and was adapted for the current project to include productive vocabulary only. Although an official German CDI is available (Grimm & Doil, 2001), it was decided not to use it since the focus of their material was more diagnostic and was slightly wider than necessary here. Another reason for developing our own German version was the fact that the official German CDI was developed with the German child environment in mind, whereas all the children tested in the current project were living in the UK. It is expected that the children have a larger vocabulary in their more dominant language (English) and, linked to this, that when a vocabulary burst occurs it might be less steep in their less dominant language. Furthermore it is assumed that their overall total scores is less than that of comparable monolingual children (as highlighted above by Bialystok, 2001).

7.1.3 Mean Length of Utterance (MLU)

Receptive vocabulary has been investigated in the last chapter and also featured in the current study (see below). As a measure of comparable productive syntactic development, the children's spontaneous speech was collected and analysed according to their mean length of utterance (see Brown, 1973, for methods and scoring of MLUs and method section). It measures the average sentence length, calculated in morphemes. Brown (1973) originally introduced this measure and divided syntactic development into five major stages. Since then it has become a standard tool in language acquisition research and normative data can be used to compare individual children to (Miller & Chapman, 1981a).

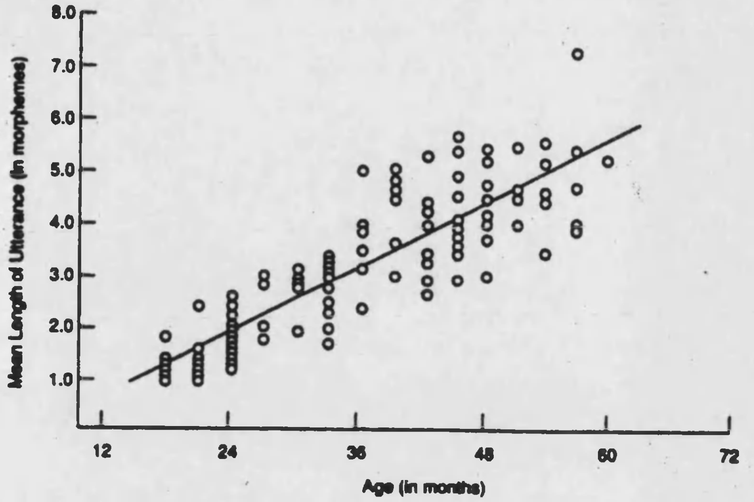


Figure 30: MLU norms – MLU is given as a function of age (in months). Figure reproduced from the normative study by Miller & Chapman (1981).

This collection of data was commenced when the child started using action words and the first function words in his or her vocabulary according to the OCDI parental questionnaire. MLU data, although obtained from a relatively crude measure, is probably the most widely used method for assessing syntax in children near language onset. It can also provide insight into the child's spontaneous use of the two languages and if or how often language switching does occur (more details of the structure and content are given in the method section). It was expected that there is a level of correspondence between the child's MLU and his / her receptive vocabulary knowledge and that MLU would be lower in the less dominant language.

7.1.4 Reception of Syntax Test (English and German)

Research has shown, on the whole, that language capabilities do not separate fluency groups (see for instance, Bernstein Ratner & Sih, 1987; Kadi-Hanifi & Howell, 1992; Nippold, Schwarz, & Jescheniak, 1991). In fact Nippold (1990) concluded, after critical evaluation of earlier research, that the evidence does not support the view that there are considerable differences between stuttering and nonstuttering children on language tasks. Most of this research assessing syntactic abilities of fluent and speakers who stutter has been carried out on children over 5 years old and adults. Bernstein Ratner (1997) reviewing these facts hypothesised that the simultaneous rapid expansion of vocabulary and syntax between ages 24 and 30 (at twice the rate in balanced bilinguals) may trigger retrieval and fluency problems.

The demands and capacities framework (see also Adams, 1990; and for a critique Siegel, 2000; Starkweather, 1987; Starkweather, 1997; Starkweather & Gottwald, 1990; Starkweather et al., 1990; Starkweather & Gottwald, 2000) has been used to highlight the multidimensional nature of fluency development and is a useful way of characterising early bilingual development. According to this model fluency failure occurs when the challenges (or demands such as speech rate, continuity of production etc) exceed the capacities (for instance speech motor control, language formulation, social and emotional maturity, and cognitive skill) of the child. For children who are brought up in bilingual environments – it is estimated that nearly half of the world's population is functionally bilingual (as pointed out by de Houwer, 1995) – the task of lexical and syntactical development is doubled for the two languages acquired. If, as Bernstein Ratner (1997) suggested, increases in vocabulary interact with syntax, leading to retrieval problems, this should be even more likely in bilingual children since they have to learn concurrently two different lexicons and grammars. To phrase it in the language of the demands and capacities model, it could be assumed that children growing up with two languages possess the same capacity for language competence but are faced with higher demands since they have to acquire two lexicons and two grammars. As highlighted throughout some speech therapists have even taken the issue of bilingualism further and have warned that mixed utterances by parents can contribute to the development of a stutter in children with a predisposition to this disorder (Lebrun & Paradis, 1984).

The ROST test (see last chapter) was given to the child on each occasion in one of the two languages only. As well as general grammatical categories, more specific categories that differ between the two languages are included. Bilingual children may, for example, find certain simple categories difficult because of cross-linguistic differences, whereas other more complex syntactical categories might not cause them a problem. One such relatively simple category is the plural marking which appears in both languages relatively early in the acquisition of grammar. This might be one of the categories which is more difficult for English – German bilingual children, because of the considerable differences in how these are marked. The German plural marking has been extensively used in the literature since it is an example of a grammatical case where the exceptions outnumber the rule. It has been used on both sides of the language acquisition argument, as evidence for rule based learning (as highlighted in Pinker, 1999, chapter 'the horrors of the German language') but also as evidence for schema learning (as summarised by Köpcke, 2001). In German plural marking, there are four plural suffixes, with an additional fifth that is phonetically zero. Three of these can be accompanied by vowel changes. It has been argued that seven out of these eight cases are irregular and only the suffix 's' is regular (Wiese, 1996). Mugdan (1977) has tried to put structural, morphological rules, based on the 'Item-and-Process' (IP) model approach (Hockett, 1954), to the German plurals. This attempt resulted in a list of 15 rules and 21 lists of exceptions. Without favouring either theoretical side of the argument, it can be concluded that it is an extremely complicated system that might lead to difficulties in the acquisition process.

7.1.5 Summary

The range of tests should provide insight into object naming processes at the time when the main vocabulary growth occurs and, additionally, give an indication of how or what is causing the bilingual child difficulties in the acquisition of the two grammars. Since the combination of rapid vocabulary acquisition and onset of syntax has been linked to stuttering (Bernstein Rantner, 1997) it was tested here whether these double demands for bilingual children have a direct influence on planning processes (i.e. naming errors). If bilingual children are more prone to disfluencies (see chapter five) then this test could indicate whether this is due to planning processes for lexical items.

The MLU data can also give additional information about the dominant language and whether / how switching between language takes place and how functional items are used.

The same children have been tested for their receptive syntax in both languages. Again certain items are included in this test to investigate cross-linguistic influence/transfer (such as the plural category, the compound nouns and word order see previous chapter). For children who acquire two syntactic systems simultaneously it is assumed again that the task demands are higher and that these demands might lead to certain differences or delays. If such delays do occur it would indicate that the concurrent demands of two linguistic systems hamper certain aspects of language acquisition.

7.2 Method

7.2.1 Participants

Five bilingual English - German speaking children participated in this study. These were four girls and one boy. These five children were aged 23 months (this child is referred to as child C), 24 (child D – the only boy in this study), 25 (child E) and 26 months (identical twins from now referred to as children A and B), at the beginning of the study. For all the children the mother was the German speaker and the fathers were native English speakers. The mothers used consistently German and fathers English when communicating with the children. Two of the children are identical twins (child A and B), and had no other siblings at the time of the study. Another girl (child C) and boy (child D) had no further siblings and the remaining girl (child E) had a younger baby sister. Parental reports showed that none of the children had any hearing or speech problems. For all of the children the mother was the main care-giver. The parents gave informed consent for the child's participation in the study.

7.2.2 Test Material

7.2.2.1 *Communicative Development Questionnaires*

Parents received the Oxford Communicative Development Inventory (OCDI) and a German translation equivalent of this inventory. It was decided that the same categories and words should be included as in the OCDI. The OCDI is separated into nineteen different semantic categories. Ten of these refer to categories of nouns. There is a section each on verbs, adjectives, adverbials (as in words describing times / timing), pronouns, interrogative pronouns, prepositions and quantifiers. A library database of pictures was set up to refer to all the objects in the different noun categories (to be used in picture naming). For each child, based on their first parental questionnaire, 30 items were selected from the library database of pictures (10 that had been learned in each of the two languages only and 10 that were known to them in both languages – see procedure section).

7.2.2.2 Reception of syntax test

The English ROST (also referred to as EngROST) tested nine syntactic categories for the under 5 year age group. ROST categories for the over 5 year olds are only briefly be mentioned here (categories and sample sentences were given in the last chapter). The nine categories for the under 5s were: (1) subject – verb – object word order; (2) pronoun (masculine / feminine); (3) singular / plural; (4) prepositions in / on; (5) article (a / some); (6) compound nouns; (7) pronoun he/they; (8) pronoun his / her and (9) prepositions others such as up / down – these are categories A-I in Table 7. In both tests only simple nouns are used, such as animal names (cat, dog, fish etc), commonly used objects (table, house, bottle etc), and people (boy, girl, child, man, woman etc) as well as simple verbs of simple actions (run, pull, carry etc). The two pictures in each trial are displayed side by side with an 18mm gap between them. Each pair is presented twice (the display of trials is randomised and the side of the picture presentation is counterbalanced) and each picture is the correct answer for one of the two trials.

The German version had both similarities and differences in categories: categories 2 (pronoun: masculine/feminine), 4 (preposition: in/on) and 6 (compound noun) were testing the same syntactic structures, which made them directly comparable. Differences between English and German categories were as follows: since there are more different ways to form plurals in German, these were tested individually in plural morpheme, plural article and plural vowel categories. Subject – verb – object was included but would only use pictures with grammatical genders neuter or feminine (since any masculine gendered object would make word order too obvious). Details and sample sentences for each category are given in Table 7. In German eight categories were presented to the under 5 year old children. These were categories, A, B, C, G, H, I, J and K.

As described in the previous chapter (and noted here again since it is used to reduce testing time in younger infants) in each syntactic category there are 8 different sentences as examples. A category is considered to be achieved if 7 from the 8 are correctly selected (sign test $p<0.05$). If a child failed two test items in a category (which meant that the child could not pass this category to meet the statistical criterion), the child failed this category and no further test items from that category would be presented. If a child passed 7 items from any one category the eighth item would not be tested because the child had already reached the statistical criterion. These restriction reduced test time, which is desirable for children aged under 5, due to the infant's limited attention span (with the added advantage that it did not compromise the power of the test).

EngRost	DeRost
Category A: SVO	Category A: Plural Morpheme
B: Article	B: Plural Article
C: Plural	C: Plural Vowel Change
D: Pronoun his/her	G: Preposition in / on
E: Preposition in/on	H: Preposition other
F: Pronoun number	I: Pronoun he/she
G: Preposition in/on	J: Pronoun he / they
H: Preposition other	K: SVO
I: Compound noun	

Table 8: Syntactic categories tested in ROST in English and German for the under 5 year old children. In English nine categories (A-I) and German eight categories (A, B, C, G, H, I, J, K) are used in this age group – for example sentences see previous chapter.

Children were tested one language at a time, alternating between the languages on subsequent visits.

7.2.3 Mean Length of Utterance (MLU)

Audio recordings were analysed to obtain the MLU data. An unstructured interaction between the researcher and the child was recorded (approximately 15 minutes). Usually books or toys were used to encourage the child to talk. Spontaneous speech samples were then recorded. The digital audio tape recording of this interaction was then orthographically transcribed and MLU calculated according to the guidelines set out by Brown (1973). In Brown's (1973) guidelines a set of rules is given which are used for the calculation of the MLUs – the whole set of rules is too lengthy to be given here, but a few of the details are highlighted in this section. For instance, fillers are not counted (such as 'mm' or 'oh') whereas short answers such as *yeah*, *no* and *hi* are counted. All compound words, proper names and ritualised reduplications count as single words (such as *birthday*, *night-night*, or *see saw* for example). All inflections are counted as separate morphemes, such as possessive, plural, third person singular, regular past and progressive. Recordings commenced at the time of ROST tests.

7.2.4 Equipment

Toshiba laptop computers (models SP 4270 and SP 4290) and OneTouch touchscreen hardware were used. Additionally, digital video recording equipment (Samsung 320X) was used. A DAT recorder (HHB PDR 1000) and microphone (Sennheiser MKH 60) were used to record MLU data. Story books and other props were used to encourage the children to talk. ROST and picture presentation software were programmed in C^{++}.

7.2.5 Procedure for Picture Naming

Prior to the first visit parents were informed about the aims of the study and the test protocol was explained (see below). The parents were then asked to indicate the productive vocabulary of the child in the OCDI and the German questionnaire. It was then determined which words had so far only been acquired in one of the two languages and which were present in both. Then for the first visit, based on the individual child's questionnaire, ten words that had been learned in each language only and ten words that had been known to the child in both languages were selected. These thirty objects were randomised and entered into a presentation program. For the naming task the parent was informed of the following test protocol. If possible one and the same parent should work with the child. Ideally the child was on the parent's lap with the computer within easy reach. The computer keyboard was covered to prevent the child from being distracted. The video camera was positioned so that both the parent and the child were in view. The parent was in control of the test and started when he/she judged that the child was ready. As the picture was presented the child had a few seconds to give a response. If after 5-10 seconds no response had been made, the parent started asking the child about the name of the picture. If still no response was made, then the parent named the picture. Any naming response was followed by a prompt from the parent to touch the screen. This was done to ensure that the child's naming was in response to the stimulus. Unclear verbal responses were confirmed by the parent. These were idiosyncratic words or sounds an individual child would use consistently for a particular object or animal, such as a cartoon character's name for a particular animal. A nod of the head or thumbs up by the parent would then indicate that this was the child's correct name for that picture. This protocol was adhered to as closely as possible, but in the case of the twins it was not always possible to separate them completely for individual testing. In these cases, one twin was tested, and the other was occupied with a game or book and encouraged to wait for her turn. This test was carried out on a three-weekly basis (different random orders were used for each of the occasions). The videos were scored according to certain categories. Whether the child spontaneously said the correct object / animal name (or spontaneously used a different word which would later be scored as a naming error) and which language the child responded in. In cases where the child did not spontaneously respond to the test item, it was noted how often the child was questioned about the object and whether he/she repeated the word when it was provided. These responses would then be entered into a standardised scoring sheet accordingly. For the purpose of the current project only the first category was used in the analysis.

7.2.6 Procedure for ROST

A similar procedure protocol was used for the syntax test. This test was not video recorded because the responses, as indicated by the picture selected, were directly recorded by the program. Depending on each child it could be carried out seated on the lap of the parent or sitting next to the researcher. For both languages there were 4 items that were used as a procedure learning phase. Once the first set of pictures appeared the researcher or parent would ask the child to touch the picture that would go with the word the child would hear. A picture of a loudspeaker appeared beneath the two pictures and the child was encouraged to touch it to start the sentence / word. Whenever the child had selected a picture by touching the computer screen the background colour of the picture changed from blue to red. In the learning phase of the experiment the child was provided with feedback by the experimenter. Once a picture was selected a green arrow would became visible underneath the pictures, which led to the next trial. Trials in the learning phase could be repeated to ensure that the child fully understood the procedure. A child was asked to repeat the learning trials if he or she selected the wrong picture. When the researcher was satisfied that the child understood the procedure the experimental trials started. Experimental trials followed, exactly, the same pattern as the learning trials. In the test trials pictures would only vary on the target syntactic dimension (rather than being different objects as in the test phase). For instance for the English SVO (subject-verb-object) – category for the test sentence 'the girl carries the boy' the two pictures indicated 'a girl carrying a boy' and 'a boy carrying a girl'. The child could not select a picture until the test sentence / word had been heard at least once, but the child was allowed to re-play the test utterance as many times as needed. If a child accidentally selected the wrong picture, but it was clear he or she really favoured the other picture (the child made a statement to that effect, for instance), they were allowed to select the other picture by touching it. The children were encouraged to control the running of the trials by touching the icons on the screen. No feedback was given for the test trials to allow re-testing to track developmental changes.

7.2.7 Behavioural Style Questionnaire

To partial out any effects of temperament on the linguistic task a data using the behavioral style questionnaire (McDevitt & Carey, 1975) was also collected . This is a parental measure in which they are asked about their impression along nine different scales (activity, rythmicity, approach, adaptability, intensity, mood, persistence, distractibility and threshold). For all these nine measures previously established population norms are available for scoring purposes.

7.2.8 Testing timetable

The diagram below gives an indication of test timetable for this longitudinal study.

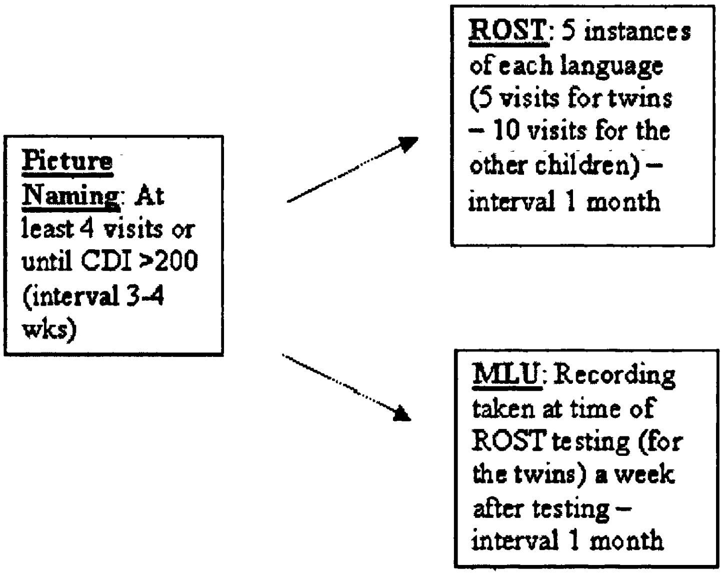


Figure 31: Diagram indicating the testing time-table. Two children were visited nine times (in case of the twins) and the remaining three children fourteen to fifteen times. All at 3-4 weekly intervals.

This diagram is a summary of the tests and intervals that was used. Starting out with picture naming which is then followed usually after 4 three weekly visits with the ROST task and MLU measurement.

7.3 Results

7.3.1 Analysis 1: Vocabulary Size of Bilingual Infants

As previously found there was considerable individual variation in the number of words that were used by each child at the start of testing. The total number of words (in both languages combined) per child are shown in Figure 32. The progression over test sessions is also indicated. In the case of the twins the parents indicated that the girls knew the same number of words. A very similar vocabulary size is likely, since the questionnaire was collected only once in three weeks and when one of the twins would learn a new word the other would hear and acquire it within a few days (according to parental report). Even though there is individual variation it can be seen that one of the children produced fewer items at the outset.

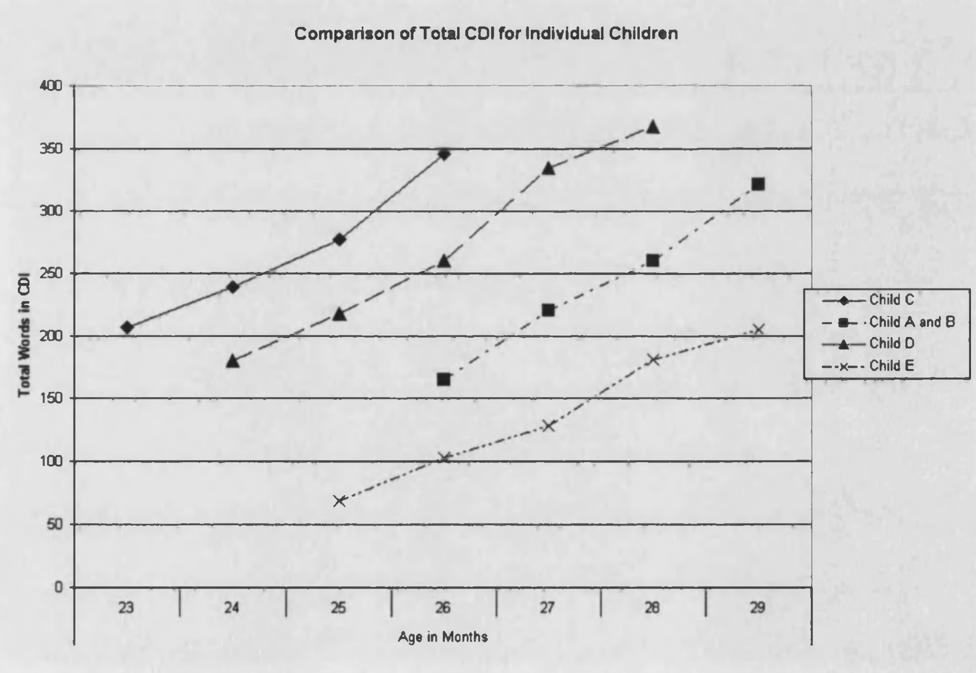


Figure 32: Total number of words (from the Communicative Development Inventory – CDI – in both languages combined) in the productive vocabulary of each child as a function of age (in month).

First of all it can be seen that only one child had over 200 words in his productive vocabulary at the start of testing. This means that these children can be considered to still be in their main period of acceleration in lexical acquisition.

Comparing this figure with the normative data from a large scale study by Bates, Dale and Thal (1995) – see Figure 29 - shows that even though the one child falls below the 1.28 standard deviation, all the others fall within that range (their numbers ranged from 89 to 534 words with a mean of 312 at the 24 month point). However, none of the children reached the upper limit of the distribution (or even the mean). This possibly indicates the effect that Bialystok (2001) was referring to, i.e. that in each of the individual languages vocabulary size would be less than that of a comparable monolingual speaker. Only one of the parents indicated that their child had started talking any later than monolingual children they knew (the one child who showed signs of being below the range). This suggests that the vocabulary size is not just an indication that these children are natural 'late talkers'.

Their vocabulary comprised words known in each language alone and words that were known in both languages from the outset of testing. Whether the vocabulary bursts happened in different rates or occurred more or less at the same rate for each language can be seen in the following figures for each child:

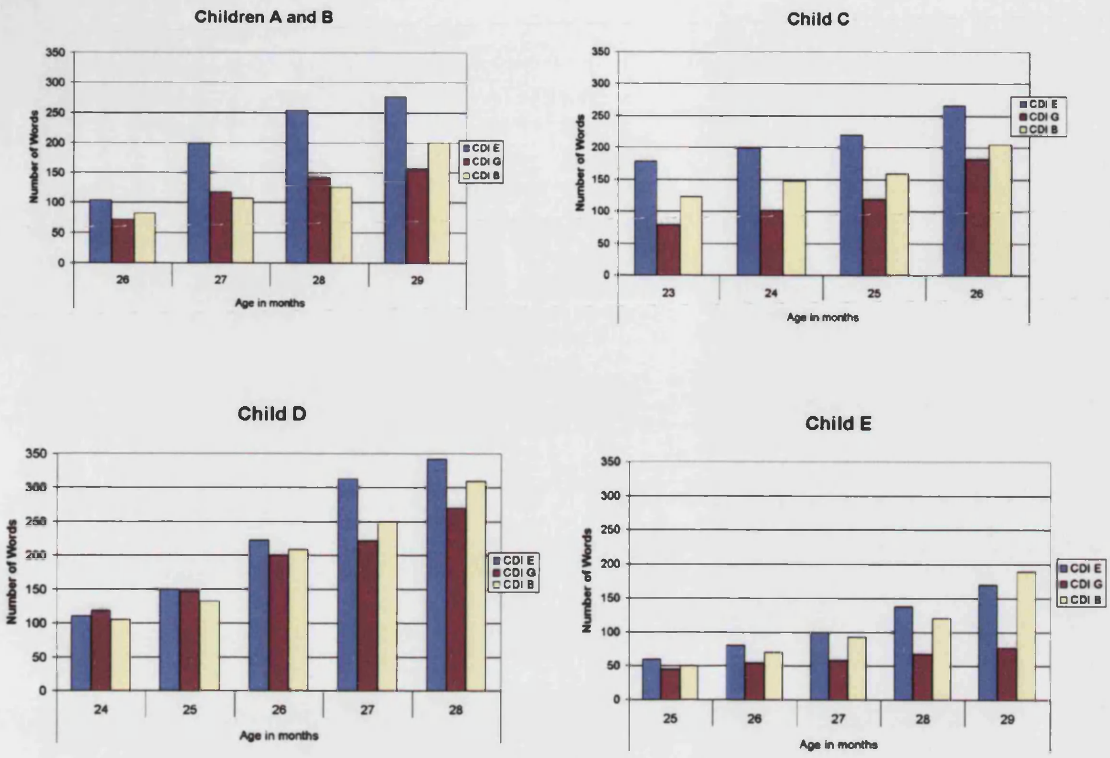


Figure 33: Individual profiles of vocabulary gain according to the different lexicons (as indicated in the differently shaded bars. Age (in month) is given along the abscissa.

These figures show that, even though at the start of data collection the gap between the lexicons was not very large, most children developed their English-only lexicon more rapidly, indicating that this was slightly more dominant. Only one of the children was reported to have a larger lexicon on words known in both languages as compared to the lexicon for English only words.

7.3.2 Analysis 2: Is there Evidence of More Naming Errors Coinciding with Large Vocabulary Bursts?

In the naming task children produced spontaneously object names from both languages; and this was also the case for the words that had been indicated as only being known to the child in one language.

Overall naming errors were only scored when the child chose the wrong word for a particular object. For instance if they focused on a different, particular aspect of the object, rather than the general object name, this was not counted as a naming error – such as *candle / Kerze* to the target word *cake.*

In each of the picture naming sessions not every picture would be named by the child, even though all objects/animals that were selected had been reported to be known to the child and in the productive vocabulary at the time. To make results comparable with each other and to the other children, a proportion of naming errors was calculated that was the total number of errors in a session divided by the total number of correct items named. The results of the total number of errors are shown in Figure 34.

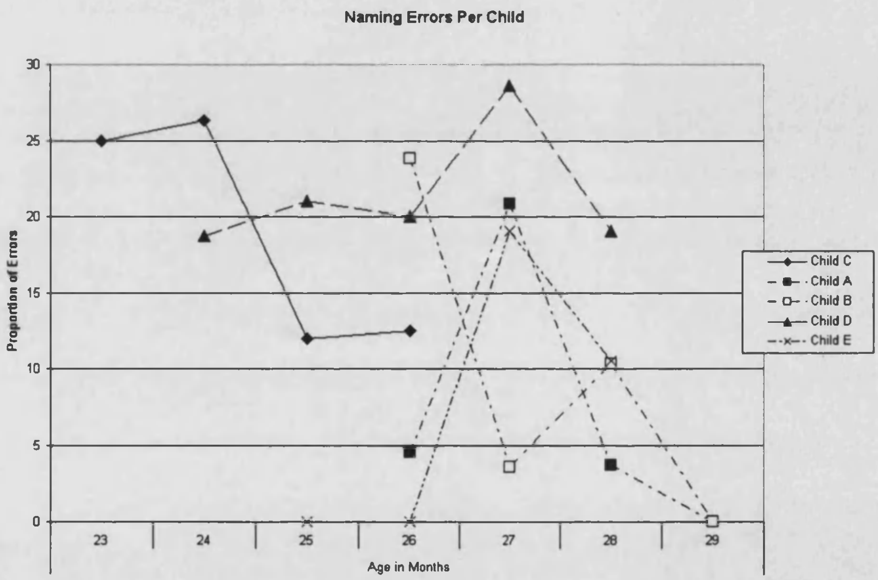


Figure 34: Proportion of naming errors per child as a function of age (in months).

Looking at this graph and comparing it to Figure 32 gives the strong impression that there is no association between acceleration rate of vocabulary and naming errors. This was confirmed by a Spearman's rho for a non-parametric correlation between the gains in vocabulary size from one session to the next and the naming errors in that session. There was virtually no correspondence between those two measures (Spearman's rho -.0.08, p=0.741). This indicates that the amount of vocabulary gained in the time between the two sessions was not a predictor of the number of naming errors the child would make.

It was also investigated whether there were any more or less naming errors in their more dominant language. For this analysis the total number of errors were divided into errors for words that were known to them in only one of the languages. These were then compared with a non-parametric Wilcoxan Signed Rank test. The language in which the child first acquired the word did not make a difference to the amount of errors exhibited (z=-1.53, p=0.126).

Figure 34 also indicates that most children had some completely error-free sessions – i.e. where their performance was at a ceiling level. It has to be pointed out that one of the children started out not naming a large number of objects (correctly or otherwise) and it is therefore not surprising that the number of errors for this child was 0 on the first two occasions. When looking at child A and B the dropping-off of errors towards the right hand side of Figure 34 seems to demonstrate that their performance on the naming task was reaching ceiling level.

After a minimum of four naming sessions, or when the child had acquired at least 200 words in his/her productive vocabulary, testing on ROST commenced.

7.3.3 Analysis 3: ROST Performance in the Two Languages

The following Table gives an indication of the number of passed block per child in each of the two languages.

Visit	*Child A*	*Child B*	*Child C*	*Child D*	*Child E*
1	E - 0	G - 0	E - 0	G - 0	G - 1
2	G - 0	E - 0	G - 0	E - 0	E - 1
3	E - 2	G - 2	G - 0	G - 0	G - 0
4	G - 0	E - 1	E - 0	E - 0	E - 0
5	E - 2	G - 4	G - 0	G - 0	G - 0
6			E - 0	E - 0	E - 0
7			G - 1	G - 0	G - 0
8			E - 3	E - 0	E - 0
9			G - X	G - 0	G - 0
10			E - 5	E - 0	E - 2

Table 9: Number of visits (in monthly intervals), language (E=English, G=German) and passed blocks per child. The X indicates one occasion where the laptop failed to start the program.

Only five sessions were carried out with the twins since they functioned as each other's individual control, whereas to obtain five instances of each language for the other children would take 10 sessions. Just glancing at Table 9, it is immediately apparent that only a few numbers of blocks were ever passed in all the visits. However, it can be seen that there is a tendency to have more blocks passed towards the end of testing. Four of the five children passed up to the highest of five blocks towards the end of testing. A norming study in English showed that children of this age should pass between an average of 1 and 4 blocks – which means that these results are not completely outside the range. However since the number of blocks passed is small, it is impossible to analyse the two different languages separately. Furthermore, since the categories are terminated when two instances are failed (since the statistical criterion of 7 out of 8 cannot be reached), it is impossible to make any more detailed calculations on the trials that were passed. One child never passed a single block of categories and only ever finished the entire test once. On all the other occasions he got distracted and restless and even when breaks were introduced, he refused to finish the test. Other problems with this test procedure are highlighted in the discussion. This might be a matter of the child's temperament, which was also measured and can then be related to the results of the ROST test. The questionnaire regarding temperament were scored according to the standardised rules and the following profiles were obtained.

Child	Activity	Rhythmicity	Approach	Adaptability	Intensity	Mood	Persistence	Distractibility	Threshold
A	3.23	3.56	4.27	2.92	3.67	3.67	2.50	3.50	3.64
B	4.31	4.00	4.27	3.42	3.67	2.83	2.60	4.33	3.91
D	4.31	2.44	2.09	3.17	5.42	3.17	3.30	4.10	3.64
C	3.82	3.3	2.27	3.70	4.33	3.85	2.91	3.64	4.50
E	3.85	4.11	2.27	2.17	4.75	3.08	2.70	3.80	3.18
Population Mean	3.56	2.75	2.99	2.55	4.52	3.31	2.87	3.89	3.98
Population SD	0.75	0.68	0.94	0.72	0.65	0.68	0.69	0.81	0.6

Table 10: Scores on the Behavioural Style Questionnaire for each child. The Population Mean and Standard Deviation are given in the two last rows of the table.

When looking at the temperament scores of the child who only finished the test once it becomes apparent that, possibly, his high level of general activity and intensity (both falling on or above 1 SD above the population mean for that characteristic) might be reasons for this.

7.3.4 Analysis 4: Grammatical Development as Indicated by MLU

As stated previously, MLU data collection commenced at the time of ROST testing. One of the factors that became apparent from the outset of MLU recording was the issue of language dominance. All the children lived in environments where the dominant language was English. In all cases the parents conversed in English with each other and, as was evident from the vocabulary sizes all of the children spontaneously produced more English words from the outset. As such it was found to be impossible to induce the children to speak German spontaneously. This was found to be the case even when the caregiver and researcher were both speaking German with the child. Sometimes occasional nouns and object names would be given in German but, on the whole, all the children spontaneously produced English utterances.

The following Figure shows the MLU for the five children tested.

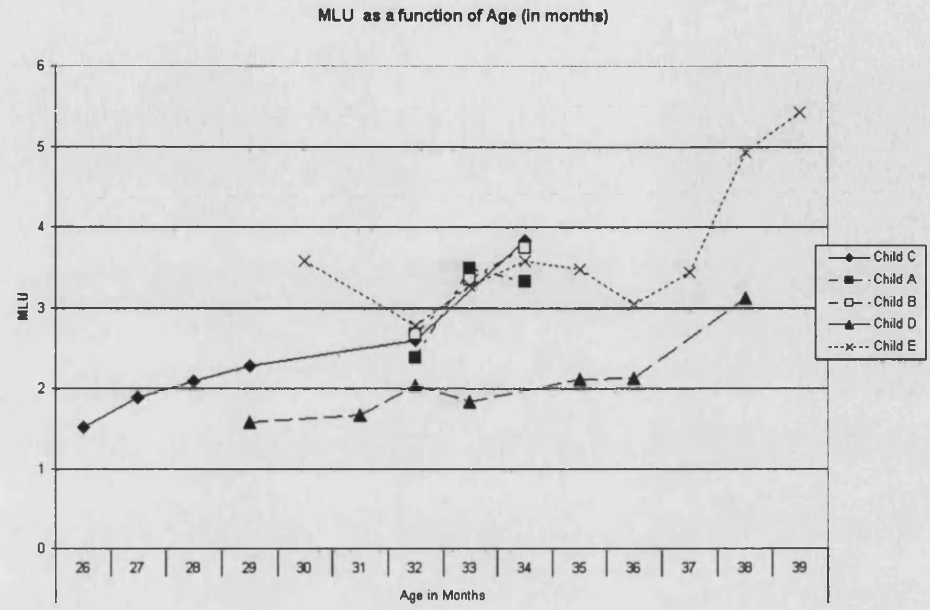


Figure 35: MLU as a function of age (in months) per child.

This, again, highlights the individual variability in language development between these children. Language switching seemed to have been more common in the earlier recording sessions and, mostly, manifested itself by the use of individual content words in German (e.g. *we took a Flugzeug* – airplane). There are a number of interesting points Figure 35 highlights, particularly in comparison to Figure 32. For instance, child E who was a 'late talker' with a comparatively large lexicon containing words known in both languages (see Figure 32 and individual profile in Figure 33), is shown to have the highest MLU of the five children (an MLU of between 5 and 6 at age 39 months is even high within the MLU norms – see Figure 30). Also, she reportedly liked the syntax task and looked forward to the testing sessions. Child D on the other hand, had a relatively large vocabulary size at the outset of the study, but was the child with the lowest MLU (possibly mainly due to his lively nature).

However, all the children fell well within the parameters of MLU norms provided for their age (see Figure 30), i.e. indicating that there is no evidence of a delay in their acquisition of syntax.

7.4 Discussion

Bates and colleagues (1995) assumed that the main acceleration in the rate of vocabulary development occurs at the stage of between 50-100 words. Bilingual children were suggested to have a smaller vocabulary size than their monolingual counterparts and it was shown that some monolingual children well into their third year of life are well below the 50-100 word boundary. Therefore the children in the current investigation were slightly older than the children tested in the original naming error study by Gershkoff-Stowe and Smith (1997). However, it was shown that at the onset of testing all but one of the children were falling in the under 200 word boundary, therefore it can be argued that these children were still in the time of the main vocabulary burst. One observation that was made in the result section is that none of the children were above the average mean word number (of the population norm) in their total CDI word scores. This finding corresponds to the frequently reported smaller vocabulary size of bilingual as compared to monolingual children (see for instance, Ben-Zeev, 1977; Bialystok, 1988; Merriman & Kutlesic, 1993; Rosenblum & Pinker, 1983; Umbel et al., 1992). Individual profiles of the children showed that when vocabulary was divided into the different lexicons (English and German-only words and words known in both), the growth in their dominant language was the steepest at that time. On the other hand, it appeared to be the case that acquisition in the less dominant language took place at a slightly flatter rate.

The second part of the result section investigated the relationship between the gain of vocabulary and the numbers of naming errors that were made in a test session. Figure 34 clearly showed that, even though errors (for each child) did occur in quite a number of sessions, the high points did not correspond to times when the rate of change in vocabulary was higher. In fact the correlation between these two measures was found to be near zero. One possibility is that the children were finding this task too easy and their motivation was not particularly high. Even though the children were more dominant from the outset in English, they did respond in both languages and made naming errors in both languages. Due to the fact that the acquisition rate was shown to be different in both languages, it was hypothesised that the two languages would also show differences in the amount of errors. However, it was found that this was not the case in the current study. A possible future direction of this work would be to look more qualitatively at the specific types of errors that were made.

It was then investigated whether their knowledge in receptive syntax was different in the two languages with the ROST test. The table of results showed that few blocks were passed in any of the test sessions. There are a number of reasons why this may be so. One possibility is that receptive syntactic development is not developing at such a rapid rate in the bilingual infants at such a young age. A reason for this could be that they are trying to differentiate the two languages from each other (as suggested by those researchers who assume this takes place in stages, Volterra & Taschner, 1978), or that their inhibitory processing is still developing (as suggested by Bialystok, 2001).

Other reasons could be more idiosyncratic or procedural in nature. One of them being that differences in individual temperament might influence a child's ability to concentrate and persist with a test. Another observation was that, at times, the children had a preference for a particular position in which the pictures would occur in (choose either only left- or right side pictures). On several occasions, with different children, there was also an issue about them knowing the right answer (because they vocalised it) but preferring the other picture. To demonstrate, in the instances of the plural category it seemed to have been the case that children often preferred pictures which had multiple examples of an item on it rather than just an individual one. It was also the case that the children did seem to prefer a gender. It was noticed on several occasions that there was a preference for pictures with girls rather than boys, or where the girl was the more active part of the picture rather than the boy. These are only observations made at the time of testing, unfortunately it is not possible from the stored responses to statistically calculate these preferences, which is one of the limitations of the current design.

The collection of the MLU data showed that, even though the children seemed to have been on the lower end of the spectrum of lexical acquisition at the outset of the study, they were well within the range of norms given for MLU data. This could mean that a faster acceleration is taking place. Furthermore results in respect of MLU showed that one of the children who was a 'late talker', according to her productive lexicon at the start of testing, had a high MLU on several recording sessions. The opposite pattern was the case for child D, who had a large lexicon but relatively low MLU compared to the other children in the current study (but within the parameters of the norms). This seems to indicate a more general dissociation between the lexicon and syntax.

A dissociation of vocabulary and grammar could indicate a more general aspect of bilingual language processing. Bialystok (2001) highlighted the fact that bilingualism is not a categorical variable and as such different performance measures should be considered and analysed. De Houwer (1995) defined bilingual language acquisition simply as learning two languages from birth (or within a month of birth). The results, reported here, corroborate more with the modular approach taken by

Bialystok (2001) rather than with the strictly categorical position of de Houwer (1995).

The more general impression from this battery of tests seems to indicate that, the language development of these bilingual children is towards the lower end of the normal range at the outset of testing (i.e. particularly in respect of vocabulary size). It is possible that this is due to the demands from the acquisition processes of the two linguistic systems. As was established by Bates and Goodman (2001) the size of the lexicon and grammatical growth are closely correlated. The later MLU data in the current study seems to suggest that they are catching up with their monolingual counterparts (as compared to the monolingual MLU norms and vocabulary size norms which were indicated in the introduction section). In sum the battery of tests seems to show some individual strengths and weaknesses; however the hypothesised planning differences and syntactical problems did not show in the current sample. In this respect the results from this chapter can be seen as relatively inconclusive.

With regards to the hypothesis of the combination of lexical growth and syntax leading directly to disfluencies (Bernstein Ratner, 1997) it was seen that a number of lexical naming errors were made by the children and that they also performed relatively at the lower end of the spectrum of the syntax test. As such the hypothesis cannot be ruled out from the current data for bilingual children.

8 Lexical access in children – an investigation of TOTs in two age groups

"I sometimes hear people who apologize for not being able to say what they mean, maintaining that their heads are so full of fine things that they cannot deliver them for want of eloquence. That is moonshine.... They themselves do not yet know what they mean. Just watch them giving a little stammer as they are about to deliver their brain-child: you can tell that they have labouring-pains not at childbirth but during conception!"
Michel de Montaigne, 1572-80, 'On Educating Children', in *The Complete Essays* (trans. M. A. Screech, 1987), I, no. 26

8.1 Introduction

The previous chapter analysed a range of measures that are indicators of fluency development in mono- as well as bilingual infants. A case had been made that the demands faced by the bilingual infant are doubled in both lexical and syntactic acquisition. Research has also claimed that the higher attentional load for infants who hear a lot of mixed utterances from one or both of the parents, can contribute to the development of a stutter in children with a predisposition to this disorder (Lebrun & Paradis, 1984). As emphasised in chapter 5, Bernstein Ratner (1997) suggested that the interaction of the increases in vocabulary and syntax could lead to retrieval problems. Word retrieval problems were initially introduced in chapter one when the tip of the tongue (TOT) phenomenon was discussed as a way of studying this, and implications of TOT research for theories of speech production and fluency development were highlighted. Some of the details relevant to the current chapter are described again below.

In this chapter a TOT study with German speaking children is carried out. The primary focus of this work is to analyse the feasibility of the technique for the use with young children. This could then become a further tool to track the fluency development of the children tested in the previous chapter and to investigate lexical retrieval in young children who stutter. In the last chapter lexical retrieval was already introduced within the picture-naming paradigm where it was analysed with naming errors. Participants in the current work are German speaking children in two age groups.

Numerous studies have tried to establish what type of partial information is available to speakers in TOT states – usually adult speakers (see Brown, 1991; and Schwartz, 2002, for comprehensive reviews) . In the original work, Brown and McNeill (1966) reported that speakers in TOT states can access phonetic information, such as the main stress pattern, first phone and number of syllables. Regarding syntactic information it has been shown, for example, that when experiencing TOT states English speakers can report count (whether a word can have a plural morpheme) or mass (words that cannot have a plural morpheme attached) information and Italian speakers a word's gender (Vigliocco et al., 1997; Vigliocco et al., 1999). There has, however, been debate about the interpretation of these findings in terms of what lexical stages they represent. A brief review of evidence that shows syntactic information is available in the TOT state is given and a short summary of the debate on the implications of these findings for the form of lexical representation is presented. This has relevance for the current chapter since the German study probed for noun gender (male, female, neuter). As the study uses child participants, a review of the (limited) research findings of TOT studies using child subjects then follows.

8.1.1 Syntactic information in TOT states

The background to the investigation of TOT states and how these have been studied previously has already been highlighted in chapter one. One important issue to recap here is the difference between positive and negative TOTs. This distinction is as follows, if subjects report that the word they are shown after a trial was the word they had in mind (*positive TOTs)* if it was not the word they tried to retrieve it is classified as a *negative TOT*. This is important since negative TOT instances are used as baseline comparison cases – which is also the case in this chapter. Previous research indicated that partial syntactic knowledge such as, for example, grammatical class, gender, auxiliary type and count / mass is available to speakers in TOT states (see for instance Vigliocco et al., 1997; Vigliocco et al., 1999).

8.1.2 TOT and Children

The four only studies of TOT states in children have already been introduced in chapter one (Butterfield et al., 1988; Elbers, 1985; Faust et al., 1997; Wellman, 1977). Recall that two different measures were used in this line of research. In contrast to the TOT state the feeling of knowing (FOK) is measured in some of these studies. Discussing these two different measures, Schwartz and colleagues (Schwartz et al., 2000; Schwartz, 2002) emphasised that they need to be clearly separated. TOTs occur at a later stage in retrieval, are involuntary and are predictors of recall. The FOK on the other hand is a predictor of recognition and is as such often used in studies that investigate met-cognitive skills. Only TOTs would contain phonological information whereas the FOK is a vague feeling of having seen or known that word at one stage.

Of the four studies researching lexical retrieval in childhood, two analysed FOK (Butterfield et al., 1988; Wellman, 1977) rather than the TOT phenomenon, and the remaining study by Elbers (1985) is mainly anecdotal in nature.

Two studies reported exactly opposing results. Wellman (1977) investigated metacognition in children in two age groups (six and eight year olds). This research used recognition rates of FOKs as a measure of accuracy. Rates were higher than chance level, and older children were significantly more accurate than younger children. Butterfield et al. (1988) however investigated developmental aspects of metacognitive accuracy (as measured by FOK recognition) in six, ten, eighteen and seventy year old subjects. In exact contrast to Wellman (1977) they reported that FOK accuracy actually decreased after age six.

Elbers (1985) reported that her son (aged two and a half years) on several occasions produced similar sounding words when he was in a TOT state. Elbers (1985) interpreted these findings as indications that partial phonological information such as syllable number and stress pattern is already available at such a young age.

Faust et al. (1997) conducted the most recent TOT study with children. They investigated TOT rates and accuracy in normal and language impaired seven to eight year old children. The procedure involved pictures of animals and objects and asked the children to name them. It was reported that language- impaired children indicated a higher rate of TOT instances, but percentage of resolved TOTs was much lower compared to the control children. These children also recognised fewer of their TOTs and recalled more incorrect phonological information.

8.1.3 Aims and procedure of the current study

TOT states are relevant in the context of this thesis since they represent cases of an acute unavailability of plan. This planning failure thus links directly to the EXPLAN model of fluency failure. Another parallel to stuttering is that in many cases the beginning of words are available in TOT states (i.e. people in TOT states can name the first phone of the word at higher than chance level – see Brown and McNeil, 1966). In stuttering it is most often the first phone that is prolonged or repeated which indicates that this is the part of the word that is already planned and ready for execution. Since there is such a general lack of TOT studies with children it is aimed in this chapter to investigate which type of information is available to children in different age groups when in TOT states. It is also investigated whether some factors that affect stuttering rate influence ease of lexical retrieval for children (such as word length and phone the word starts with).

The present study uses a similar methodology to that in previous studies, presenting pictures and definitions / descriptions to under 7-year old and over nine-year old German speaking children. In the previous chapter picture naming was described and used as one paradigm to examine early language production (word retrieval) and possible naming errors. With increasing age the types of language tests a child can perform changes. Subsequently more information about their language development and more general development can be made. There are a number of reasons why the TOT paradigm in particular, could constitute an avenue to further investigate fluency development:

Pictures and definitions can be used as TOT-probes that would directly link it to the picture naming experiment of the previous chapter. It could then be clarified whether children who had naming errors would also have more TOTs later in their language acquisition.

With regards to children's lexicon two issues need to be considered. A) The vocabulary of children is growing and B) the children have used words infrequently relative to adults. Each of these observations alone would allow predictions to be made about lexical retrieval during the time when lexical items are still acquired. Note, though, that both the growth of the lexicon and the frequency of usage interact during early language development. The influences of these two factors are to some extent in opposite directions, and as such the nature of this interaction needs further specification and investigation. Researchers have suggested (see for instance Barrett, 1995; Clark, 1993; Fenson et al., 1994) that rapid increases in vocabulary size lead to a widening gap between the child's productive vocabulary (i.e. the words a child would frequently use) and the same child's comprehension lexicon (i.e. the words a child would understand but would not produce). An adult's comprehension lexicon is around 85,000 words of which only 20,000 to 50,000 words are included in the production lexicon (see Clark, 1993; 1995). A gap between these two vocabulary types is one of the reasons for lexical retrieval failures. Do children, therefore, experience fewer TOTs, since their vocabularies are smaller and the gap between the two vocabulary types is not as large as in adulthood?

Language production theories generally agree that the links between stages / nodes in the retrieval process are strengthened through repeated use and association (see for instance Dell, 1986; MacKay, 1987; Roelofs, 2000). As such an investigation with younger children would show whether they experience more TOTs as compared to older children and adults since they have weaker links between semantic and phonological stages / nodes of lexical access.

Some methodological considerations should be briefly highlighted here. There are two main issues that make the paradigm used in the current study difficult to adapt to younger age groups as compared to the adult procedure. One concerns the fact that the vocabulary size of children is not as large as the adults' mental lexicon. With a larger vocabulary size there are more possible words that would be likely to produce TOT states. In the case of children a smaller vocabulary size leads to fewer possible words that would generate this phenomenon. Another important issue in the work with children is the motivational aspect. Both too-easy and too-hard test material would be de-motivating for children, particularly (but not exclusively) at younger ages. Too-easy would be especially problematic because not enough TOTs would be produced. However, too-hard would prevent children from holding their attention. The balancing out of both these factors means that a lower level of TOT instances per word is to be expected. The motivational issue also indicates that there is a need to prepare different sets of stimuli for the two age groups, since nouns for the younger age group would be too easy for the older group and vice versa. This has the effect that some results of comparisons might be due to the different stimuli sets used. However, different linguistic word factors (phone the word starts with, gender of the word and length) were standardised across groups which makes interpretation of these factors easier.

The current study aimed to investigate how much phonological and syntactic information is available to children in two age groups. Details about syllables in initial position and, in addition, gender were probed. Gender is included since it has previously reported to be available to Italian speakers (Vigliocco et al., 1997). The TOT picture-naming paradigm is going to be used with two groups of children. These children have not been tested in the current thesis work previously. This project is seen as preliminary work which is used for the continuation of the work in the previous chapter and extending to failures of lexical retrieval when the bilingual children are old enough for TOT-tests. As such it assesses whether young children can report TOT states and how much phonological or syntactic information they might have available. It is also studied how, or if, rates and accuracy of TOT change over age groups.

8.2 Method

8.2.1 Participants

Twenty-one pupils of the *Deutsche Schule London* took part in this experiment. Of these ten were male and eleven female. All had German as their first and dominant language. Their ages ranged from 9 years 6 month to 11 years 10 month with a mean age of 10 years 4 months. Ten further children participated in a play group in Germany (five girls and five boys). Their ages ranged from 3 years and 5 months to 7 years and 4 months (mean age of 5 years and 9 months). Parents were informed of the aims and procedures of the study and gave their consent for their children to participate.

8.2.2 Apparatus

Stimuli were presented on a laptop computer. The computer program running the presentation of the stimuli was written in C++ programming language.

8.2.3 Material

Words were selected from the German part of the Celex Lexical Database (Second Release August 1995). The word frequency measures given were based on both written and spoken sources. German definitions were taken (and when necessary adapted, i.e. shortened or simplified) from Mayer's online lexicon (http://www.iicm.edu/meyers) and the *Ravensburger Lexicon der Natur & Technik* (1994). Pictures representing the selected words were found on the internet and, when necessary, were resized. A selection was made from nouns below a frequency of 50 and of 100 from the Celex catabase and below 50 for the older children and 100 for the younger age group. A number of word factors were selected. There were equal numbers of words starting with vowels and consonants. These were selected from words of different syllable lengths. The different levels of the word length factor were one, two, three syllables. Noun gender (male, female and neutral) was represented in equal numbers in the set. For each age group there were 54 words with definitions. A complete set of the words and their definitions are given in Appendix 3.

8.2.4 Procedure

The participants were tested individually, in a quiet room at the school or in the play group. TOTs were briefly described and example situations given to illustrate the meaning of the concept to the child. The child was seated in front of the laptop monitor. The computer was operated by the experimenter. Prior to the start of the trials the concept of syllables was explained to the child and examples of different words with different syllable lengths were given on the computer. For each stimulus trial the child would see a picture on the laptop monitor and the experimenter read out the definition for this picture, which appeared below the picture. If the child could name the object or animal, the word given by the child was entered into the computer by the experimenter, consequently naming latencies were not made. Whenever the child could not retrieve a word but showed signs of knowing (such as 'yes we had that in a lesson' etc) one of the buttons on the screen (entitled 'I sort of know') was clicked and the child was encouraged to provide guesses for a number of questions. These guesses referred to questions about the sounds at the beginning / end of the word, number of syllables and gender of the word. The answers were entered into the computer by the experimenter. Then the 'OK' button was clicked and the answer to the picture / definition would be given. The child was asked to indicate whether that word was the word that they were thinking of. Once they had answered this question the next trial was initiated.

8.3 Results

8.3.1 Analysis 1: Are older Children Becoming More Accurate?

Generally the children reported that they enjoyed the test / test material and, apart from one child in the younger age group, all children completed it.

Table 11 shows the pattern of responses in each age group – broken down into total, positive and negative TOTs and instances where the child did not know the answer.

	younger children (3-7)		older children (9-11)	
	Mean	Std Deviation	Mean	Std Deviation
total number of TOTs	5.70	3.09	7.52	3.92
positive TOTs	1.60	2.46	4.05	2.84
negative TOTs	4.20	2.74	3.52	2.02
don't knows	4.50	4.67	15.29	6.99

Table 11: Average response number per child in each age group (younger age group N=10, older age group N=21). The response values are separated into total number of TOTs, number of positive TOTs, number of negative TOTs and instances where the child did not know the correct answer.

From reading Table 11 it can be seen that children were quite willing to report that they were in a TOT state – indicated by the average of more than 5 reported TOTs per child (and more than 7 in the older age group). When the total instances are broken down into positive and negative TOTs it can be seen that, on average, the younger children reported more negative TOTs than positive, whereas the opposite was true for the older children. Size of vocabulary and motivational issues, as pertaining to the methodological paradigm, have already been mentioned in the introduction. The large difference in numbers of 'don't know' responses should be highlighted in this respect. Older children, with larger vocabularies, allowed for a wider selection of, and also less frequent, items than were presented to the younger age group.

Difference of group size for the two ages and the methodological issue meant that the majority of statistical calculations were carried out using non-parametric tests. This was calculated using both Mann Whitney and Wilcoxon signed rank tests.

The relationship between the number of positive and negative TOTs in each age group is graphically represented in Figure 36.

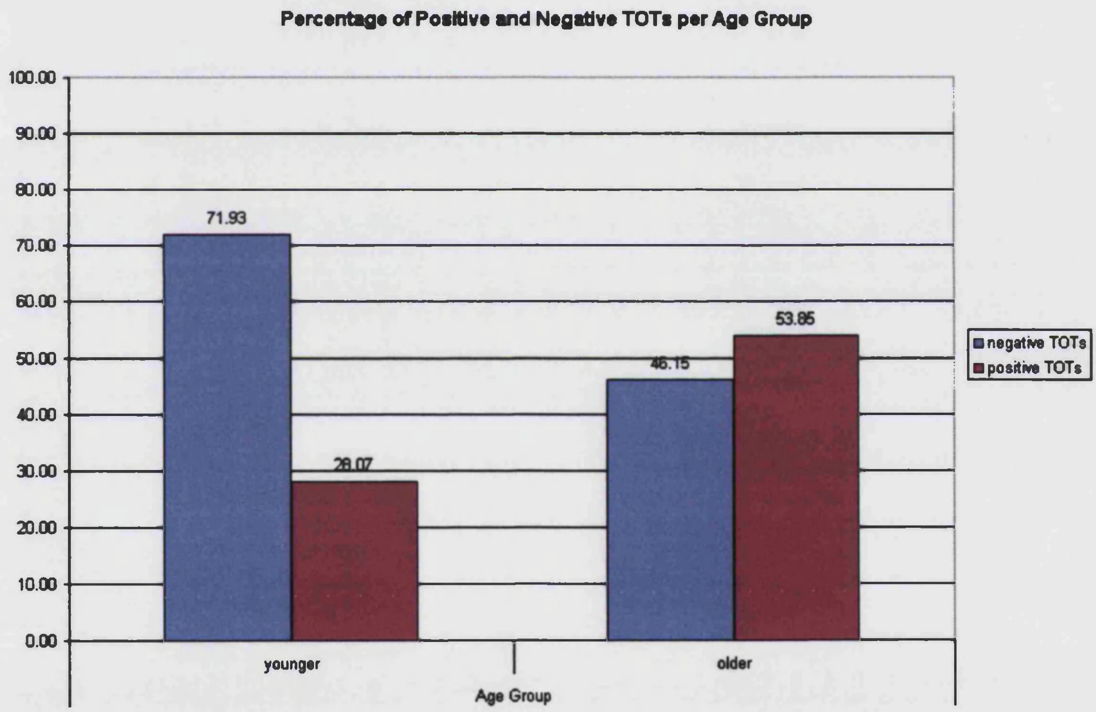


Figure 36: Percentage of positive and negative TOT instances per age group (younger children N=10, older children N=21).

Rather than a test on the frequencies the nonparametric Mann Whitney test used the raw TOTs reported per subject and confirmed that the average number of positive TOTs reported in the older age group was larger than the number average number reported by the younger age group ($U(10, 21)=42$, $p<0.01$). The average number of negative TOTs was, however, not found to be significantly larger for the older age group ($U(10, 21)=91$, $p>0.5$). When repeated measures analysis was carried out for positive and negative TOTs within each age group it was shown that for the younger age group seven out of the ten reported more negative than positive TOTs, a difference that was approaching significance (Wilcoxon test approx. $z=-1.846$, $p=0.065$). For the older children there was no difference between the number of positive and negative TOTs reported (Wilcoxon test approx. $z=-0.629$, $p>0.5$). In sum with increasing age the number of positive TOTs seems to increase and there is a trend for younger children to report more negative than positive TOT states.

8.3.2 Analysis 2: Can Children Report Gender in TOT States?

To see whether children could report the noun gender of a word correctly, a different calculation was made. For each child the number of correct noun genders was counted according to their positive or negative TOT status. From these numbers proportions of correct gender in both categories could be calculated. It was hypothesised that noun genders in positive TOTs would be reported correct above chance level. Whereas noun gender in negative TOTs should be at chance level. The proportions of correct gender in both categories were calculated separately for positive and negative TOTs using a one sample t-test comparing it to the value of 33.33% (due to the fact that there are three noun genders in German). Using this procedure it was found that correct noun gender was only above chance level in the older children for positive TOTs (the proportion of positive noun genders was 78.45% - $t(20)=7.74$, $p<0.001$). None of the other proportions were significantly different from the chance value, i.e. correct gender in negative TOTs for the older children as well as both correct gender in positive and negative TOTs for the younger age group.

8.3.3 Can children report the correct beginning, ends and syllable number in TOT states?

A similar procedure was used as in the previous analysis. The number of correct beginnings, endings and syllable numbers were recorded per person and separately for positive and negative TOTs. Since the overall number of positive and negative TOTs varied correct beginnings, ends and syllables were calculated as a proportion of the total number of positive and negative TOTs. The following proportions were obtained for each age group:

	younger children (3-7) N=10			older children (9-11) N=21		
	Mean	Median	Std Deviation	Mean	Median	Std Deviation
correct beginning positive TOT	.03	.00	.09	.32	.20	.44
correct beginning negative TOT	.10	.00	.32	.02	.00	.06
correct ending positive TOT	.03	.00	.09	.32	.25	.35
correct ending negative TOT	.00	.00	.00	.00	.00	.00
correct syllable positive TOT	.34	.10	.43	.63	.50	.52
correct syllable negative TOT	.27	.23	.31	.18	.00	.26

Table 12: Proportion of correctly guessed beginning, ends and syllable numbers per positive and negative TOTs for each age group.

It can be seen that in the younger age groups these values are very small – apart from syllable number which was correctly guessed a third of the time (both in positive and negative TOT states). In the older age group these values are a lot higher. Correct beginnings and ends are reported in on average one third of all positive TOT states and correct number of syllables even close to two thirds of the time for this age group. These values were then analysed by repeated measures non parametric Wilcoxon analysis for each age group. For the younger age group the differences for correctly reported beginnings, ends and number of syllables between positive and negative TOT states were not significant (Wilcoxon signed rank test approx. z=-.447, z=1.00,z=-.560 all p values >0.3). These difference for the older age group were all found to be significant (Wilcoxon signed rank test approx. z -3.071, z= -3.3080,z= -3.377 all p values < 0.01). This again shows that with increasing age the children become more accurate not just in the reporting gender (see previous analysis) but also can report correctly more often than chance beginnings, ends and syllable numbers.

8.3.4 Analysis 3: Are There any Effects of Word Sound (Phoneme Onset) and Word Length?

Since the numbers are already small, they become even smaller when categorised into different linguistic classes (length and phoneme the word starts with). To analyse these two characteristics it was thus decided to concentrate on the positive TOTs only and collapse the data across age groups. The relationship between these two variables can best be represented graphically. The frequencies of positive TOTs were separated into the four cells of the table, i.e. long and starting with a vowel, long and starting with a consonant, short and starting with a vowel and, finally, short and starting with a consonant. The contingency between these two variables is shown below - see Figure 37.

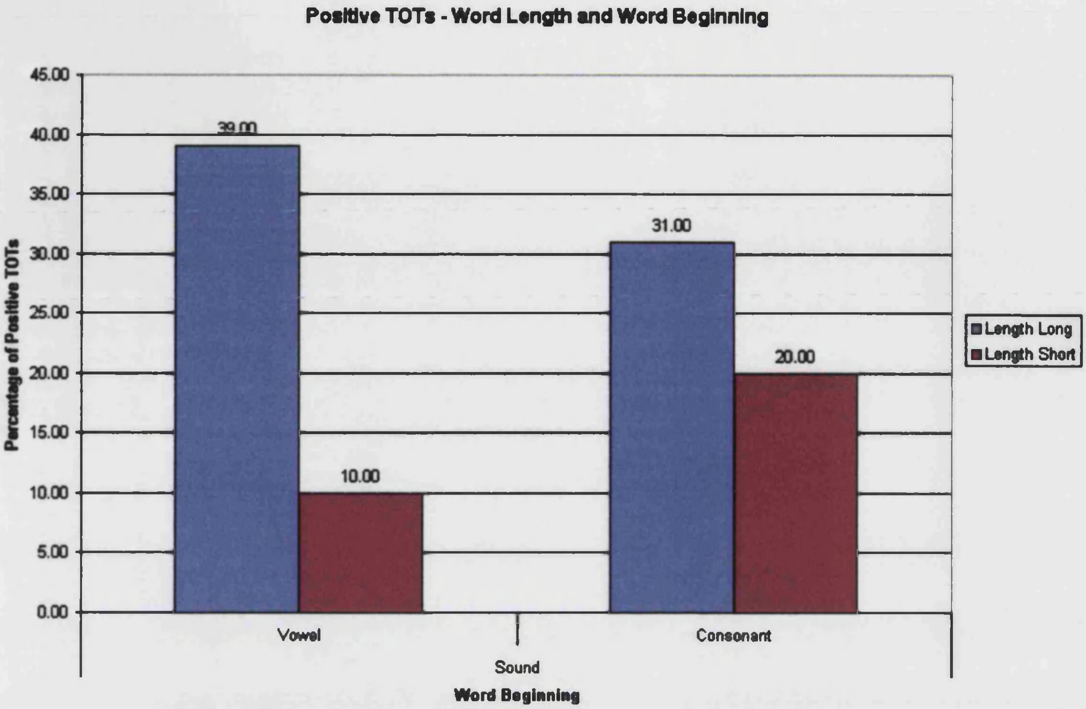


Figure 37: Percentage of total positive TOTs starting with vowels or consonants and being either short (one and two syllables) or long (three syllables) – notice that the numbers presented are percentages of total whereas the previous graph showed percentages within each age group.

The relationship between these factors was analysed using the nonparametric Wilcoxon signed rank test. But whereas the figure is showing frequency percentages the tests are carried out on the raw number per subject. The figure clearly shows that there seems to be an effect of word length. This was confirmed in the analysis (Wilcoxon signed rank test approx. $z=-3.564$, $p<0.001$) – equivalent to the 70% compared to 30% difference. The figure also indicates that the combined vowel and consonant percentages are very similar (49% and 51% respectively) – the Wilcoxon signed rank test on the raw number of positive TOTs for words starting with vowels and consonants was not found to be significant (Wilcoxon signed rank test approx. z approaching 0 and p approaching 1).

8.4 Discussion

Even though there are the motivational and vocabulary size issues to consider, it can be seen from the results that the paradigm does achieve the aim of inducing TOT states in children. As such the goal of devising an avenue to research fluency development with this paradigm and compare it to lexical retrieval failure in the earliest word production stages (the technique with the children used in the previous chapter) has been achieved.

The means showed that, collapsed over all reported TOTs, the younger children showed similar incidence of TOT states (5.7 and 7.5 on average per child for the two age groups respectively). However, there was a difference when TOT states were broken down into positive and negative instances (for the younger group: 1.60 positive and 4.20 negative and the older children: 4.05 positive and 3.52 negative TOTs). This was confirmed in the first analysis showing that the accuracy of TOTs increased with age (rather than Butterfield et al., 1988 concerning metacognitive improvements in FOK; i.e. corroborating with Wellman, 1977). Wellman (1977) had found a similar increase of TOT accuracy in his study whereas that had not bee the case in the investigation by Butterfield and colleagues. Younger children showed a trend to have more negative than positive TOTs, whereas the results of the older children seemed to suggest that they could discriminate positive and negative TOTs more clearly. The number of "don't knows" might also be mentioned in this respect. It is possible that the younger children indicated more frequently that they were in a TOT state when they actually did not know the word – possibly because they thought that was the aim of the test, or because they did not want to admit not knowing (i.e. the demand characteristics of this particular task). The considerably higher average number of 'don't knows' for the older children might indicate that they could already, better distinguish when they were in a TOT state as compared to when they did not know a word.

It was also shown that the older children could guess more correctly the gender, word beginnings, ends and syllable numbers when in positive TOT states. This again shows that with increasing age accuracy increases.

Some limitations of these results should also be emphasised here. The fewer numbers of 'don't know' responses in the younger age group could also indicate that their test words were a lot simpler. In that respect they might well be more motivated, but it might under-represent the number of TOTs in comparison with the more difficult words for the older children. This might well be a reason for their higher negative TOT responses, i.e. because most words were easy they did not want to admit that they did not know a word. A connected issue is the problem of frequency databases in case of child language. The numbers recorded are solely based on adult language usage, both written and spoken: for example, both the German words 'Esel' (donkey) and 'Justiz' (judiciary) have roughly equivalent frequency rates. As such they are misleading for child usage. Even though this fact was taken into consideration when selecting the nouns, it could have led to the inclusion of some items that were either too easy, or too hard, for both groups.

However, the pattern of increased accuracy is also confirmed in the analysis of gender assignment to nouns in the positive TOT states. Correct gender is assigned from the earliest stages of the lexicon, for instance even when acquiring two linguistic systems with two different word genders these are correctly assigned as early as age two (Müller, 1990). Even though gender assignment acquisition has been reported to be a problem for German speaking children elsewhere (see for instance Szagun, 2004 though this report is mainly referring to case marking rather than strictly gender) in view of the other findings for the younger children it seems to be an accuracy increase with age. In the current study only the older children could assign noun gender correctly in their positive TOT states. That the accuracy of correctly specified noun gender was at chance level for their negative TOTs indicates also that the words were not simply biased for a specific gender. The finding of the ability to correctly identify a word's gender when in a TOT state corresponds to the results by Vigliocco et al. (1997) for Italian speaking adults.

The last part of the analysis ties the results of the current investigation in with issues highlighted in previous chapters. The fact that longer words showed more positive TOT instances indicates that these seem to be more difficult to retrieve. This fits in with the issue of content and function word processing. The fact that content words are usually longer and, specifically, the issue of the high amount of compound nouns in German might explain fluency failures in respect of retrieval failure for these words. The overall number of vowel and consonant proportions of overall positive TOTs was not that different from each other. This again suggests, for instance, that consonants compared to vowels is not as big a difference in complexity in German as compared to English. This was a characteristic discussed in depth in chapter 2 of this thesis.

For future research the test procedure would benefit from closer inspection of the words and their responses in the current test group. Words that were always correctly recognised should be replaced by more difficult items. Words that receive a majority of 'don't know' responses should be replaced by easier nouns. Obviously a larger sample size should also be used in both age groups to induce a higher number of TOT states, which would also make direct parametric comparisons possible.

This chapter established that the TOT paradigm is a useful technique to investigate lexical access, particularly in the older age group of children. This makes it a particularly good research tool to investigate populations of children where retrieval differences may be expected. For instance, it would be interesting to investigate whether the slight delay shown by early bilingual infants (see previous chapter) in vocabulary acquisition also has an influence on lexical retrieval later on in language acquisition. Furthermore it could be a test instrument to investigate lexical retrieval in children who stutter.

In relation to the EXPLAN model of stuttering (Howell& Au-Yeung, 2001) the study of lexical retrieval problems is important since TOT states represent acute cases of when the plan for a word is not ready for execution. As such it could parallel cases where planning and execution are not in synchrony as in stuttering.

9 Summary and Implications

[Arthur] 'Ford' he said. 'Yeah?' 'What's this fish doing in my ear?' 'It's translating for you. It's a babel fish.'
Douglas Adams, 1979, *The Hitch-Hiker's Guide to the Galaxy*, Ch. 6

9.1 Report of Findings

Examination of stuttering patterns that appear relatively straightforward to characterize linguistically at the outset, on completion of the thesis, seem to be a complex issue involving both language and developmental differences.

The first part of the thesis (chapters two to four) was concerned with the investigation of the linguistic and phonetic factors that predict stuttering in German.

The first empirical chapter (two) looked at four basic characteristics that were reported by Brown in (1945) for English, to see the extent to which they apply to German. Brown's four "factors" are characteristics of words that predict the likely loci of disfluency in English-speaking adults who stutter. Thus a word is more likely to be stuttered for these speakers if it is 1) a content word (the factor Brown referred to as grammatical class but which has since been interpreted as the function/content word distinction) 2) starts with a consonant, 3) is positioned at the beginning of a sentence, and 4) if it is a long word. These same factors were examined in native German-speaking children and adults who stutter. For the adult group, it was predicted that words starting with consonants would not lead to as much of an increase in disfluencies compared with English samples, because of cross-linguistic differences in syllable onset properties. It was predicted that stuttering would be more likely in later sentence positions in German because the verb is usually near the end of a sentence. There were no obvious reasons to expect differences on the two remaining factors, content words and word length. With children, it was hypothesised that Brown's factors that specify level of linguistic difficulty would not be such a good predictor of stuttering rate. Specifically, it was predicted that the difference in stuttering rate between function and content words would be lower in children. For the adults both word type (content/function) and word length increased stuttering rate significantly, whereas changes in stuttering rate for the other two factors were non-significant. It was also found that when word difficulty (based on a combined measure of all factors) increased, stuttering rate rose. With children, only the word-length factor was significant, and stuttering rate was not governed to the same extent by overall word difficulty.

Phonetic complexity was based on a simple measure in Brown' work (whether a word started with a consonant or vowel). As this measure is crude, phonetic complexity as a determinant of stuttering in German was investigated in detail in chapter three, using a potentially better measure. The aim of this chapter was to investigate how this measure of phonetic complexity affects stuttering rate in German and to assess changes developmentally over different age groups (i.e. over a wider age group than was used in chapter two). The measure of phonetic difficulty was that used by Jakielski's (1998), and is called the index of phonetic complexity (IPC). The IPC gives a composite score summed over eight individual characteristics. For German, as previously shown for Spanish and English, stuttering rates for function words were not predicted by phonetic difficulty. This supports the interpretation that disfluencies on these words are qualitatively different to those of content words and are not influenced by their linguistic properties. Significant linear trends (when stuttering rate was divided into IPC sum categories) were found for content words in both children over the age of six and adults. A comparative linguistic analysis was carried out to establish whether there were overall differences in IPC scores and which of the factors would have most impact in each language. This was analysed both in respect to average IPC score and frequency of individual characteristics within each language. It was shown that German content words have a higher mean IPC sum compared to their English counterparts. This also affected stuttering rates since the mean IPC score increase was greater between fluent and stuttered words in German. As previously found for Spanish but not for English, factor 5 (word shape) influenced stuttering rates in both German age groups. This finding is particularly interesting since it refers to a word final phonetic characteristic (i.e. a place where stuttering does not occur). However, from a planning and execution perspective it could affect the timing of the plan.

It was shown (in chapters two and also three) that age-related changes occurred in function and content word stuttering in German (as was previously shown for English). Also, function words that precede a content word are significantly more likely to be stuttered than those that follow content words in English (Au-Yeung, Howell, & Pilgrim, 1998; Howell et al., 1999). These studies have used the concept of the phonological word as a means of investigating these phenomena. Phonological words help to determine the position of function words relative to content words, and to establish the origin of the patterns of disfluency with respect to these two word classes. Chapter four analysed German speech for similar patterns. German contains many long compound nouns. Accordingly, German content words are more complex than English ones – as was shown in chapter three. Even when subdivided into individual components, the patterns of disfluency within phonological words may differ between German and English. Results indicated three main findings. Firstly, function words that occupy an early position in a PW have higher rates of disfluency than those which occur later in a PW, this being most apparent for the youngest speakers. Second, function words that precede the content word in a PW have higher rates of disfluency than those that follow the content word. Third, young speakers exhibit high rates of disfluency on function words, but this drops off with age and, correspondingly, disfluency rate on content words increases.

The second part of the thesis (chapters 6-8) is analysing fluent and bilingual children's lexical and semantic development. The reason for this being: 1) the reviewed evidence of a link between bilingualism and the onset of stuttering; 2) to get a clearer understanding whether the sudden growth in the lexicon combined with the development of the grammatical system would lead to errors in naming and / or lexical retrieval; 3) To device test material that can shed light on the processes that operate at the time when stuttering first occurs. Chapter six examined how receptive syntax developed in bilingual as compared to monolingual children in school-age children. The aim of the investigation was the creation of a test instrument that could assess children over a wide age range. Furthermore it was analysed whether there were delays in the acquisition of grammar for the bilingual children as was suggested by previous research (e.g. Hulk, 2000). A group of bilingual children (from ages 5 to 10 years) was tested in both English and German and was then compared to monolingual groups of English- and German-speaking children. Questionnaire data showed that most of the bilingual children were English dominant. Irrespective of this, no overall group differences were found. However, analysis of individual categories revealed differences with regards to compound word and word order performance. Bilingual children performed better than their monolingual English counterparts in the compound noun category, but performed worse than their monolingual German peers in the word order category (both of which make sense in terms of the differences between English and German). The broad similarity between language groups, as a measure of external validity of the German version of the test, was seen as suggesting that this test was a useful tool to investigate syntactical progress in language development.

Chapter seven reports preliminary efforts in starting to track early fluency development in five children being brought up as English-German bilinguals (from age 23 months onwards). The tests conducted so far include a measure of productive vocabulary, a picture naming task, syntactic development in both languages, and MLU recordings. It was shown that the vocabulary development was at the lower end of the distribution for these children. The picture naming task did not seem to indicate a relationship between the amount of vocabulary gained - and the number of errors made in a session, possibly supporting the independence of lexical and syntactic development. Picture-naming errors occurred at the same rate in both languages. Future work with more subjects is needed to examine absolute error rates to see how they compare with Gershkoff-Stowe and Smith's (1997) work. It should also examine whether error rate peaks later in bilingual children. Both of the above could indicate some delay at early ages in fluency measures. The receptive syntax task showed that, even though most children started of passing none, a small numbers of blocks were passed at the very end of testing. However in productive vocabulary the children were within the norms for English counterparts. The subjective feeling was that those children started slightly later but then caught up more rapidly, although further work is needed (more subjects and longer range) to assess this. The MLU data showed somewhat surprising (and contrasting patterns) for two individual children. One child had a relatively small lexicon at the start of testing but had high MLU, another showed the direct opposite.

Lexical retrieval (using the TOT paradigm) in German speaking children in two age groups was investigated in chapter eight. One of the aims was to establish a test instrument that would make it possible to investigate word access in child populations where differences would be hypothesised. Another goal was the investigation of whether or not accuracy of lexical knowledge would increase with age and, if so, what type of information or which word factors affect this process. Results indicated that accuracy did increase with age (as previously found by Wellman, 1977, but not Butterfield et al. 1988). The number of positive TOTs and don't knows increased with age. It was also shown that the older children could report the gender of a word (significantly above chance level) when they were in a TOT state. It was also shown that word length was associated with more positive TOT states, whereas the phone a word starts with did not seem to have an effect.

9.2 Theoretical and Wider Implications

The implication from the first chapter is that subtle cross linguistic differences in linguistic structure could have an effect on the pattern of disfluencies observed. Furthermore there was a large grammatical class difference for the adults, that can also be linked to comparative differences in the two languages. For the children this difference was not found to be significant. Phonetic difficulty of grammatical class was more closely investigated in the second chapter which highlighted a more complex structure of German content words. The patterns within phonological words may be general to German and English, as was shown in chapter four. All these results can be accounted for by the EXPLAN model, assuming lexical class operates equivalently across these languages, or that lexical categories contain some common characteristic that is associated with fluency across the languages.

The findings on German speakers who stutter indicate a lot of "continuity" between the young children who stutter and what has been reported with young speakers who are fluent. Jakielski's work on fluency development (which was extended here in chapter three) can be linked to the growing body of evidence on the relationship between distributional properties of languages and their impact on fluency development / language acquisiton (Demuth, 2001; Tomasello & Abbot-Smith, 2000; Tomasello & Abbot-Smith, 2002). It also builds on and extends Clark and Clark's (1977) work on function word repetition. These facts are reasonably well documented. There are also other influences reported in fluent speakers that have been hypothesized to relate to the breakdown of fluency development that could then lead to stuttering. These include previous research on fluent speakers in "snap-shot" type studies (one or two measures per study - CDI and picture naming Gershkoff-Stowe & Smith, 1997) and, also, areas where stuttering or developmental work indicate new experimental tools that are needed to investigate other possible determinants of stuttering (lexical retrieval/TOT). The goal of the line of research in the second part of this thesis, was to develop a battery of tests that can be made with very young children whose fluency development is followed up over a period of time. When selecting young children, apart from possibly genetic factors (as suggested by Bloodstein, 1993; but critically evaluated by Wingate, 2002), there are no established predictors of who is likely to stutter. Hence it was decided to use children (likely to be fluent) being brought up as English-German bilinguals whose fluency development may be delayed (like CWS). These children were then assessed as to how their lexicon and syntax develops. The full plan of future work is (in rough chronological order) CDI, MLU, picture naming, onset of syntax and lexical retrieval.

The EXPLAN model makes specific reference to developmental patterns and specifically highlights points of difficulties where fluency is more likely to break down, which is not the case in CRH. Chapters two and four highlight the points of difficulty for German speakers and chapter four looked at the developmental pattern and the relationship between function and content words within an articulatory context

Another appealing factor of EXPLAN is the implicit assumption of the continuity between fluent speakers and persons who stutter. Empirical work has established factors in fluency development that should extend to stuttering (and future work is needed to establish the factors that are implicated in this process such as TOT and syntax).

In respect to theories of bilingualism, the thesis highlighted certain aspects of lexicon (compound nouns) and acquisition of syntax (i.e. sentence position). In this respect both chapters six and seven highlight the fact that generalisations about bilingualism cannot be made, but rather specific processes need to be investigated separately (as suggested by Bialystok, 2001).

This thesis has provided more evidence on the continuity between early stuttering and normal fluency development (here in bilinguals). It is work that has been conducted from the perspective of the EXPLAN theory. This is the only current theory that offers any indication of what happens in late development that possibly relates to the persistence of stuttering. It emphasizes that what children who stutter do, changes over development. There is broad support for the theory throughout and the studies in chapter four offer some specific evidence (for German) in support of it.

Bilingual speakers have two sets of "plans" that they feed through the same execution system. Work on older bilingual speakers has been commenced (Howell et al., 2003), but here work was done to provide broad brush stroke indicators of where bilingual children may have problems in fluency development. Some of the variables relate to planning (ROST, TOT), and others to planning + execution (MLU, picture naming).

For the first time in the literature a detailed analysis on the linguistic factors associated with incidences of stuttering in German is provided in this thesis. The first four chapters provide evidence that for German adult speakers who stutter the complexity of content words (which as compared to English could be classed as *super-content* words) almost exclusively determines the words that are stuttered. They are phonetically more complex and on average longer, for instance. In other words wherever the execution of a word takes longer in planning time, in EXPLAN terms, there is also more disfluency. The first part of the thesis also provides first evidence of a developmental exchange pattern of stuttering in German. Even though function word disfluencies have been reported previously in German children (Rommel, Häge, Johannsen, & Schulze, 1997), this developmental trend was not made explicit. This indicates that this exchange pattern is not just a one language occurrence but happens throughout the development in people who stutter. For theories of stuttering this causes problems as highlighted previously throughout the thesis. Many theories are not built around a framework that could account for children's stuttering on the less complex function words. For instance, if as assumed in Postma and Kolk's (Kolk & Postma, 1997; Postma & Kolk, 1993) covert repair hypothesis, speakers who stutter have a slower phonological system why would that speed affect different word types in different age groups. Furthermore why would that lead to different patterns across languages? There is also Wingate's (1988; 2002) ongoing assertion that linguistic stress is the underlying problem for speakers who stutter. The analyses of the first part of the thesis seem to disprove this point. Children stutter frequently on the function words which are in general not receiving linguistic stress. As such this cannot be the underlying problem for speakers who stutter. Moreover, genetic models of stuttering (see for instance Kidd, 1980; 1984) have difficulty accounting for developmental differences in stuttering patterns for different word classes. Going back to the theories and positions on disfluencies introduced in chapter one the diagram that was introduced is reproduced here.

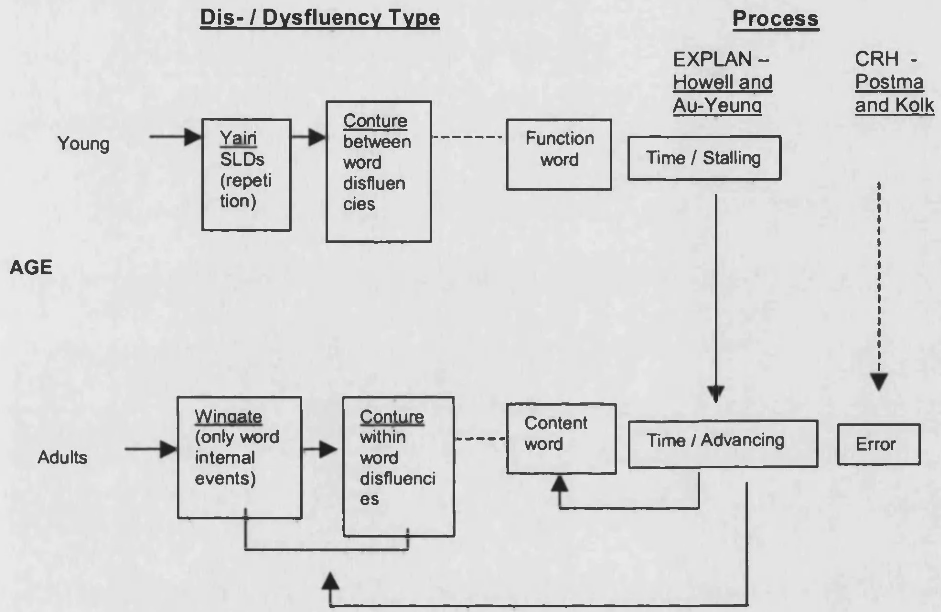


Figure 38: Diagrammatic representation (see chapter one) summarising the different positions taken by researchers with regard to age group, word type, and the processes leading to disfluencies.

The results of the thesis are now considered in relation to the different positions in this diagram. The German results reported here provide further evidence that theories have to take a stance with regard to the different patterns of children and adults who stutter. This highlights the one sided focus of both the positions of Yairi and co-workers (see for instance Yairi, 1982; 1983; Yairi & Ambrose, 1992a; 1992b) and Wingate (1988; 2002). The former almost exclusively focus on childhood patterns of stuttering whereas the opposite is true for the latter – indicated in the diagram by the arrows leading from adults horizontally across to these positions. If a position cannot account for the whole range of patterns it is neither a feasible description of the disorder nor could provide an underlying theoretical framework.

Conture (1990) uses the descriptive terms of between- and within word disfluencies. These two terms broadly correspond to function and content word disfluencies which could then be linked to the age related differences. However, Conture (1990) does not explicitly link these terms to age and has not proposed any specific processes that lead to either type of disfluency – in the figure the lack of an arrow across shows that this relationship is not accounted for.

Kolk and Postma's (1997; Postma & Kolk, 1993) position (see previous paragraph) has problems accounting for a number of patterns reported in this thesis. As mentioned above CRH has difficulties accounting for word class differences. Furthermore, the often different patterns across languages (as reported here for content words in chapter three) and for bilingual speakers who stutter (findings that were summarised in the introduction) cannot be explained by this model. Also the serial position effect reported in chapter four for German phonological words cannot be accounted for by CRH.

This then leaves us to consider the EXPLAN model of fluency failure (Howell, 2002; Howell & Au-Yeung, 2001; 2002) in relation to the results reported here. The vertical arrow in the diagram down from function to content word disfluencies indicates that EXPLAN does take the different patterns of child and adult speakers into consideration. Children are similar in their disfluency pattern to speakers who produce normal non-fluencies. They are assumed to use a process by which they stall the planning with the repetitions of function words until the plan for the next word is available. Adult speakers on the other hand advance to early and use a process of prolonging, repeating or blocking on the first part of the word to gain time for the rest of the word's planning to be completed. The results from this thesis are consistent with these processes. Particularly the positional results from the PW (chapter four) provide evidence that this processes also seems to take place in German speakers who stutter. Adult speakers who stutter rarely, if ever, would stutter on a function word that follows a content word within the same PW. When stuttering does occur on function words, which is less frequently the case for adults (in German even less so than in English), then it would almost exclusively affect the first function word within a PW positioned before the content word. The arrows referring back to word internal events described as stuttering by Wingate (1988, 2002) and Conture's within word disfluencies (1990) correspond directly to the content word disfluencies in EXPLAN. In the German results that showed an even higher frequency of content word stuttering in adults compared to the English speakers, adults seem to try to attempt the first part of the word (classified as advancing type disfluencies in EXPLAN) which they would then need to prolong or repeat to gain time for the rest of the plan to be finalised. For German the longer planning is necessary due to the phonetically more complex nature of this word class (see chapter three).

However, EXPLAN does not take a stance on what processes lead to a change in the pattern at around age 8-11 years which seems to be the time in both English and German speakers who stutter when this takes place (see chapter four). This is an area that needs to ideally be investigated longitudinally to analyse when and why this change takes place. This indicates that research in this area should not only focus on the time of stuttering onset but also at the time where the pattern of disfluencies changes.

One of the pre-eminent researchers investigating stuttering, Nan Bernstein-Ratner (1997), has linked the onset of stuttering directly to sudden increases in lexical- combined with the start of grammatical development. In the second part of the thesis which is mainly exploratory in nature, the aim was to investigate these two aspects in relation to fluency development in bilingual children at the age equivalent to stuttering onset. These were children that do not have a speech problem. The reasons for the use of bilingual speaking children were given in chapter five. A new syntax task was developed with the aim to eventually use it to look at both the age when stuttering first occurs in childhood and could also be used to analyse what takes place at the age where the pattern of stuttering changes. When comparing bilingual German/English speaking children with their monolingual counterparts this test showed differences in syntax processing that can be linked to the first part of the thesis. The bilingual children in chapter six showed superior compound word performance and did less well in a word order category. Interestingly these correspond to the factors that were hypothesised to influence stuttering in the first part of the thesis.

Not necessarily a delay, but lexical development at the lower end of the spectrum, was observed for the bilingual children aged between two and three years – chapter seven. This is consistent with previous reports of delays in this population (as found in numerous studies, such as Ben-Zeev, 1977; Bialystok, 1988; Merriman & Kutlesic, 1993; Rosenblum & Pinker, 1983; Umbel et al., 1992). Since all the children had total vocabulary sizes below 250 words they could still be assumed to be in the phase of the main vocabulary burst. Unfortunately there was no relationship found between big increases in vocabulary and picture naming problem. Even though naming errors did occur relatively frequently. The syntax task performance was also at the lower end of spectrum – compared to norms from a wide scale English sample. Since it is a very small sample it is difficult to draw more general conclusions from this finding. It is however, suggesting that the simultaneous acquisition of two languages is a more demanding task for the child which has consequences on the rate of development. The fact that there was evidence of cross linguistic influences / transfer even in school aged children reported in chapter six corroborates with this conclusion.

Parallels can also be drawn between the results of chapter eight and the first part of the thesis. Children had more difficulties accessing lexical items that were longer than shorter items (even when word frequency was controlled). This again highlights the time for the planning of longer words. This chapter also provided evidence for the first time in the literature that the older age group of German speaking children could correctly report word gender, beginnings, ends and syllable number more often than for their negative TOT states. That words starting with consonants and vowels showed no difference in lexical retrieval corresponds to the findings in chapter two of a null effect for stuttering on words beginning with consonants and vowels. It is possible that because vowel onsets in German are stronger they also require more planning which might be a reason for lexical retrieval to be equal across these categories.

Even though the second part of the thesis did not provide any new insights into fluency development per se, it did show instances of cross linguistic influences and thus provided converging evidence for compound noun and word order processing differences between the two languages. Furthermore the test material opens up new avenues to explore the language development of German speaking infants and school children.

9.3 Problems encountered and future directions for this research area

9.3.1 Problems encountered

In the first part of the thesis the main problem encountered was the fact that in Germany treatment for children who stutter start generally later compared to children seen by speech therapists in the UK. This made it extremely difficult to obtain speech samples of this age group. Eventually research links with one group in Germany were established who were also interested in these younger speakers who stutter. This provided the main breakthrough in this part and only then was it possible to fully assess the developmental pattern for German.

There were a number of difficulties with studies in the second part of the thesis. The main problems were encountered in the work with bilingual infants. Even though they did start off naming pictures in both languages language switching became less common during the time of later visits which made MLU for the less dominant language impossible. Having only read the research indicating that language switching does occur even in young children (see the review of this issue by Köppe & Meisel, 1995) it was surprising to the author that in practice this did not seem to occur all that frequently. However, one of the positive aspects of the ROST task is that the child hear only one language at a time which makes it possible to assess syntax comprehension in both individual languages.

The design of new tests for a wider age range of bilingual children brings with it a number of problems, too. It has to include enough categories for the older children not to reach ceiling level and should also not be too challenging for the youngest age group to become de-motivated. Inevitably this leads to only a very broad overview of development because specific syntactic issues are not explored in greater detail to give the test a wider range. This worked reasonably well with school aged children where both the growth of syntactic competence and also specific language related difference were observed. However children in the youngest age group would have benefited from fewer categories which could have been given more depth. This would have then made it possible to compared specific language differences in more detail. The number of categories included in their age group might have had the effect that it was demanding longer attention and might have lead to them not being credited categories even when they had acquired it.

9.3.2 Future directions

Further work is needed on cross linguistic patterns of stuttering that would test the findings in more detail. Ideally suited for this would be test cases of bilingual German/English speakers who stutter. It would be interesting to see whether individuals would show the patterns that were analysed here. Would these speakers show a pattern where they would stutter almost exclusively in German on content words whereas in comparison also stutter in higher frequency on English function words, too. Furthermore would they show a greater difference words starting with consonants in English but about equal amounts of words starting with consonants or vowels in German. Also, if they have one more dominant language what would the pattern be of stuttering in the two languages. More case studies are still needed to clarify this within the same individual.

The review of literature has also emphasised throughout that there is still a lot more to find out about the processes that take place at the time of stuttering onset and what exactly leads to the changes in the pattern of stuttering in later development. Ideally suited to this question is longitudinal research. The latter part of the thesis was intended to construct material that would be suitable to test these stages. It would be ideal to test a large group of children that would have a predisposition to the disorder, be it due to bilingualism or due to familial history of stuttering or both. Individual profiles can then be analysed to see whether there is anything separating children who stay fluent and those that develop the disorder. Furthermore specific difficulties at the time of onset can be explored and analysed if these can be generalised across children.

On a more general note the EXPLAN model of fluency failure needs to be more specific about the processes that lead to a change in the pattern of stalling and advancing disfluencies. If children display a pattern that is more like the normal non-fluencies of fluent speakers then why and at what point do they change to the pattern of advancing disfluencies. EXPLAN has in this respect been purely descriptive rather than provide a theoretical reason for this occurrence. Since EXPLAN's focus is on the synchrony of the two components the shift of disfluency patterns should be explored with a more specific description of the timing mechanisms of execution and planning. Future work is needed to build on these findings by extending the subject base and also, as a firmer test of EXPLAN, to design tests that look systematically at planning / execution variables and their interaction in detail.

Another interesting future direction in this research area would be the possibility of an integration of the BIMOLA model within the EXPLAN theory which could provide a way forward in looking at bilingualism, the development of fluency and its breakdown in stuttering. A combination of both these models could benefit each since BIMOLA is not specifically designed for the timing aspects of speech planning and execution whereas EXPLAN does not specifically account for cases of bi- and multilingualism. An integration of the features eventually might shed light on the different patterns of stuttering observed in some bilinguals and also give a better idea of how and why there is a link between second language onset and first occurrence of speech disfluencies.

10 References

Abudarham, S. (1987). Fact and Fiction. In S. Abudarham (Ed.), Bilingualism and the bilingual: An interdisciplinary approach to pedagogical and remedial issues. (pp. 15-34). Worcester, UK: Billing & Sons Ltd.

Adams, M. R. (1982). Fluency, nonfluency, and stuttering in children. Journal of Fluency Disorders, 7, 171-185.

Adams, M. R. (1990). The demands and capacities model: I. Theoretical elaborations. Journal of Fluency Disorders, Vol 15(3), 135-141.

Andrews, G., & Harris, M. (1964). The syndrome of stuttering. London: Pamphlet for the Spastic Society Medical Education and Information Unit in association with W. Heinemann Medical Books.

Andrews, G., Hoddinott, S., Craig, A., Howie, P., Feyer, A.-M., & Nielson, M. (1983). Stuttering: A review of research findings and theories circa 1982. Journal of Speech and Hearing Disorders, 48, 226-246.

Au-Yeung, J., Howell, P., Davis, S., Charles, N., & Sackin, S. (2000). UCL survey of bilingualism and stuttering. In H.-G. Bosshardt, J. S. Yaruss, & H. F. M. Peters (Eds.), Fluency disorders: Theory, research, treatment and self-help. Nijmegen: Nijmegen University Press.

Au-Yeung, J., Howell, P., Davis, S., Sackin, S., & Cunniffe, P. (2000). Introducing the Preschoolers Reception of Syntax Test (PROST). Proceedings of conference on Cognitive Development: Besancon France.

Au-Yeung, J., Howell, P., & Pilgrim, L. (1998). Phonological words and stuttering on function words. Journal of Speech, Language, and Hearing Research., Vol 41(5), 1019-1030.

Bar, A., Singer, J., & Feldman, R. G. (1969). Subvocal muscle activity during stuttering and fluent speech: A comparison. Journal of the South African Logopedic Society, 16, 9-14.

Barrett, M. (1995). Early lexical development. In P. Fletcher & B. MacWhinney (Eds.), The handbook of child language. (pp. 362-392). Oxford: Blackwell.

Bastiaansen, M., Van Der Linden, M., ter Keurs, M., Dijkstra, T., & Hagoort, P. (2002). Open-class words produce a larger theta power increase than closed-class words. Cognitive Neuroscience Ninth Annual Meeting, San Francisco, 14-16 April.

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). Oxford: Blackwell.

Bates, E., & Goodman, J. (2001). On the inseparability of grammar and the lexicon: Evidence from acquisition. In M. Tomasello & E. Bates (Eds.), Language development: The essential readings. (pp. 134-162). Oxford UK: Blackwell Publishers Ltd.

Bates, E., & MacWhinney, B. (1989). Functionalism and the competition model. In B. MacWhinney & E. Bates (Eds.), The crosslinguistic study of sentence processing. (pp. 3-73). Cambridge: Cambridge University Press.

Ben-Zeev, S. (1977). The influence of bilingualism on cognitive strategy and cognitive development. Child Development, 48, 1009-1018.

Berman, R. A., & Clark, E. (1989). Learning to use compounds for contrast: Data from Hebrew. First Language, 9, 247-270.

Bernstein Ratner, N. (1997). Stuttering: A psycholinguistic perspective. In R. F. Curlee & G. M. Siegel (Eds.), Nature and Treatment of Stuttering: New Directions. Needham Heights, MD: Allyn & Bacon.

Bernstein Ratner, N., & Benitez, M. (1985). Linguistic analysis of a bilingual stutterer. Journal of Fluency Disorders, 10, 211-219.

Bernstein Ratner, N., & Sih, C. C. (1987). Effects of gradual increases in sentence length and complexity on children's dysfluency. Journal of Speech and Hearing Disorders, 52, 278-287.

Bialystok, E. (1988). Levels of bilingualism and levels of linguistic awareness. Developmental Psychology, 24, 560-567.

Bialystok, E. (2001). Bilingualism in development. New York: Cambridge University Press.

Biesalski, P. (1978). Ärztlicher Rat bei Sprachstöhrungen im Kindesalter. Stuttgart: Georg Thieme Verlag.

Bishop, D. V. M. (1983). The Test of Receptive Grammar. Published by the author and available from Age and Cognitive Performance Research Centre: University of Manchester, M13 9PL.

Bloodstein, O. (1987). A Handbook on Stuttering. (4th edition ed.). Chicago: The National Easter Seal Society.

Bloodstein, O. (1993). Stuttering: The search for a cause and a cure. Boston: Allyn & Bacon.

Bloodstein, O. (1995). A Handbook on Stuttering. (5th ed.). Chicago: The National Easter Seal Society.

Bloodstein, O. (2002). Early stuttering as a type of language difficulty. Journal of Fluency Disorders, Vol 27(2), 163-167.

Bloodstein, O., & Gandtwerk, B. F. (1967). Grammatical function in relation to stuttering in young children. Journal of Speech and Hearing Research., 10(4), 786-789.

Bloodstein, O., & Grossman, M. (1981). Early stutterings: Some aspects of their form and distribution. Journal of Speech and Hearing Research., Vol 24(2), 298-302.

Bock, K., & Levelt, W. J. M. (1994). Language production: Grammatical encoding. In M. A. Gernsbacher (Ed.), Handbook of Psycholinguistics. (pp. 945-984).

Booij, G. (1983). Principles and parameters in prosodic phonology. Linguistics, 21, 249-280.

Booij, G. (1985). Coordination reduction in complex words, a case for prosodic phonology. In van der Hulst H. & N. Smith (Eds.), *Advances in Non-linear Phonology.* Dordrecht: Foris.

Booij, G. (1996). Cliticization as prosodic integration: The case of Dutch. The Linguistic Review, 13, 219-242.

Browman, C. P., & Goldstein, L. (1986). Towards an articulatory phonology. Phonology Yearbook, 3, 219-252.

Browman, C. P., & Goldstein, L. (1989). Articulatory gestures as phonological units. Phonology, 6, 151-206.

Browman, C. P., & Goldstein, L. (1992). Articulatory phonology: An overview. Phonetica, 49, 155-180.

Browman, C. P., & Goldstein, L. (1997). The gestural phonology model. In W. Houlstijn, H. F. M. Peters, & P. H. H. M. van Lieshout (Eds.), Speech Production: Motor Control, Brain Research and Fluency Disorders. (pp. 120-135). Amsterdam Elsevier:

Brown, A. S. (1991). A review of the tip-of-the-tongue experience. Psychological.Bulletin., Vol 109(2), 204-223.

Brown, C. M., Hagoort, P., & ter Keurs, M. (1999). Electrophysiological signatures of visual lexical processing: Open- and closed-class words. Journal of Cognitive Neuroscience, Vol 11(3), 261-281.

Brown, R. (1973). A first language: the early stages. Cambridge MA: Harvard University Press.

Brown, R., & McNeill, D. (1966). The "tip of the tongue" phenomenon. Journal of Verbal Learning and Verbal Behavior., 5(4), 325-337.

Brown, S. F. (1937). The influence of grammatical function on the incidence of stuttering. Journal of Speech Disorders, 2, 207-215.

Brown, S. F. (1938a). A further study of stuttering in relation to various speech sounds. Quarterly Journal of Speech, 24, 390-397.

Brown, S. F. (1938b). Stuttering with relation to word accent and word position. Journal of Abnormal and Social Psychology., 33, 112-120.

Brown, S. F. (1938c). The theoretical importance of certain factors influencing the incidence of stuttering. Journal of Speech Disorders, 3, 223-230.

Brown, S. F. (1943). An analysis of certain data concerning loci of 'stutterings' from the viewpoint of general semantics. American Congress on General Semantics Papers., 2, 194-199.

Brown, S. F. (1945). The loci of stutterings in the speech sequence. Journal of Speech Disorders, 10, 181-192.

Brown, S. F., & Moren, A. (1942). The frequency of stuttering in relation to word length during oral reading. Journal of Speech Disorders, 7, 153-159.

Butterfield, E. C., Nelson, T. O., & Peck, V. (1988). Developing aspects of the feeling of knowing. Developmental Psychology, 24, 654-663.

Butterworth, B. (1989). Lexical access in speech production. In W. Marslen-Wilson (Ed.), Lexical Access Representation and Process. (pp. 108-135). Cambridge, MA.: MIT Press.

Cabrera, V., & Bernstein Ratner, N. (2000). Stuttering patterns in the two languages of a bilingual child. Paper presented at the ASHA annual convention, Washington DC: November 16-19.

Caramazza, A. (1997). How many levels of processing are there in lexical access? Cognitive Neuropsychology, 14(1), 177-208.

Caramazza, A., & Miozzo, M. (1998). More is not always better: A response to Roelofs, Meyer, and Levelt. Cognition, 69, 231-241.

Caramazza, A., & Miozzo, M. (1997). The relation between syntactic and phonological knowledge in lexical access: Evidence from the "tip-of-the-tongue" phenomenon. Cognition, Vol 64(3), 309-343.

Charman, T., Drew, A., Baird, C., & Baird, G. (2003). Measuring early language development in preschool children with autism spectrum disorder using the MacArthur Communicative Development Inventory (Infant Form). Journal of Child Language, Vol 30(1), 213-236.

Chomsky, N. (1965). Aspects of the theory of syntax. Cambridge MA: MIT Press.

Chomsky, N. (1995). The minimalist program. Dordrecht: Foris.

Chomsky, N., & Halle, M. (1968). The sound pattern of English. New York: Harper & Row.

Clahsen, H. (1995). German plurals in adult second language development: Evidence for a dual mechanism model of inflection. In L. Eubank, L. Selinker, & M. Sharwood Smith (Eds.), The current state of interlanguage. (pp. 75-94). Amsterdam: John Benjamin.

Clark, E. V. (1981). Lexical innovation: How cildren learn to create new words. In W. Deutsch (Ed.), The child's construction of language. (pp. 299-328). New York: Academic Press.

Clark, E. V. (1985). The acquisition of Roamnce, with special reference to French. In D. I. Slobin (Ed.), The crosslinguistic study of language acquisition:Vol. 1. The data. (pp. 687-782). Hillsdale, NJ: Lawrence Erlbaum Ass.

Clark, E. V. (1993). The lexicon in acquisition. Cambridge: Cambridge University Press.

Clark, E. V. (1995). Later lexical development and word formation. In P. Fletcher & B. MacWhinney (Eds.), The handbook of child language. (pp. 393-412). Oxford: Blackwell.

Clark, E. V. (1998). Lexical creativity in French speaking children. Cahiers de Psychologie Cognitive, 17, 513-530.

Clark, E. V., & Berman, R. A. (1987). Types of linguistic knowledge: Interpreting and producing compound nouns. Journal of Child Language, 14, 547-567.

Clark, E. V., Gelman, S. A., & Lane, N. M. (1985). Compound nouns and category structure in young children. Child Development, 56, 84-94.

Clark, H., & Clark, E. (1977). Psychology and Language: An Introduction to Psycholinguistics. New York: Harcourt Brace.

Conture, E. G. (1990). Stuttering. (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Conway, J. K., & Quarrington, B. J. (1963). Positional effects in the stuttering of contextually organized verbal material. Journal of Abnormal.Psychology, 67(3), 299-303.

Cordes, A. K., & Ingham, R. J. (1999). Effects of time-interval judgment training on real-time measurement of stuttering. Journal of Speech, Language, and Hearing Research, Vol 42(4), 862-879.

Dale, P. S., Bates, E., Reznick, J. S., & Morisset, C. (1989). The validity of a parent report instrument of child language at 20 months. Journal of Child Language, 16, 239-249.

Dayalu, V. N., Kalinowski, J., & Stuart, A. (2003). A flawed memory with one's own work: A reply to Wingate. Journal of Speech, Language and Hearing Research, in press

de Houwer, A. (1995). Bilingual language acquisition. In P. Fletcher & B. MacWhinney (Eds.), The Handbook of Child Language. (pp. 219-250). Cambridge MA: Blackwell Ltd.

de Villiers, J. G., & de Villiers, P. A. (1973). A cross-sectional study of the acquisition of grammatical morphemes in child speech. Journal of Psycholinguistic Research, 2, 267-278.

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. Psychological Review, 93, 283-321.

Dell, G. S. (1988). The retrieval of phonological forms in production: Tests of predictions from a connectionist model. Journal of Memory and Language, 27, 124-142.

Dell, G. S., & O'Seaghdha, P. G. (1991). Mediated and convergent lexical priming in language production: A comment to Levelt et al. (1991). Psychological Review, 98, 604-614.

Demuth, K. (2001). Prosodic constraints on morphological development. In J. Weissenborn & B. Höhle (Eds.), Approaches to Bootstrapping: Phonological, Syntactic and Neurophysiological Aspects of Early Language Acquisition. (pp. 3-21).

Deuchar, M. (1999). Are function words non-language-specific in early bilingual two-word utterances? Bilingualism: Language and Cognition, 2(1), 23-34.

Deuchar, M., & Quay, S. (2000). Bilingual acquisition: Theoretical implications of a case study. Oxford: Oxford University Press.

Dickson, S. (1971). Incipient stuttering and spontaneous remission of stuttered speech. Journal of Communication Disorders, 4, 99-110.

Dixon, R. M. W. (1977a). A Grammar of Yidn. Cambridge UK: Cambridge University Press.

Dixon, R. M. W. (1977b). Some phonological rules of Yidn. Linguistic Inquiry, 8, 1-34.

Döpke, S. (1998). Competing language structures: the acquisition of verb placement by bilingual German-English children. Journal of Child Language, 25, 555-584.

Döpke, S. (1999). Cross-linguistic influences on the plaecement of negation and modal particles in simultaneous bilingualism. Language Sciences, 21, 143-175.

Döpke, S. (2000). The interplay between language-specific development and crosslinguistic influence. In S. Doepke (Ed.), Cross-linguistic structures in simultaneous bilingualism. Studies in bilingualism, vol. 21. (pp. 79-103). Amsterdam, Netherlands: John Benjamins Publishing Company.

Dworzynski, K., & Howell, P. (2004). Predicting stuttering from phonological complexity in German. Journal of Fluency Disorders, 29(2), 151-176.

Dworzynski, K., Howell, P., Au-Yeung, J., & Rommel, D. (2003). Stuttering on function and content words across age groups of German speakers who stutter. Journal of Multilingual Communication Disorders, in press

Dworzynski, K., Howell, P., Au-Yeung, J., & Rommel, D. (2004). Stuttering on function and content words across age groups of German speakers who stutter. Journal of Multilingual Communication Disorders, 2(2), 81-101.

Eisenson, J. (1984). Stuttering as an expression of insufficient language development. In L. J. Raphael, C. B. Raphael, & M. R. Vasovinos (Eds.), Language and Cognition: Essays in Honor of Arthur Bronstein. (pp. 77-95). New York: Plenum.

Eisenson, J. (1986). Language and speech disorders in children. New York: Pergamon Press.

Elbers, L. (1985). A tip-of-the-tongue experience at age two? Journal of Child Language, 12, 353-365.

Elbers, L., & Wijnen, F. (1992). Effort, production skill, and language learning. In C. A. Ferguson, L. Menn, & C. Stoel-Gammon (Eds.), Phonological Development: Models, Research, Implications. (pp. 337-368). Baltimore: York Press.

Faust, M., Dimitrovsky, L., & Davidi, S. (1997). Naming difficulties in language-disabled children: Preliminary findings with application of the tip-of-the-tongue paradigm. Journal of Speech, Language, and Hearing Research, 40, 1026-1036.

Fenson, L., Bates, E., Dale, P., Goodman, J., Reznick, J. S., & Thal, D. (2000). Measuring variability in early child language: Don't shoot the messenger. Child Development, Vol 71(2), 323-328.

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., & Pethick, S. J. (1994). Variability in early communicative development. Monographs.of the.Society.for.Research.in Child Development, Vol 59(5), v-173

Fowler, C. A. (1986). An event perception approach to the study of speech perception from a direct-realist perspective. Journal of Phonetics, 14, 3-28.

Fowler, C. A. (1995). Speech production. In J. L. Milner & P. D. Eimas (Eds.), Speech, Language and Communication. (pp. 1-27). San Diego: Academic Press.

Freeman, F. J., & Ushijima, T. (1975). Laryngeal activity accompanying the moment of stuttering: A preliminary report of EMG investigations. Journal of Fluency Disorders, 1, 36-55.

Fromkin, V. A. (1971). The non-anomalous nature of anomalous utterances. Language, 47, 27-52.

Garrett, M. F. (1975). The analysis of sentence production. In G. H. Bower (Ed.), The Psychology of Learning and Motivation: Vol. 9. London: Academic Press.

Garrett, M. F. (1980). Levels of processing in sentence production. In B. Butterworth (Ed.), Language production, vol.1, speech and talk. New York: Academic.

Garrett, M. F. (1982). Production of speech: Observations from normal and pathological language use. In A. W. Ellis (Ed.), Normality and pathology in cognitive functions. (pp. 19-76). London: Academic Press.

Garrett, M. F. (1993). Errors and their relevance for models of language production. In G. Blanken, J. Dittman, H. Grim, J. Marshall, & C. Wallesch (Eds.), Linguistic disorders and pathologies. (pp. 72-92). Berlin, Germany: Walter de Gruyter.

Gawlitzek-Maiwald, I. (1997). Der monolinguale und bilinguale Erwerb von Infinitivstruktionen: ein Vergleich von Deutsch und Englisch. Tübingen: Niemeyer.

Genesee, F. (1989). Early bilingual development: One language or two? Journal of Child Language, 16, 161-179.

Gershkoff-Stowe, L. (1997). Early changes in lexical processes: Evidence from children's naming errors. Dissertation.Abstracts.International.: Section.B: The.Sciences.and Engineering., Vol 58(3-B), 1564

Gershkoff-Stowe, L., & Smith, L. B. (1997). A curvilinear trend in naming errors as a function of early vocabulary growth. Cognitive.Psychology., Vol 34(1), 37-71.

Goldfield, B. A., & Reznick, J. S. (1990). Early lexical acquisition: Rate, content, and the vocabulary spurt. Journal of Child Language, 17(171), 184

Golinkoff, R. M., & Hirsh-Pasek, K. (1995). Reinterpreting children's sentence comprehension: Toward a new framework. In P. Fletcher & B. MacWhinney (Eds.), The Handbook of Child Language. (pp. 430-461). Oxford: Blackwell.

Golinkoff, R. M., Hirsh-Pasek, K., Cauley, K. M., & Gordon, L. (1987). The eyes have it: Lexical and syntactic comprehension in a new paradigm. Journal of Child Language, 14, 23-45.

Granda-Rodriguez, G. B. (1998). Bilingualism and the onset of language. Dissertation.Abstracts.International.: Section.B: The.Sciences.and Engineering., Vol 59(4-B), 1883

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. Bilingualism.: Language and Cognition, 1, 67-81.

Grimm, H., & Doil, H. (2001). Elternfragebögen für die Früherkennung von Risikokindern (ELFRA). Göttingen: Testzentrale.

Grosjean, F. (1988). Exploring the recognition of guest words in bilingual speech. Language and Cognitive Processes, 3, 233-274.

Hahn, E. S. (1956). Stuttering: Significant Theories and Therapies. (2nd ed.). Stanford Calif: Stanford University Press.

Hall, C. (1992). Modern German Pronunciation. Manchester: Manchester University Press.

Hall, T. A. (1999a). Phonotactics and the prosodic structure of German function words. In T. A. Hall & U. Kleinhenz (Eds.), Studies on the Phonological Word. (pp. 99-132). Amsterdam: Benjamins.

Hall, T. A. (1999b). The phonological word: A review. In Anonymous, Studies on the Phonological Word. (pp. 1-22). Amsterdam: Benjamins.

Hall, T. A., & Kleinhenz, U. (1999). Studies on the phonological word. Amsterdam: John Benjamins.

Hamilton, A., Plunkett, K., & Schafer, G. (2000). Infant vocabulary development assessed with a British Communicative Development Inventory: Lower scores in the UK than in the USA. Journal of Child Language, 27(3), 689-705.

Hannahs, S. J. (1995a). Prosodic Structure and French Morphonphonology. Tübingen Germany: Niemeyer.

Hannahs, S. J. (1995b). The phonological word in French. Linguistics, 33, 1125-1144.

Harley, T. A., & Brown, H. E. (1998). What causes a tip-of-the-tongue state? Evidence for lexical neighbourhood effects in speech production. British Journal of Psychology, 89, 151-174.

Harley, T. A., & MacAndrews, S. B. G. (1995). Interactive models of lexicalization: Some constraints from speech error, picture naming, and neuropsychological data. In J. Levy, D. Bairaktaris, J. Bullinara, & D. Cairns (Eds.), (pp. 311-331). London: UCL Press.

Hartley, T., & Houghton, G. (1996). A linguistically constrained model of short-term memory for non-words. Journal of Memory and Language, 35, 1-35.

Hartmann, R. R. K., & Stork, F. C. (1972). Dictionary of language and linguistics. London: Applied Science Publishers.

Haynes, W. O., & Hood, S. B. (1977). Language and disfluency variables in normal speaking children from discrete chronological age groups. Journal of Fluency Disorders, 2, 57-74.

Hedge, M. N. (1982). Antecedents of fluent and dysfluent oral reading: A descriptive analysis. Journal of Fluency Disorders, 7, 323-341.

Hirsh-Pasek, K., & Golinkoff, R. M. (1996a). The intermodal preferential looking paradigm: A window onto emerging language comprehension. In D. McDaniel, C. McKee, & H. S. Cairns (Eds.), Methods for Assessing Children's Syntax. (pp. 105-127). Cambridge MA: MIT Press.

Hirsh-Pasek, K., & Golinkoff, R. M. (1996b). The Origins of Grammar: Evidence from Early Language Comprehension. Cambridge MA: MIT Press.

Hockett, C. F. (1954). Two models of grammatical description. Word, 10, 210-231.

Howell, P. (2002). The EXPLAN theory of fluency control applied to the treatment of stuttering by altered feedback and operant procedures. In E. Fave (Ed.), Current Issues in Linguistic Theory series: Pathology and therapy of speech disorders. (pp. 95-118). Amsterdam: John Benjamins.

Howell, P., & Au-Yeung, J. (2001). Application of EXPLAN theory to spontaneous speech control. University of Edinburgh: Edingburgh.

Howell, P., & Au-Yeung, J. (2002). The EXPLAN theory of fluency control and the diagnosis of stuttering. In E. Fave (Ed.), Pathology and therapy of speech disorders. In Current Issues in Linguistic Theory series. (pp. 74-94). Amsterdam: John Benjamins.

Howell, P., Au-Yeung, J., & Pilgrim, L. (1999). Utterance rate and linguistic properties as determinants of lexical dysfluencies in children who stutter. Journal of the Acoustical Society of America, 105, 481-490.

Howell, P., Au-Yeung, J., & Sackin, S. (2000). Internal structure of content words leading to lifespan differences in phonological difficulty in stuttering. Journal of Fluency Disorders, 25, 1-20.

Howell, P., Au-Yeung, J., Sackin, S., Glenn, K., & Rustin, L. (1997). Detection of supralexical dysfluencies in a text read by child stutterers. Journal of Fluency Disorders, 22, 299-307.

Howell, P., Au-Yeung, J., & Sackin, S. (1999). Exchange of stuttering from function words to content words with age. Journal of Speech, Language, and Hearing Research., Vol 42(2), 345-354.

Howell, P., Davis, S., & Au-Yeung, J. (2003). Norms for Reception of Syntax Test (ROST). Child Language and Teaching, 19, 1-21.

Howell, P., Ruffle, L., Fernández-Zúòiga, A., Gutiérrez, R., Fernández, A. H., O'Brien, M. L., Tarasco, M., Vallejo-Gomez, I., & Au-Yeung, J. (2003). Comparison of exchange patterns of stuttering in Spanish and English monolingual speakers and a bilingual Spanish-English speaker . IFA conference presentation: Montreal.

Hulk, A. (2000). Non-selective access and activation in child bilingualism: The syntax. In S. Döpke (Ed.), Cross-linguistic structures in simultaneous bilingualism. (pp. 57-78). Amsterdam: John Benjamins.

Hulk, A., & Müller, N. (2000). Bilingual first language acquisition at the interface between syntax and pragmatics. Bilingualism.: Language and Cognition, Vol 3(3), 227-244.

Hulk, A., & van der Linden, E. (1996). Language mixing in a French-Dutch bilingual child. In E. Kellerman, B. Weltens, & T. Bongaerts (Eds.), Eurosla 6: A selection of papers. (pp. 89-101).

Ingham, R. J., & Cordes, A. K. (1997). Identifying the authoritative judgments of stuttering: Comparisons of self-judgments and observer judgments. Journal of Speech and Hearing Research, Vol 40(3), 581-594.

Jakielski, K. J. (1998). Motor organization in the acquisition of consonant clusters. Dissertation / PhD thesis: University of Texas Austin.

Jankelwitz, D. L., & Bortz, M. A. (1996). The interaction of bilingualism and stuttering in an adult. Journal of Communication Disorders, 29, 223-234.

Jarayam, M. (1983). Phonetic influences on stuttering in monolingual and bilingual stutterers. Journal of Communication Disorders, 16, 278-297.

Jayaram, M. (1981). Grammatical factors in stuttering. Journal of the Indian Institute of Science of Science, 63, 141-147.

Jescheniak, J. D., & Levelt, W. J. M. (1994). Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. Journal of Experimental.Psychology.: Learning., Memory., and Cognition, Vol 20(4), 824-843.

Johnson, G. F. (1987). A clinical study of Porky Pig cartoons. Journal of Fluency Disorders, Vol 12(4), 235-238.

Johnson, S. (1998). Exploring the German language. London: Arnold Publishers.

Johnson, W. (1938). The role of evaluation in stuttering behavior. Journal of Speech Disorders, 3, 85-89.

Johnson, W. (1961). Stuttering and what you can do about it. Minneapolis: University of Minnesota Press.

Johnson, W., & Associates. (1959). The onset of stuttering. Minneapolis: The University of Minnesota Press.

Johnson, W., & Brown, S. F. (1935). Stuttering in relation to various speech sounds. The Quarterly Journal of Speech, 21, 481-496.

Johnson, W., & Brown, S. F. (1939). Stuttering in relation to various speech sounds: a correction. Quarterly.Journal of Speech, 25, 20-22.

Kadi-Hanifi, K., & Howell, P. (1992). Syntactic analysis of the spontaneous speech of normally fluent and stuttering children. Journal of Fluency Disorders, 17, 151-170.

Karniol, R. (1992). Stuttering out of bilingualism. First.Language, Vol 12(36, Pt 3), 255-283.

Kidd, K. K. (1980). Genetic models of stuttering. Journal of Fluency Disorders, 5, 187-201.

Kidd, K. K. (1984). Stuttering as a genetic disorder. In R. F. Curlee & W. H. Perkins (Eds.), Nature and treatment of stuttering. (pp. 149-169). San Diego: College Hill.

Kolk, H., & Postma, A. (1997). Stuttering as a covert repair phenomenon. In R. F. Curlee & G. M. Siegel (Eds.), The nature and treatment of stuttering: New directions. Boston MA: Allyn & Bacon.

Köpcke, K.-M. (2001). The acquisition of plural marking in English and German revisited: Schemata versus rules. In M. Tomasello & E. Bates (Eds.), Language development: The essential readings. Essential readings in developmental psychology. (pp. 203-226). Malden, MA, US: Blackwell Publishers Inc.

Köppe, R., & Meisel, J. M. (1995). Code-switching in bilingual first language acquisition. In L. Milroy & P. Muysken (Eds.), One Speaker - Two Languages: Cross-Disciplinary Perspectives on Code-Switching. (pp. 276-301). Cambridge UK: Cambridge University Press.

Korzybski, A. H. S. (1933). Science and Sanity: An Introduction to Non-Aristotelian Systems and General Semantics. Lancaster PA: Science Press Printing Co.

Kroger, B. J. (1993). A gestural production model and its application to reduction in German. Phonetica, 50, 213-233.

Kuhl, P. K., & Meltzhoff, A. (1997). Evolution, nativism and learning in the development of language and speech. In M. Gopnik (Ed.), The inheritance and innateness of grammars. (pp. 7-44). New York: Oxford University Press.

Lanza, E. (1992). Can bilingual two-year-olds code-switch? Journal of Child Language, 19, 633-658.

Lebrun, Y., & Paradis, M. (1984). To be or not to be an early bilingual? In Y. Lebrun & M. Paradis (Eds.), Early bilingualism and child development. (pp. 9-18). Lisse: Swets & Zeitlinger.

Leopold, W. F. (1950). The speech development of a bilingual child. Evanston, Illinois: Northwestern University Press.

Leske, M. C. (1981). Prevalence estimates of communicative disorders in the U.S.: Speech disorders. ASHA, 23, 217-225.

Levelt, W. J. M. (1983). Monitoring and self-repair in speech. Cognition, 14, 41-104.

Levelt, W. J. M. (1989). Speaking: From intention to articulation. Cambridge, MA: MIT Press.

Levelt, W. J. M. (1992). Accessing words in speech production: Stages processes and representations. Cognition, 42, 1-22.

Levelt, W. J. M. (2000). Introduction to the section on language processing. In M. S. Gazzaniga (Ed.), Cognitive neuroscience : a reader. (pp. 843-844). Malden, Mass.: Blackwell.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. Behavioral.and Brain Sciences., Vol 22(1), 1-75.

Lofqvist, A. (1990). Speech as audible gesture. In W. J. Hardcastle & A. Marchal (Eds.), Speech Production and Speech Modelling. (pp. 289-322). Dordrecht, Netherlands: Kluwer Academic Publishers.

Louttit, C. M., & Halls, E. C. (1936). Survey of speech defects among public school children of Indiana. Journal of Speech Disorders, 1, 73-80.

MacKay, D. G. (1981). The problem of rehearsal or mental practice. Journal of Motor Behaviour, 13, 274-285.

MacKay, D. G. (1987). The organization of perception and action: A theory for language and other cognitive skills. New York: Springer-Verlag.

Maclay, H., & Osgood, C. E. (1959). Hesitation phenomena in spontaneous English speech. Word, 15, 169-182.

Macnamara, J. (1966). Bilingualism and primary education. Edinburgh: Edinburgh University Press.

MacNeillage, P., & Davis, B. (1990). Acquisition of speech production: Frames, then content. In M. Jeannerod (Ed.), Attention and performance XIII: Motor representation and control. Hillsdale: Lawrence Erlbaum.

MacWhinney, B. (1995). The CHILDES project.Tools for analyzing talk. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

MacWhinney, B. (1997). Second language acquisition and the competition model. In A. De Groot & J. Kroll (Eds.), Tutorials in bilingualism: psychological perspectives. (pp. 113-142). Hillsdale, NJ: Lawrence Erlbaum:

Martin, N., Dell, G. S., Saffran, E. M., & Schwartz, M. F. (1994). Origins of paraphasias in deep dysphasia: Testing the consequences of a decay impairment to an interactive spreading activation model of lexical retrieval. Brain and Language, 47, 609-660.

McClelland, J., & Elman, J. (1986). The TRACE model of speech perception. Cognitive Psychology, 18(1), 86

McDevitt, S. C., & Carey, W. B. (1975). Behavioral Style Questionnaire. Scottsdale, AZ: Behavioral-Developmental Initiatives.

Mehler, J., Jusczyk, P. W., Lambertz, G., Halsted, N., Bertoncini, J., & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. Cognition, 29, 143-178.

Meisel, J. M. (1989). Early differentiation of languages in bilingual children. In K. Hyltenstam & L. Obler (Eds.), Bilingulaism across the lifespan: Aspeacts of acquisition, maturity, and loss. (pp. 13-40). Cambridge: Cambridge University Press.:

Meisel, J. M. (1990). Grammatical development in the simultaneous acquisition of two first languages. In J. M. Meisel (Ed.), Two first languages: Early grammatical development in bilingual children. (pp. 5-22). Dordrecht: Foris.

Meisel, J. M. (2001). The simultaneous acquisition of two first languages: Early differentiation and subsequent development of grammars. In J. E. Cenoz & F. E. Genesee (Eds.), Trends in bilingual acquisition. Trends in language acquisition research, vol. 1. (pp. 11-41). Amsterdam, Netherlands: John Benjamins Publishing Company.

Meisel, J. M., & Müller, N. (1992). Finiteness and verb placement in early child grammars. Evidence from simultaneous acquisition of French and German in bilinguals. In J. M. Meisel (Ed.), The acquisition of verb placement: Functional categories and V2 phenomena in language acquisition. (pp. 109-138). Dordrecht: Kluwer Academic Publishers.

Merriman, W. E., & Kutlesic, V. (1993). Bilingual and monolingual children's use of two lexical acquisition heuristics. APPLIED PSYCHOLINGUISTICS, 14, 229-249.

Metz, D. E., Conture, E. G., & Colton, R. H. (1976). Temporal relations between the respiratory and laryngeal systems prior to stuttered disfluencies. ASHA, 18, 664

Miller, J. F., & Chapman, R. S. (1981a). The relations between Age and Mean Length of Utterance. Journal of Speech and Hearing Research, 24, 154-161.

Miller, J. F., & Chapman, R. S. (1981b). The relations between age and mean length of utterance. Journal of Speech and Hearing Research, 24, 154-161.

Miller, N., & Abudarham, S. (1984). Management of communication problems in bilingual children. In N. Miller (Ed.), Bilingualism and language disability-assessment and remediation. London: Croom Helm.

Mugdan, J. (1977). Flexionsmorphologie und Psycholinguistik. Untersuchungen zu sprachlichen Regeln und ihrer Beherrschung durch Aphatiker, Kinder und Ausländer, am Beispiel der deutschen Substantivdeklination. Tübingen: Narr.

Müller, N. (1990). Developing two gender assignment systems simultaneously. In J. M. Meisel (Ed.), Two first languages: Early grammatical development in bilingual children. (pp. 193-234). Dordrecht: Foris.

Müller, N. (1998). Transfer in bilingual first language acquisition. Bilingualism.: Language and Cognition, 1(3), 151-171.

Mussafia, M. (1967). Plurilingisme et troubles du langage. Folia Phoniatrica, 19, 63-68.

Muysken, P. (2000). Bilingual speech. Cambridge UK: Cambridge University Press.

Natke, U., Sandrieser, P., Pietrowsky, R., & Kalveram, K. T. (2001). Stuttering and syllabic stress in preschool children: Preliminary observations. In B. Maasen, W. Hulstijn, R. D. Kent, & P. H. H. M. van Lieshout (Eds.), Speech Motor Control in Normal and Disordered Speech. (pp. 258-261). Nijmegen, The Netherlands: Uitgeverij Vantilt.

Nespor, M., & Vogel, I. (1986). Prosodic Phonology. Dortrecht - Holland: Foris Publications.

Nicoladis, E. (2002). The cues that children use in acquiring adjectival phrases and compounds nouns: Evidence from bilingual children. Brain and Language, Vol 81(1-3), 635-648.

Nicoladis, E. (2003a). Cross-linguistic transfer in deverbal compounds of preschool bilingual children. Bilingualism.: Language and Cognition, 6(1), 17-31.

Nicoladis, E. (2003b). What compound nouns mean to preschool children. Brain and Language, 84(1), 38-49.

Nicoladis, E., & Genesee, F. (1997). Language development in preschool bilingual children. Journal of Speech-Language Pathology and Audiology, 21(4), 258-270.

Nippold, M. (1990). Concomitant speech and language disorders in stuttering children: A critique of the literature. Journal of Speech and Hearing Disorders, 55, 51-60.

Nippold, M., Schwarz, I. E., & Jescheniak, J. (1991). Narrative ability in school-age stuttering boys: A preliminary investigation. Journal of Fluency Disorders, 16, 289-308.

Nwokah, E. E. (1988). The imbalance of stuttering behavior in bilingual speakers. Journal of Communication Disorders, 13, 357-373.

Oller, D. K. (1988). The emergence of sounds of speech in infancy. In G. H. Yeni-Komshian, J. F. Kavanaugh, & C. A. Ferguson (Eds.), Child phonology: Vol. 1. Production. (pp. 93-112). New York: Academic Press.

Osterhout, L., Allen, M., & McLaughlin, J. (2002). Words in the brain: Lexical determinants of word-induced brain activity. Journal of Neurolinguistics., Vol 15(3-5), 171-187.

Paradis, J. C., & Genesee, F. (1995). Language differentiation in early bilingual development. Journal of Child Language, 22, 611-631.

Paradis, J. C., & Genesee, F. (1996). Syntactic acquisition in bilingual children: Autonomous or independent? Studies in Second Language Acquisition, 18, 1-15.

Paradis, J. C., & Genesee, F. (1997). On continuity and the emergence of functional categories in bilingual first language acquisition. Language Acquisition, 6, 91-124.

Pearson, B. Z., & Fernández, S. C. (1994). Patterns of interaction in the lexical growth in two languages of bilingual infants and toddlers. Language Learning, 44, 617-653.

Pearson, B. Z., Fernández, S. C., & Oller, D. K. (1993). Lexical development in bilingual infants and toddlers: Comparison to monolingual norms. Language Learning, 43, 93-120.

Peters, T., & Guitar, B. (1991). Stuttering, an integrated approach to its nature and treatment. In Anonymous, Baltimore: Williams & Wilkins.

Pichon, E., & Borel-Maisonny, S. (1964). Le bégaiement. Sa nature e son traitment. Paris: Masson.

Pinker, S. (1995). Language Acquisition. In L. R. Gleitman & M. Liberman (Eds.), An Invitation to Cognitive Science.Vol. 1: Language. Cambridge, MA: MIT Press.

Pinker, S. (1999). Words and Rules. London: Weidenfeld & Nicolson.

Postma, A., & Kolk, H. (1993). The covert repair hypothesis: Prearticulatory repair processes in normal and stuttered disfluencies. Journal of Speech and Hearing Research., Vol 36(3), 472-487.

Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). A comprehensive grammar of the English language. London: Longman.

Raffelsiefen, R. (1999). Diagnostics for prosodic words revisited: The case of historically prefixed words in English. In A. T. Hall & U. Kleinhenz (Eds.), Studies on the Phonological Word. (pp. 133-202). Amsterdam: John Benjamins Publishing Co.

Richards, I. A., Schmidt Mackey, I., Mackey, W. F., & Gibson, C. (1960). German through pictures. New York: Washington Square Press, Inc.

Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. Cognition, 42, 107-142.

Roelofs, A. (2000). WEAVER++ and other computational models of lemma retrieval and word-form encoding. In L. Wheeldon (Ed.), Aspects of language production. Studies in cognition series. (pp. 71-114). Philadelphia, PA, US: Psychology Press.

Roelofs, A., Meyer, A. S., & Levelt, W. J. M. (1998). A case for the lemma/lexeme distinction in models of speaking: comment on Caramazza and Miozzo (1997). Cognition, 69, 219-230.

Rogers, J. (2000). The phonatory correlates of juncture in German. Proceedings of the 5th Seminar on Speech Production: Models and Data: May 1st to 4th, 2000, Kloster Seeon, Bavaria, Germany.

Romaine, S. (1999). Bilingual language development. In M. Barrett (Ed.), The development of language. (pp. 251-275). Hove, UK: Psychology Press.

Rommel, D. (2001). The influence of psycholinguistic variables on stuttering in childhood. In H.-G. Bosshardt, J. S. Yaruss, & H. F. M. Peters (Eds.), Fluency disorders: Theory, research, treatment and self-help. Proceedings of the third world congress of fluency disorders in Nyborg, Denmark. (pp. 195-202). Nijmegen: Nijmegen University Press.

Rommel, D., Häge, A., Johannsen, H. S., & Schulze, H. (1997). Linguistic aspects of stuttering in childhood. In W. Hulstijn, H. F. M. Peters, & P. H. H. M. van Lieshout (Eds.), Speech Production: Motor Control, Brain Research and Fluency Disorders. (pp. 603-610). Amsterdam: Elsevier Science Publishers.

Rommel, D., Häge, A., Kalehne, P., & Johannsen, H. S. (2000). Development, maintenance, and recovery of childhood stuttering: Prospective longitudinal data 3 years after first contact. In K. L. Baker, L. Rustin, & F. Cook (Eds.), Proceedings of the Fifth Oxford Dysfluency Conference. (pp. 168-182). Leicester UK: Kevin L. Baker.

Rosenblum, T., & Pinker, S. (1983). Word magic revisited: Monolingual and bilingual children's understanding of the word-object relationship. Child Development, 54, 773-780.

Ruke-Dravina, V. (1967). Mehrsprachigkeit im Vorschulalter. Lund: Gleerup.

Rustin, L., Botterill, W., & Kelman, E. (1996). Assessment and therapy for young dysfluent children: Family ineraction. London: Whurr Publishers.

Saltzman, E. L., & Munhall, K. G. (1989). A dynamical approach to gestural patterning in speech production. Ecological Psychology, 1, 333-382., 1, 333-382.

Sander, E. K. (1972). When are speech sounds learned? Journal of Speech and Hearing Disorders, 37, 55-63.

Santiago, J., MacKay, D. G., Palma, A., & Rho, C. (2000). Sequential activation processes in producing words and syllables: Evidence from picture naming. Language and Cognitive Processes, 15(1), 1-44.

Schriefers, H. (1990). Lexical and conceptual factors in the naming of relations. Cognitive Psychology, 22, 111-142.

Schwartz, B. L., Travis, D. M., Castro, A. M., & Smith, S. M. (2000). The phenomenology of real and illusory tip-of-the tongue states. Memory & Cognition, 28, 18-27.

Schwartz, B. L. (2002). Tip-of-the-tongue states: Phenomenology, mechanism, and lexical retrieval. Mahwah, NJ: Lawrence Erlbaum Associates.

Scott Trautman, L., & Keller, K. (2000). Bilingual intervention for stuttering: a case in point. Paper presented at the ASHA annual convention, Washington DC: November 16-19.

Segalowitz, S. J., & Lane, K. C. (2000). Lexical access of function versus content words. Brain and Language, Vol 75(3), 376-389.

Selkirk, E. (1978). On prosodic structure and its relation to syntactic structure. In T. Fretheim (Ed.), Nordic Prosody II. (pp. 75-88). Trondheim: TAPIR.

Selkirk, E. (1980a). Prosodic domains in phonology: Sanskrit revisited. In M. Aronoff & M. Kean (Eds.), Juncture. A collection of Original Papers. (pp. 42-58). Saratoga: Amma Libri.

Selkirk, E. (1980b). The role of prosodic categories in English word stress. Linguistic Inquiry, 11, 563-605.

Selkirk, E. (1984). Phonology and syntax: The relation between sound and structure. Cambridge, MA: MIT Press.

Selkirk, E. (1996). The prosodic structure of function words. In J. L. Morgan & K. Demuth (Eds.), Signal to syntax: Bootstrapping from speech to grammar in early acquisition. (pp. 187-215). Mahwah NJ: Lawrence Erlbaum Associates Publishers.

Shapiro, A. I. (1980). An electromyographic analysis of the fluent and dysfluent utterances of several types of stutterers. Journal of Fluency Disorders, 5, 203-231.

Shattuck-Hufnagel, S. (1992). The role of word structure in segmental serial ordering. Cognition, 42, 213-259.

Sheehan, J. G. (1953). Theory and treatment of stuttering as an approach-avoidance conflict. Journal of Psychology, 36, 27-29.

Shenker, R. C., Conte, A., Gingras, A., Courcey, A., & Polomeno, L. (1998). The impact of bilingualism on developing fluency in a preschool child. In E. C. Healey & H. F. M. Peters (Eds.), Second world congress on fluency disorders proceedings, San Francisco, August 18-22. (pp. 200-204). Nijmegen: Nijmegen University Press.

Shi, R., & Werker, J. F. (2001). Six-month old infants' preference for lexical words. Psychological Science, Vol 12(1), 70-75.

Siegel, G. M. (2000). Demands and capacities or demands and performance? Journal of Fluency Disorders, Vol 25(4), 321-327.

Silverman, F. H. (1974). Disfluency behavior of elementary school stutterers and non-stutterers. Language, Speech and Hearing Services in Schools, 5, 32-37.

Slobin, D. I. (1970). Universals of grammatical development in children. In G. B. Flores d'Arcais & W. J. M. Levelt (Eds.), Advances in Psycholinguistics. (pp. 87-131). Amsterdam: North Holland Publishing.

Smith, A. (1999). Stuttering: A Unified Approach to a Multifactorial, Dynamic Discorder. In N. Bernstein Ratner & E. C. Healey (Eds.), Stuttering Research and Practice: Bridging the Gap. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Smith, G. (1996). Das 'phonologische Wort' in der nicht-linearen Phonologie. M.A. thesis. Berlin: Free University of Berlin.

Sommer, M., Koch, M., Paulus, W., Weiller, C., & Büchel, C. (2002). Disconnection of spech-relevant brain areas in persistent developmental stuttering. Lancet, 360, 380-383.

Spencer, A. (1996). Phonology. Oxford: Blackwell Publishers Ltd.

Starkweather, C. W. (1987). Fluency and stuttering. Englewood Cliffs, NJ: Prentice-Hall.

Starkweather, C. W. (1997). Therapy for younger children. In R. F. Curlee & G. M. Siegel (Eds.), Nature and treatment of stuttering: New directions. (pp. 257-279). Boston: Allyn and Bacon.

Starkweather, C. W., & Gottwald, S. R. (1990). The demands and capacities model: II. Clinical applications. Journal of Fluency Disorders, Vol 15(3), 143-157.

Starkweather, C. W., Gottwald, S. R., & Halfond, M. H. (1990). Stuttering prevention: A clinical method. Englewood Cliffs, NJ.: Prentice-Hall.

Starkweather, C. W., & Gottwald, S. R. (2000). The Demands and Capacities Model: Response to Siegel. Journal of Fluency Disorders, Vol 25(4), 369-375.

Stemberger, J. P. (1985). Phonological rule ordering in a model of language production. Bloomington: Indiana University Linguistics Club.

Stern, E. (1948). A preliminary study of bilingualism and stuttering in four Johannesburg schools. J.Logopaed., 1, 15-25.

Sternberg, S., Knoll, R. L., Monsell, S., & Wright, C. E. (1988). Motor programs and hierachical organization in the control of rapid speech. Phonetica, 45, 175-197.

Strand, E. A. (1992). The integration of speech motor control and language formulation in process models of acquisition. In R. S. Chapman (Ed.), Processes in language acquisition and disorders. (pp. 78-92). St. Louis: Mosby-Year Book, Inc.

Szagun, G. (2004). Learning by ear: on the acquisition of case and gender marking by German-speaking children with normal hearing and cochlear implants. Journal of Child Language, 31, 1-30.

Throneburg, R. N., Yairi, E., & Paden, E. P. (1994). Relation between phonologic difficulty and the occurrence of disfluencies in the early stage of stuttering. Journal of Speech and Hearing Research, Vol 37(3), 504-509.

Thürmer, S., Thurmfart, W., & Kittel, G. (1983). Elektomyographische Untersuchungsbefunde bei Stotterern. Sprache-Stimme-Gehör, 7, 125-127.

Tomasello, M., & Abbot-Smith, K. (2000). First steps toward a usage-based theory of language acquisition. Cognitive Linguistics, Vol 11(1-2), 61-82.

Tomasello, M., & Abbot-Smith, K. (2002). A tale of two theories: Response to Fisher. Cognition, Vol 83(2), 207-214.

Tracy, R. (1995). Child language in contact: Bilingual language acquisition (English/German) in early childhood. Habilitationsschrift: University of Tübingen.

Travis, L. E., Johnson, W., & Shover, J. (1937). The relation of bilingualism to stuttering. Journal of Speech Disorders, 2, 185-189.

Umbel, V. M., Pearson, B. Z., Fernandez, M. C., & Oller, D. K. (1992). Measuring bilingual children's receptive vocabularies. Child Development, Vol 63(4), 1012-1020.

Vaane, E., & Janssen, P. (1978). Different types of disfluencies and some phonological factors. Logoped.Foniat., 50, 14-20.

Van Borsel, J., Maes, E., & Foulon, S. (2001). Stuttering and bilingualism: A review. Journal of Fluency Disorders, 26, 179-205.

Van Riper, C. (1937). The preparatory set in stuttering. Journal of Speech Disorders, 2, 149-154.

Van Riper, C. (1971). The nature of stuttering. Englewood Cliffs: Prentice Hall.

Van Riper, C. (1981). The nature of stuttering. (2nd ed.). Englewood Cliffs: Prentice Hall.

Vigário, M. (1999). On the prosodic status of stressless function words in European Portuguese. In T. A. Hall & U. Kleinhenz (Eds.), Studies on the Phonological Word. (pp. 255-295). Amsterdam: Benjamins.

Vigliocco, G., Antonini, T., & Garrett, M. F. (1997). Grammatical gender is on the tip of Italian tongues. Psychological Science, 8, 314-317.

Vigliocco, G., & Hartsuiker, R. J. (2002). The interplay of meaning, sound, and syntax in sentence production. Psychological Bulletin, 128(3), 442-472.

Vigliocco, G., Vinson, D., Martin, R. C., & Garrett, M. F. (1999). Is "count" and "mass" information available when the noun is not? An investigation of tip of the tongue states and anomia. Journal of Memory and Language, 40, 534-558.

Vihman, M. M. (1996). Phonological development: The origins of language in the child. Oxford: Blackwell, Ltd.

Volterra, V., & Taeschner, T. (1978). The acquisition and development of language by bilingual children. Journal of Child Language, 5, 311-326.

Walker, M. (1999). Flutter stutter. New Scientist, 164(2209), 23-24.

Wallin, J. E. W. (1916). A census of speech defectives among 89,057 public-school pupils-a preliminary report. Scholastic Scociety, 3, 213-216.

Weber-Fox, C. (2001). Neural systems for sentence processing in stuttering. Journal of Speech, Language, and Hearing Research, Vol 44(4), 814-825.

Wellman, H. M. (1977). Tip of the tongue and feeling of knowing experiences: A developmental study of memory monitoring. Child Development, 48, 13-21.

Wells-Jensen, S. B. (2000). Cognitive correlates of linguistic complexity: A cross-linguistic comparison of errors in speech. Dissertation.Abstracts.International.Section.A: Humanities.and Social.Sciences., Vol 60(8-A), 2899

Wexler, K. B. (1982). Developmental disfluency in 2-, 4- and 6-year-old boys in neutral and stress situations. Journal of Speech and Hearing Research, 25, 229-234.

Wexler, K. B., & Mysak, E. D. (1982). Disfluency characteristics of 2-, 4- and 6-year-old males. Journal of Fluency Disorders, 7, 37-46.

Whorf, B. L. (1940). Science and Linguistics. Technological Review (MIT), 42, 229-231.

Wiese, R. (1996). The Phonology of German. Oxford: Clarendon Press.

Wijnen, F. (1990). The development of sentence planning. Journal of Child Language, Vol 17(3), 651-675.

Wijnen, F. (1992). Incidental word and sound errors in young speakers. Journal of Memory.and Language, Vol 31(6), 734-755.

Wingate, M. E. (1979). The loci of stuttering: grammar or prosody? Journal of Communication Disorders, 12, 283-290.

Wingate, M. E. (1984). Stutter events and linguistic stress. Journal of Fluency Disorders, 9, 295-300.

Wingate, M. E. (1988). The Structure of Stuttering. New York: Springer-Verlag.

Wingate, M. E. (2002). Foundations of Stuttering. San Diego CA: Academic Press.

Wingate, M. E. (2003). Major Problems with a revisit. Journal of Speech, Language and Hearing Research, *in press*

Yairi, E. (1982). Longitudinal studies of disfluencies in two-year-old children. Journal of Speach and Hearing Research, 25, 155-160.

Yairi, E. (1983). The onset of stuttering in two- and three-year old children: A preliminary report. Journal of Speech and Hearing Disorders, 48, 171-178.

Yairi, E., & Ambrose, N. (1992a). A longitudinal study of stuttering in children: A preliminary report. Journal of Speech and Hearing Research, Vol 35(4), 755-760.

Yairi, E., & Ambrose, N. (1992b). Onset of stuttering in preschool children: Selected factors. Journal of Speech and Hearing Research, Vol 35(4), 782-788.

Yairi, E., Ambrose, N., & Cox, N. (1996). Genetics of stuttering: A critical review. Journal of Speech and Hearing Research, 39, 771-784.

Yairi, E., & Ambrose, N. G. (1999). Early childhood stuttering I: Persistency and recovery rates. Journal of Speech, Language, and Hearing Research, Vol 42(5), 1097-1112.

Yairi, E., Paden, E. P., Ambrose, N., & Throneburg, R. (1994). Pathways of chronicity and recovery: Longitudinal studies of early childhood stuttering. Presentation at American Speech-Language-Hearing Association annual convention: New Orleans.

Yairi, E., Watkins, R., Ambrose, N., & Paden, E. (2001). What is stuttering? Journal of Speech, Language, and Hearing Research, Vol 44(3), 585-592.

Yu, S. Y. (1992). Unterspezifikation in der Phonologie des Deutschen. Tübingen: Niemeyer.

Zebrowski, P. M. (1991). Duration of the speech disfluencies of beginning stutterers. Journal of Speech and Hearing Research, Vol 34(3), 483-491.

Zebrowski, P. M. (1994). Duration of sound prolongation and sound/syllable repetition in children who stutter: Preliminary observations. Journal of Speech and Hearing Research, Vol 37(2), 254-263.

Zebrowski, P. M. (1995). The topography of beginning stuttering. Journal of Communication Disorders, Vol 28(2), 75-91.

Zürn, S. R. (1994). Das Ravensburger Lexicon der Natur & Technik. Ravensburg (Germany): Ravensburger Buch Verlag.

11 Appendices

Appendix 1:

Language Usage Survey

I. General Information

Subject's name: ______________________________

Today's Date: ______________________________

Gender: ☐ Male ☐ Female

Date of birth: ______________________________
Place of birth: ______________________________

If not born in the UK,
when did you come to the UK ______________________________

Number of brothers and sisters ______________________________

Handedness: ☐ Right-handed ☐ Left-handed ☐ Both

II. Language Usage:

First Language (L1): (the best language you speak)

Language: ☐ English ☐ other ____________________

Age when you first used this language: ________________
(enter 0 if from birth)

How good are you in using this language?

Speaking: ☐Very well ☐ Well ☐ Not very well ☐ Not at all

Listening: ☐ Very well ☐ Well ☐ Not very well ☐ Not at all

Reading: ☐ Very well ☐ Well ☐ Not very well ☐ Not at all

Writing: ☐ Very well ☐ Well ☐ Not very well ☐ Not at all

Second Language (L2): (if any)

Language: ☐ English ☐ other ____________________

Age when you first used this language: ________________
(enter 0 if from birth)

How good are you in using this language?

Speaking: ☐Very well ☐ Well ☐ Not very well ☐ Not at all

Listening: ☐ Very well ☐ Well ☐ Not very well ☐ Not at all

Reading: ☐ Very well ☐ Well ☐ Not very well ☐ Not at all

Writing: ☐ Very well ☐ Well ☐ Not very well ☐ Not at all

Do you speak any other languages? Yes ☐ No ☐
If 'yes' what are they?
a. ________________ b. ____________________ c. __________________

Reason for learning more than 1 language: (you may tick more than one box)

- ☐ Family members do not speak English.
- ☐ Parents speak more than one language.
- ☐ Learning a second language as a subject at school
- ☐ other reasons ______________________________

III. Language use in different environments:

1. The language(s) used when **you** speak to **your mother**
   a. ______________ b. ________________ c. ________________

2. The language(s) used when **you** speak to **your father**
   a. ______________ b. ________________ c. ________________

3. The language(s) used when **your mother** speaks to **you**
   a. ______________ b. ________________ c. ________________

4. The language(s) used when **your father** speaks to **you**
   a. ______________ b. ________________ c. ________________

5. The language(s) used when **you** speak to **your brothers/sisters**
   a. ______________ b. ________________ c. ________________

6. The language(s) used when **you** speak to **your friends**
   a. ______________ b. ________________ c. ________________

7. The language(s) used when **your mother** speaks to **your father**
   a. ______________ b. ________________ c. ________________

8. The language(s) used when **your father** speaks to **your mother**
   a. ______________ b. ________________ c. ________________

9. The language(s) spoken by **your mother**
   a. ______________ b. ________________ c. ________________

10. The language(s) spoken by **your father**
    a. ______________ b. ________________ c. ________________

11. The language(s) spoken by **your grandparents**

    A. Grandfather 1
    a. ______________ b. ________________ c. ________________

    B. Grandmother 1
    a. ______________ b. ________________ c. ________________

C. <u>Grandfather 2</u>
a. ________________ b. ___________________ c. __________________

D. <u>Grandmother 2</u>
a. ________________ b. ___________________ c. __________________

Appendix 2:

Graph indicating variability in the three language groups – chapter 6

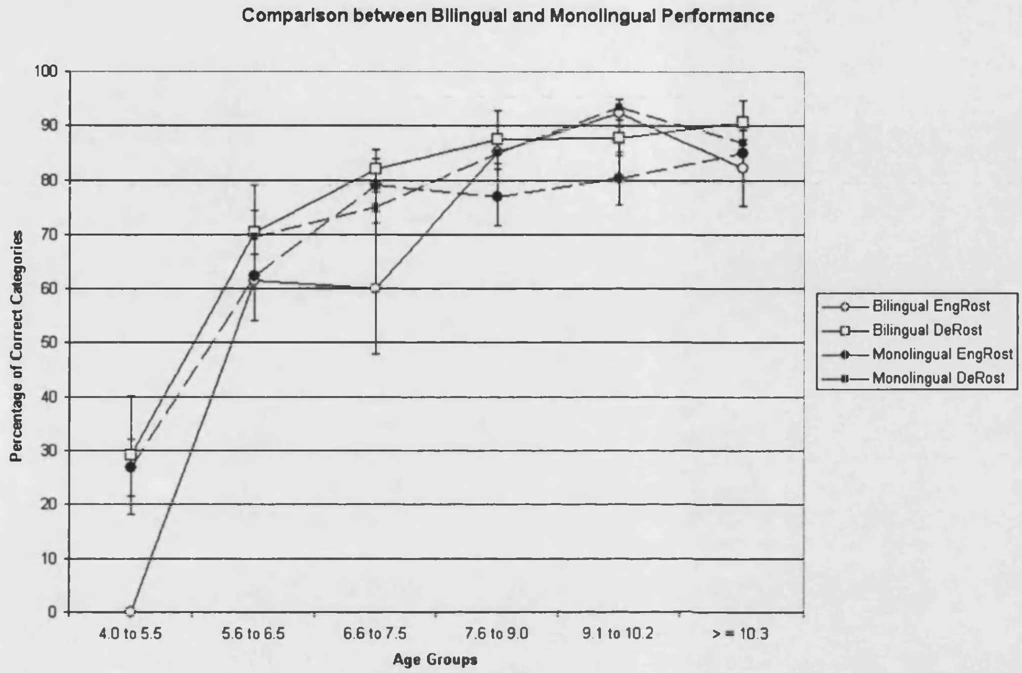


Figure 39: Appendix 2 – individual group means plus/minus one standard error. See Figure 25 chapter 6.

Appendix 3:

TOT stimuli used in chapter 8

TOT Stimuli – children aged 9-11 (these were randomised in presentation –presentation is here in order of syllable length) – equal numbers of vowel/consonants and word gender each in one, two and three syllable words	
Word	Definition
der Elch	Größte und schwerste Hirschart, bei dem die männlichen Tiere oft mächtig entwickelte (bis zu 20kg schwere) meist schaufelförmige Geweihe besitzen.
der Aal	Schlangenförmiger räuberischer Knochenfisch, der auch ein wertvoller Speisefisch in Süßgewässern Europas ist.
der Ast	Der Teil eines Baumes an dem die Blätter hängen.
der Bug	Vorderteil einen Schiffes oder Flugzeugs.
der Dachs	Gattung der Marder, etwa 70cm lang, plump und relativ kurzbeinig mit langem Schwanz, es gräbt einen Erdbau.
der Strauß	Größter heute lebender Vogel der aus Afrika stammt.
die Alm	In Hoch- und Mittelgebirgen vielfach auch unter der natürlichen Waldgrenze gelegene Bergweide - auch von Heide bekannt.
die Axt	Ein Gerät zum Zerkleinern von Holz.
die Uhr	Etwas womit man die Zeit ablesen kann.
die Tracht	Traditionelle Kleidung verschiedener Völker, Stämme, Volksgruppen oder Berufsgruppen.
die Laus	Ein Tier welches in Haaren lebt und Juckreitz hervorruft.
die Yacht	Eine Art von Sport- oder Freizeitboot.
das Pult	Eine Art Schultisch oder auch Stehtisch bei Vorträgen.
das Pik	Eine der schwarzen Farben bei Spielkarten.
das Schilf	Pflanzen am Ufer von Gewässern.
das All	Ein anderes Wort für den Welltraum.
das Öhr	Obere Teil eine Nähnadel durch den man den Faden zieht.
das Erz	Allgemeine Bezeichnung für ein Gestein, dass Metalle enthält.
die Arche	Ein biblisches Schiff worin Menschen und Tiere der Sinnflut entkamen.
die Bahre	Eine Liege auf denen man kranke Menschen transportiert.
die Ebbe	Eine Gezeit des Meeres, dass sich an der Küste durch Absinken des Wasserspiegels äußert.
die Öse	Ein meist runder Metallring zum Befestigen und Einharken anderer Materialien.
die Boje	Tonnenförmiger, an der Wasseroberfläche schwimmender Hohlkörper, zur Markierung des Fahrwassers.
die Kreide	Weiches, feinkörniges Gestein mit dem man auf Tafeln schreiben kann.
das Cello	Ein Musikinstrument, größer als eine Geige und kleiner als ein Kontrabaß.
das Iglu	Runde Schneehütte der Eskimos.
das Atom	Kleinster, elektrisch neutraler Bestandteil eines chemischen Elements.
das Fossil	Versteinerte Überreste eines Lebewesens (auch Versteinerungen genannt).
das Ferkel	Bezeichnung für junge Schweine.
das Abitur	Das Schulabschlußzeugniss, dass man am Ende des Gymnasiums bekommt.
der Erker	Ein oder mehrgeschossiger Vorbau (im Obergeschoßbereich).
der Imker	Jemand der Bienen hält und züchtet zur Gewinnung von Honig und Wachs.
der Sockel	Etwas auf das man ein Denkmal oder ein wichtiges Object stellt.
der Otter	Ein Wassertier, dass Fische fängt.
der Falke	Eine weit verbreitete Art der Raubvögel.
der Puma	Eine Großkatze mit braunem bis silbergrauem Fell, die wild nur noch in Amerika zu finden ist.
der Abakus	Seit der Antike verwendetes Rechenbrett mit frei beweglichen Steinen für die vier Grundrechnungsarten.
der Bumerang	Gewinkeltes oder leicht gebogenes Wurfholz der Eingeborenen Australiens, das beim Verfehlen des Zieles zum Werfer zurückkehrt.

der Ultraschall	Für den Menschen unhörbare Töne, die in der Medizin und Technik vielseitig verwendet wird.
der Akrobat	Menschen die im Zirkus arbeiten und virtuose Körperbeherrschung und Geschicklichkeit besitzen.
der Pelikan	Wasservogel mit großem Schnabel für die Fischbeute.
der Samurai	Bewaffnete Krieger aus dem alten Japan.
die Achterbahn	Eine Attraktion in Vergnügungsparks.
die Oktave	Das Interval, das vom Grundton acht diatonische Stufen entfernt ist.
die Erdkunde	Ein Schulunterricht in dem man Landkarten lesen lernt.
die Reißzwecke	Etwas das man in einem Büro benutzt um Papier an die Wand zu heften.
die Mumie	Durch Einbalsamierung vor Verwesung geschützter Leichnam, besonders im alten Ägypten zu finden.
die Tapete	Papier mit dem man die Wand behängt.
das Bullauge	Dickverglastes, rundes Schiffsfenster.
das Abendrot	Die Farbe des Himmels in der Dämmerungszeit kurz nach Sonnenuntergang.
das Meerschweinchen	Ein kleines Nagetier, dass auch als Haustier gehalten wird.
das Elfenbein	Das Zahnbein der Stoßzähne des Afrikanischen und Indischen Elefanten, wurde unter anderem verwendet für Klaviertasenbelag.
das Mosaik	Flächendekoration wobei Steinchen nach einer Vorzeichnung dicht in ein feuchtes Mörtelbrett gesetzt und später poliert werden.
das Megaphon	Ein Sprachrohr in Trichterform mit dem man die Stimme lauter machen kann.

Younger age group	
Word	Definition
der Hai	Ein gefährlicher Fisch mit scharfen Zähnen.
der Blitz	Ein Licht, dass plötzlich bei einem Gewitter erscheint.
der Aal	Schlangenförmiger Fisch den man essen kann (zum Beispiel geräuchert).
der Brief	Eine schriftliche Mitteilung, die verschlossen in einem Umschlag verschickt wird.
der Arm	Ein Teil des Oberkörpers.
der Ast	Ein Teil des Baumes an dem die Blätter hängen.
die Uhr	Etwas womit man Zeit ablesen kann.
die Axt	Ein Werkzeug, dass man zum Fällen von Bäumen benutzt.
die Alm	Eine Wiese in den Alpen (auch von Heidi bekannt).
die Faust	Eine Hand bei der die Finger zusammengeballt sind.
die Gans	Ein Tier mit Federn, dass häufig zur Weihnachtszeit gegessen wird.
die Nuß	Eine Frucht mit einer harten Schale die geknackt werden muss.
das Rad	Ein Teil einer Maschine oder eines Autos, dass man drehen oder rollen kann.
das Obst	Ein anderes Wort für Früchte.
das As	Eine bestimmte Spielkarte die meist den höchsten Wert hat.
das Öl	Eine Flüssigkeit, die man in der Küche zum Braten benutzt.
das Kalb	Der Name für eine junge Kuh.
das Huhn	Gefiedertes Tier, dass Eier legt und auf dem Bauernhof lebt.
die Schnecke	Ein langsames Tier, dass sein Haus mit sich trägt.
die Ente	Ein meist schwimmendes Tier, dass man in Parks findet.
die Eule	Eine Vogelart die am Abend oder in der Nacht auf Jagd geht.
die Fliege	Ein Insekt dass im Haus oft läßtig ist und auch als Fischköder benutzt wird.
die Feder	Etwas das Vögel zum Fliegen hilft. Man kann es auch zum Schreiben gebrauchen.
die Ampel	Ein Licht, dass im Autoverkehr anzeigt wenn man fahren darf.
der Esel	Ein störrischer Verwandter der Pferde mit langen Ohren.
der Fächer	Etwas dass man in der Hand hält und wedelt um sich mit Luft zu kühlen.
der Igel	Ein stacheliges Tier, dass sich im Notfall zusammenrollen kann.
der Affe	Ein Tier dass man im Zoo sehen kann und dass mit uns verwandt ist.
der Delphin	Ein intelligentes Wassertier dem man im Zoo oft Tricks beibringt.
der Besen	Ein Gerät zum Putzen und Fegen.

das Klavier	Ein Musikinstrument mit Tasten.
das U-Boot	Ein Schiff mit dem man komplett tieftauchen kann.
das Armband	Ein Schmuckstück, dass man am Handgelenk trägt.
das Nilpferd	Ein riesiges, schweres Tier dass sich meist in Flüssen aufhält.
das Einhorn	Eine Art Pferd die man in Märchenerzählungen findet.
das Kamel	Ein einhöckriges oder zweihöckriges Tier dass man in Wüsten zum Tragen benutzt.
die Eidechse	Schnelle, langgestreckte Kriechtiere, deren Schwänze leicht abfallen.
die Kokosnuß	Eine an Palmen wachsende tropische Frucht mit harter Schale und weißem Inhalt.
die Antenne	Etwas mit dem man eine Radio oder Fernsehübertragung verbessern kann.
die Giraffe	Ein Tier mit einem langen Hals und einem gefleckten Fell.
die Ananas	Eine tropische Frucht die an Palmen wächst mit einem süß-sauren Saft.
die Schildkröte	Ein vierbeiniges Kriechtier, dass einen Rückenpanzer trägt und seinen Kopf einziehen kann.
der Schnürsenkel	Ein Band mit dem man sich die Schuhe zubindet.
der Elefant	Ein großes, graues Tier (aus Indien oder Afrika) mit einem Rüssel und großen Ohren.
der Astronaut	Teilnehmer an einem Weltraumflug.
der Federball	Etwas dass man beim Spielen mit Schlägern benutzt.
der Hubschrauber	Eine bestimmte Art von Flugzeug mit Drehflügeln.
der Eskimo	Eine Volksgruppe die in kalten Gegenden wohnt und Iglus baut.
das Alphabeth	Die gesamten Schriftzeichen einer Sprache.
das Eichhörnchen	Ein auf Bäumen lebendes Tier mit einem buschigen Schwanz, dass Nüsse sammelt.
das Radieschen	Eine bestimmte Art von Gemüse, von der man die Wurzel ißt.
das Telefon	Ein Gerät mit dem man über weite Entfernungen miteinander sprechen kann.
das Mikroskop	Ein Gerät womit man sich kleinste Dinge vergrößert anschauen kann.
das Orchester	Eine Gruppe von Menschen die zusammen Musik spielen