



National Library
of Canada

Acquisitions and
Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale
du Canada

Direction des acquisitions et
des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file - Votre référence

Our file - Notre référence

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

Canada

**SPEECH PRODUCTION PROCESSING
IN THE SECOND LANGUAGE**

by

MELANIE LUKACH

A thesis submitted to the Department of Linguistics
and the Graduate School of the University of Ottawa
in partial fulfilment of the requirements
for the degree of

MASTER OF ARTS

in

LINGUISTICS



National Library
of Canada

Acquisitions and
Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale
du Canada

Direction des acquisitions et
des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file *Votre référence*

Our file *Notre référence*

The author has granted an irrevocable non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of his/her thesis by any means and in any form or format, making this thesis available to interested persons.

L'auteur a accordé une licence irrévocable et non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de sa thèse de quelque manière et sous quelque forme que ce soit pour mettre des exemplaires de cette thèse à la disposition des personnes intéressées.

The author retains ownership of the copyright in his/her thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without his/her permission.

L'auteur conserve la propriété du droit d'auteur qui protège sa thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

ISBN 0-315-89648-5

Canada



UNIVERSITÉ D'OTTAWA
UNIVERSITY OF OTTAWA

ABSTRACT

The phenomenon of foreign accent has long been a topic of linguistic theory. Non-native-like speech was previously interpreted as being the result of incomplete acquisition. The Critical Period Hypothesis, the Contrastive Analysis Hypothesis, competing cognitive systems and affective factors were all cited as preventing native-like acquisition of the L2, and especially its phonology. Research has shown, however, that even L2 speakers who match native speakers on tests of grammatical competence cannot match native performance in speech or writing.

Neufeld proposes that speech production, especially at the phonological level, is hampered by the use of (conscious or unconscious) knowledge that speakers have about the L2 — metalinguistic knowledge. Those who begin acquiring an L2 after the age of five focus more on structural correctness than younger learners, and tend to use this metalinguistic knowledge more often. Thus even among balanced bilinguals, on an experiment designed to induce focus on form, older learners should perform more speech errors and dysfluencies than native speakers or early bilinguals, and tend to correct more. This pattern should be even more pronounced in learners who have acquired their L2 in a formal (school) context.

An experiment consisting of five tasks was designed to test these three points of Neufeld's Pre- and Post-Articulatory Verification (PAV) model. Five Anglophone Controls and four groups (N=5) of highly-proficient French-English bilingual Ss — simultaneous Bilinguals (acquisition between ages 0-5), older learners who had learned in street (SC) or formal (FC) contexts, and those who learned English after age 10 (10+) — performed tasks designed to elicit speech errors. Task 1 measured reading speed and category evaluation (the sentences were normal,

morphologically malformed, and there were two semantic and syntactic categories — anomalous and bizarre); Task 2 measured oral reading speed; Task 3 required Ss to read aloud the stimulus items, correcting anything wrong; Task 4 used a set of three key words to elicit sentence production, and Task 5, a French version of Task 4, was not performed by the Control group. Ss were requested to respond as quickly and as correctly as possible to all items.

As predicted by Neufeld's model, the Bilinguals did perform more like the Controls than any other group, and the SC and FC groups seemed to be negatively affected by metalinguistic interference (many errors and dysfluencies), the SC group being less severely affected. The FC group performed least like the Controls. Other data supporting these findings are presented. Problems for the model include: the performance of the 10+ group; the discrepancy between the performance of the Bilinguals and the Controls; the Francophone groups did not meet the proposed criteria for balancedness, and response time was not at all significant. Modifications to the methodology are proposed, and more in-depth analysis of the data is also suggested to overcome these difficulties. Overall, Neufeld's PAV model is supported; the hypothesis of age-of-acquisition and metalinguistic interference is weakly supported, and the effect of context of acquisition is more strongly supported.

Acknowledgements

First and foremost, I would like to thank Gerald Neufeld, my thesis supervisor, for his patience and academic and moral support over the past several years. Without his continued belief in my ability to get this thesis written, I would have quit years ago.

My friends and housemates — John and Margaret Jensen, Clara Broten, Alice Lang, Tim Lethbridge, Tim Brennan, Spice, Xi Ming, and Chris Heick — have all provided supportive shoulders to cry on, puffs of encouragement, and distraction when needed. Also, thanks for saying "Get back to work!" Régis Foré, of the Phonetics Laboratory, has also provided support both technical and moral, both of which are warmly appreciated. Jocelyne De Bellefeuille was of great assistance and support in the early stages of the experiment. I am happy to be including my thanks to Nancy Beaudry, my counsellor.

Last, but certainly not least, I would like to express heartfelt thanks to my family — my sister and mother especially, the Clarks (Lynne and Gerald and Shannon and Graham), Aunt Liz and my father. All have provided distraction, without uglification and derision, encouragement, and sometimes warm places to run away to and relax when it all got to be too much for me.

I would also like to thank all my subjects.

Table of Contents

Introduction	1
 Chapter One	
Origins of Foreign Accent	6
1.1 Neurophysiological Explanations of Accent	7
1.2 Contrastive analysis and language transfer.	10
1.3 Competing Cognitive Systems	12
1.4 Social and Psychological Factors	18
1.5 Performance and Foreign Accent	20
1.6 Summary	24
 Chapter Two	
Speech Production Models	27
2.1 L1 Speech Production Models	27
2.1.1 Fromkin's Model	28
2.1.2 Garrett's Model	31
2.1.3 Expansions upon Garrett's Model	34
2.1.4 Neurologically-based models	37
2.2 Models of Second Language Speech Production	40
2.2.1 Performance and Second Language Speech Production	43

Chapter Three

Method	49
3.1 Hypotheses	49
3.1.1 Working:	49
3.1.2 Tested:	50
3.2 Subjects	51
3.2.1 Control Group	51
3.2.2 Experimental Group	51
3.3 General Procedure	52
3.3.1 Task 1	55
3.3.2 Task 2	55
3.3.3 Task 3	56
3.3.4 Task 4	57
3.3.5 Task 5	59
3.5 Measures	59

Chapter Four

Results	63
4.1 Task 1	63
4.2 Task 2	65
4.3 Task 3	68

4.4	Task 4	72
4.5	Task 5	79
 Chapter Five		
Discussion and Conclusions		86
5.1	Discussion	86
5.2	Problems with Experimental Methodology and Design	89
5.3	Conclusions	91
 APPENDIX A		96
 APPENDIX B		101
 APPENDIX C		104
 APPENDIX D		106
 APPENDIX E		110
 References		114

List of Tables

Table 1: Task 1, Response-Time Latency	64
Table 2: Task 1, Category Evaluation	65
Table 3: Task 2 Data	66
Table 4: Task 2, Errors, Corrections, Dysfluencies, Substitutions, Additions and Omissions	67
Table 5: Task 3, Response-Time Latency	69
Table 6: Task 3, Errors, Corrections, Dysfluencies, Substitutions, Additions and Omissions	70
Table 7: Task 4, Response-Time Latency	73
Table 8: Task 4, Errors, Corrections, Dysfluencies, Substitutions, Additions and Omissions	75
Table 9: Task 4, Type-Token Data	78
Table 10: Task 5, Response-Time Latency	80
Table 11: Task 5, Errors, Corrections, Dysfluencies, Substitutions, Additions and Omissions	81
Table 12: Task 5, Type-Token Data	84

Introduction

One of the highest compliments second language (L2) speakers can be paid is to be told by native speakers that because they seem to have no accent, the native speakers were not aware that they were not native speakers. Likewise, when hearing a person speak one's own language with an accent, people sometimes try to identify the person's native language (L1) by listening carefully to their accent. The phenomenon of foreign accent receives much attention.

It is also the focus of much academic research. The Critical Period hypothesis is based upon the idea that when one learns¹ a language after a certain age or 'critical period', one will not be able to learn the language well enough to be able to pass as a native speaker of that language. This is ascribed to maturationally-induced changes in the brain. After these changes have begun, it is proposed that a learner will not have the same kind of representation of knowledge in the brain as one who began to acquire the L2 at an earlier age, and thus one's performance will necessarily be different.

The Contrastive Analysis hypothesis (CA) was also used to account for the phenomenon of foreign accent. Originally falling within the domain of behaviourism, this approach held that when learning an L2, the habits of the L1 will necessarily affect one's acquisition of the L2. While the affects of the L1 are not always adverse, such transfer is believed to account adequately for foreign accent. More recently, researchers using the framework of Universal Grammar have proposed a similar scenario, involving, however, the setting or resetting of natural parameters.

¹ I will be using the terms 'learning' and 'acquisition' interchangeably in this document.

Social and psychological factors have also been used to account for foreign accent. One's attitude toward the speakers of the L2 and their language, one's reasons for learning it, and the existence of family or societal support are said to affect, positively or negatively, one's ability to acquire the language. The more advanced the student is, it is predicted, the more likely that they will feel removed from their L1 and culture — although such an identity problem can be mitigated by retaining a foreign accent when speaking the L2. Yet no matter how these factors affect L2 acquisition, there are few who would use this approach alone to account for foreign accent.

One theoretical approach to foreign accent that combines various aspects of the hypotheses described above is the theory of Competing Cognitive Systems. Within this framework, it is proposed that after a certain age (critical period) one tends to treat learning an L2 as if it were a problem, and one uses general problem solving techniques (e.g. knowledge of the L1 — CA) that are inadequate for the task of language acquisition. While these problem solving techniques cannot entirely supplant natural language processing, they are seen as interfering with it. Some who use this approach also propose that psychological factors affect the use of problem solving, although others do not take this in to account.

It is an anecdotal commonplace, however, to note that one cannot always use the L2 knowledge one has when speaking or writing. This observation was scientifically corroborated when researchers found that even when L2 speakers perform similarly to native speakers on tests of grammatical knowledge (linguistic competence), they are not able to approach native-like levels of performance when either writing or speaking. The distinction made by these researchers between their Subjects' competence and performance is one that is rarely maintained. The

approaches to foreign accent described above have all used performance as evidence of competence, generally without acknowledging that one's performance is not a perfect reflection of one's competence.

While some who have tried to account for foreign accent have used CPH-like constructs, or interference, others have found that even when taking lesser syntactic knowledge or proficiency into account, performance scores were unaccountably lower than comprehension scores. Moreover, experimentation has also shown that older learners can pass for native speakers, and that people can mimic accurately L2 sounds when unburdened by the processing of semantic and syntactic knowledge.

The current models of speech production — developed on the basis of evidence from (usually English) speech error data — rarely do little more than mention that such information may be used. The major models of speech production are linear; that is, they treat speech production as following a sequence of processors. They generally start with some form of semantic or cognitive communicative idea, that undergoes semantic, syntactic, phonological and phonotactic processing. Most models divide syntactic processing into more than one stage, with different functions being performed at each stage. Retrieval of the lexical items is proposed at different times and in different ways in different models, while morphological processing is usually subsumed under syntactic and/or phonological processing. Editing and self-correction, while commonly acknowledged to exist, are not paid much attention. More recent, but less developed, neurologically-based models contain the same processors proposed by the linear models, but they allow parallel and/or simultaneous processing on all levels. They also tend to

place more stress on editing, and have less difficulty allowing both the linguistic processors and other phenomena to affect the speech production process.

Very little work has been done on second language speech production. Many papers that seem to be on this topic deal more with speech production in languages other than English, rather than with second language processing per se. This data can be useful, nevertheless. Most researchers in this area do not assume that there are any radical differences between L1 and L2 speech production. Much of the data examined so far is not highly detailed, and is fairly general, covering hesitation phenomena and corrections rather than errors. While there have been differences noted in the distribution of phenomena, these errors have not required any changes to the usual models of speech production.

This data does not, however, have any impact upon the phenomenon of foreign accent, or the effect of performance factors upon foreign accent. Based on the data presented above, Neufeld has proposed a model of L2 speech production, the Pre- and Post-Articulatory Verification model (PAV), that takes performance factors into account. Like other models, it proposes no major difference in speech production processing. Instead, the older learner's tendency to concentrate on producing grammatically correct utterances, and to use metalinguistic knowledge — the knowledge one has about a language — to do so, interferes with natural linguistic processing. Neufeld proposes that there is a processor that attempts to feed this information, along with other cognitive and perceptual data, into the linguistic process. This interferes with normal pre- and post-articulatory verification, such that segments that are not entirely appropriate are not edited out or corrected, and the phenomenon of foreign accent occurs or is more pronounced. Neufeld also suggests that those who learn the L2 in a formal (school)

setting, where there is a constant focus on 'correct' language, will focus more on form than those who have learned the L2 in an informal (street) setting, and will suffer more metalinguistic interference.

The purpose of the experiment described in the following pages is to examine the effect of metalinguistic processing on foreign accent. If Neufeld's hypotheses are correct, it is hoped that age of acquisition and acquisition context will correspond to notable differences in speech error patterns on the experiment, on which the Subjects were asked to respond quickly, and to focus on the grammatical form of their answers.

Chapter One

Origins of Foreign Accent

The characteristic that most readily distinguishes the speech of a non-native from that of a native speaker of a language is accent. The non-native speaker may also differ in lexical choice or the range of syntactic structures used, but these differences are not very noticeable (Seliger, 1978; Johnson and Newport, 1989). Perhaps it is for this very reason that the phenomenon has not received much attention. The phenomenon of accent, however, has received much attention, and many theories have been advanced to account for it. Some have attributed accent to an inability to articulate foreign sounds (Flege 1981; Major 1987a), while others have claimed that problems in pronouncing a second language (L2) are the result of not being able to distinguish these sounds aurally (Oyama 1976; Flege 1981). Others have proposed that there is an inability to internalize the phonological rules of the L2, for linguistic (Lado 1957; Richards 1971; Dulay and Burt 1974; Eckman 1981; Broselow 1983) and psychological reasons (Jakobovits 1970; Guiora et al. 1972; Gardner and Smythe 1975), or both (Stevick 1976; Krashen 1982, 1985; Major 1987a,b). As foreign accent is most marked in those who acquire the L2 after early childhood (Oyama 1976; Flege 1981; Neufeld 1988; Scovel 1988; Johnson and Newport 1989), most theories have proposed a 'critical period' for language learning of one sort or another. The critical period is that age after which, for most people, it is more or less impossible to learn an L2 well enough to perform in all aspect like a native speaker. Although the concept of the critical period has received much attention, there is little empirical evidence to support this hypothesis (Neufeld 1988; Patkowski 1990) — it has, rather, not been disproved.

The distinction that can be made between the underlying linguistic knowledge of an individual (competence) and how this knowledge is used in performance is rarely maintained by those who work in this area (Flanigan 1991). Researchers who discuss competence in L2 usually use performance on oral or written tests to assess it (Courchene and Bagheera 1985). The errors produced on these tests are usually ascribed directly to lacunae in competence, without considering that the rule(s) may merely be difficult to access or implement during speech. Difficult-to-use rules are, it is claimed (Krashen 1979, 1982; Neufeld 1988; Green and Hecht 1992), likely to be late-learned and/or metalinguistic in nature. As L2 acquisition after age five is considered, by some, to be dependent upon such rules (Krashen 1979, 1982), it is possible that non-native-like speech may be due to developmentally-induced changes in language processing of a linguistic rather than cognitive nature (Goodluck and Birch 1988).

1.1 Neurophysiological Explanations of Accent

One theory that deals with imperfect L2 acquisition is the Critical Period Hypothesis (CPH). Introduced by Penfield and Roberts (1959), based upon data from studies of epilepsy, aphasia and cortical mapping, they proposed that the ability to recover speech after brain damage (plasticity — different areas of the brain taking on functions of damaged areas of the brain) or to acquire an L2 after puberty was impaired due to maturation of the brain. They found that children of up to 11 or 12 years seemed to recover the ability to speak without impairment after brain damage. Older aphasics usually sustained lasting speech impairment. This, combined with their observation that most people who learn an L2 after puberty do not attain a native-like accent, Penfield and Roberts attributed to loss of plasticity of the brain.

Although they claimed (p. 240) that plasticity began to decrease after age 5, by which time the L1 has been for the most part acquired (p. 243), it was only after maturation was complete that language could not be successfully learned or relearned (p. 240). Penfield and Roberts proposed that younger children acquired language 'directly', as an outgrowth of thought. Between 6 and 10 years, although L2 learning is usually less direct, due to teaching methods, a native-like accent could eventually be attained (p. 240-241). After 11 or 12 years of age, when cerebral maturation is largely complete, the L2 can only be learned 'indirectly' via the L1, using a translation-like process. At this point, it is impossible for either L1 or L2 to be acquired 'directly', and thus those (re-)learning a language after this age usually never achieve L1 production norms (p. 251-252).

Lenneberg (1967) expanded upon Penfield and Roberts' ideas. Not only did he include updated neuropathological data, but he also used information from language development in other cultures and among the deaf and retarded, as well as from children in 'abnormal' home situations. He also explored the issues of cognitive capacity, whether or not practice was needed for language acquisition, and whether children were or were not 'taught' their L1. Lenneberg agreed with Penfield that native-like language acquisition, first or second, seemed limited to a critical period that ended about puberty. Unlike Penfield and Roberts, however, who never defined in detail what constituted maturation of the brain, Lenneberg associated the end of the critical period with completion of lateralization (p. 178). Lateralization, in turn, was defined as a complex phenomenon encompassing lateralization (localization of functioning in one hemisphere of the brain, usually the left) of language and handedness (also referred to as cerebral dominance), changes in brain chemistry, neural morphology, and EEG patterns. The most

important diagnostic criterion was, however, loss of plasticity. Lenneberg's examples of the effects of the loss of plasticity included the inability to continue acquiring L1 in the case of abnormal development, as well as the inability to recover from aphasia. While Lenneberg did acknowledge that these factors could be coincidental rather than causative (p. 169), he seemed to attribute the strong, almost predictive correlation to a single underlying cause.

Over the next few years, some researchers' data supported Lenneberg's conclusions (Asher and Garcia 1969; Scovel 1969; Oyama 1976; Seliger 1978). Some proposed modified versions of the CPH. This was due, in part, to subsequent researchers using the term 'lateralization' to refer only to neurophysiological lateralization, or cerebral dominance. Asher and Garcia (1969), among others, noted that some older learners could learn an L2 quite well. Scovel (1969) suggested that the Critical Period affected only the acquisition of the phonological system of a language, as almost all the evidence for the Critical Period was articulatory in nature. Yet Seliger (1978) claimed that although evidence for the Critical Period was largely articulatory, the possibility that syntax and semantics might also have critical periods, occurring later than that of phonology, was not precluded.

There was also opposition to Lenneberg's book. One of the more important of these papers is Krashen (1973). In this article, Krashen noted that there was evidence for lateralization of language before the age of five as early as 1963 (Kimura 1963). This finding was supported by later studies (Knox and Kimura 1970; Geffner and Hochberg 1971). Krashen concluded that the Critical Period, then, could not be the result of lateralization. Although Krashen concedes that interhemispheric transfer may occur up to puberty, he claims that evidence of aphasia from

right-hemisphere damage (Kinsbourne 1971; Best 1987) renders doubtful the mechanisms for and existence of transfer.

Today it is generally accepted that the hemispheres are neurophysiologically asymmetric at birth (Best 1987; Springer and Deutsch 1988). Since the brain undergoes considerable development after birth, however, when linguistic (and other) stimuli could affect development, the question of to what extent innate asymmetry may influence the acquisition of language is still open to debate (Best 1987).

The CPH currently as described above receives little support today (Patkowski 1990). This is due more to the nature of current theories of language and the questions they raise than due to advances in neurophysiology. The CPH, itself, does not presuppose a specific processing system, cognitive or UG-like, except that it be innate. Current research centres upon how, if at all, the underlying organizational system (competence) changes, and when this change might occur (Patkowski 1990). The CPH, then, seems to be neither a competence nor a performance model of L2 acquisition, but an attempt to describe how the L2 is internalized and stored. In the event that this process changes, to what extent already-acquired knowledge, as well as the process of acquiring new knowledge, also changes, is an open question.

1.2 Contrastive analysis and language transfer.

The thesis of the Contrastive Analysis (CA) hypothesis and related approaches is that L2 acquisition is affected and/or mediated by L1 knowledge. CA is a cognitive approach, in that using current knowledge to aid in acquiring new knowledge is a general principle of problem

solving. It is distinct from what I will call the 'cognitivist' position in that imperfect L2 acquisition is usually accounted for by CA alone.

Originally, an offshoot of behaviourism, the CA position was that the linguistic 'habits' of the L1 (phonology, syntax, etc.) were transferred to the L2 situation (Lado, 1957; Wardhaugh 1970). Where the grammars of the two languages had identical rules it was assumed that the rules would be transferred automatically. As this would aid acquisition, it was called 'positive transfer'. Where the grammars did not match, the L1 rule was believed to interfere with learning the new L2 rule. This situation was known as 'negative transfer'. Lado claimed that by comparing the grammars of different languages and finding where they differed, one could predict where positive and negative transfer would occur (Lado 1957). To perform such a comparison, even on only a part of the language, took much time and required much painstaking research into the grammar of both languages concerned. Instead, many researchers examined the performance of L2 learners to see if the errors they made could be ascribed to negative transfer. This much-abbreviated form of CA became known as Error Analysis (James 1971; Richards 1971).

One of the results of error analyses, however, was that it was found that language structures did not always transfer, either positively or negatively (James 1971; Dulay and Burt 1974). In some areas, it was found, L2 learners reacted more like L1 acquirers (Richards 1971; Bailey, Madden and Krashen 1974; Dulay and Burt 1974). The similarities between L1 and L2 learners started a wave of research into universals of language acquisition (Dulay and Burt 1974; Eckman 1977, 1981). At the same time it divided the field into those who proposed a more general, 'cognitive' system for language acquisition after the L1 was acquired (section 1.3), and

those who proposed that an innate, linguistic system of restrictive parameters, like that proposed by the proponents of Universal Grammar (UG), is still used but is somehow affected by the L1 structure.

Over the last decade, much work within the area of second language theory has been done within the UG framework. In the area of syntax, work has focused on supporting the existence of parameters and principles and how they are set (Flynn 1984; White 1985, 1986; Bley-Vroman, Felix and Ioup 1988; Zobl 1988; duPlessis et al. 1989; Tomaselli and Schwartz 1990) or reset (Flynn 1987, 1989; White 1988). Although generative phonologists have used speech samples as their data base (Broselow 1983, 1984, 1988), this research was intended to give an account of the syllable structure of the L1 as much as it was intended to show how this competence was transferred to the L2. This is entirely in keeping with the Chomskyan approach to linguistics, which is seen as the description of competence (see Gregg 1989 for a thorough review of this topic). Thus performance is rarely mentioned by generative linguists, except to define the competence-performance distinction used to explain their approach to acquisition data (Gregg 1989; White 1990; Lennon 1990, 1991).

1.3 Competing Cognitive Systems

As mentioned in the previous section, cognitive explanations for older learners' imperfect L2 acquisition are, historically, related to CA. Unlike purely psychological explanations for language acquisition, such as those of Piaget and Slobin, these approaches generally propose that there is some form of innate language acquisition device for L1 acquisition. Older learners, however, are presumed to use general problem-solving techniques when acquiring L2, an

approach suggested by Rosansky as early as 1975. Within this framework lie Selinker's Interlanguage hypothesis (1972), Krashen's Monitor Model (1977, 1979), Felix's Competing Cognitive Systems Model (1985), and Bley-Vroman's Fundamental Difference hypothesis (1989), among others.

Selinker's Interlanguage hypothesis (1972) was one of the earliest 'cognitive' (see previous section) accounts of non-native L2 performance. While accepting Lenneberg's (1967) 'latent language structure' (LLS), Selinker also claimed that there was a 'latent psychological structure' (LPS) that was responsible for foreign accent and structural errors in L2 speech (p. 211). Although he did not state that it was impossible to fully 'reactivate' the LLS, Selinker claimed that only 5% of the population could do so and thus achieve native-like performance (p. 212). Imperfect reactivation of the LLS resulted in the use of the LPS during language acquisition. The non-native-like speech that resulted from the use of the LPS was called interlanguage (IL) (p. 214).

Although Selinker posited that there were five central processes within the LPS — namely language transfer, transfer of training, strategies of second language learning, strategies of second language communication, and overgeneralization of L2 rules — the hallmark of interlanguage was fossilization. Fossilization was defined as the surfacing of an incorrect rule after the correct version had been acquired (p. 215), which is known as 'backsliding'.

The Interlanguage hypothesis seems, on the surface, to be entirely performance-oriented. The only permissible evidence of interlanguage or fossilization is speech language data, and the LPS was likewise held to underlie L2 speech. Statements to the effect that the proposed universal, interlingual structural units become

available to the idealized second language learner who will not achieve native-speaker competence in the TL [L2], ... whenever he attempts to produce a TL norm (p. 228)

seem to indicate that the sole criterion for competence is performance. Yet the very concept of fossilization contributes to the confounding of competence and performance. A fossil is merely the relic of an old IL rule, it is not proof that one's competence is currently lacking (Lennon 1991) — yet fossilization is also considered proof that one is still speaking IL and has not attained the L2 rule. Subsequent work that seems to be within this framework, due to the use of the terms interlanguage and fossilization and backsliding have, for the most part, focused entirely upon the rules of IL competence. The cognitive Interlanguage hypothesis has largely fallen by the wayside.

The basis of Krashen's model lies in the distinction he makes between 'learning' and 'acquiring' a language. Acquisition, in this sense, refers to the natural, subconscious way a child acquires its mother tongue. Learning, then, is the conscious, rule-based approach to language that results in knowing about language, rather than knowing the language itself. According to Krashen (1977, 1985), this learned knowledge functions solely as a Monitor. The role of the Monitor is to ensure that output conforms to the rules within the Monitor.

Krashen proposed that the only time that learned material alone may be used for speech production is when the learner is forced to speak while still in the early stages of L2 acquisition, for example in a classroom setting. In such a situation, one can use the Monitor to shape L1-based utterances into a more L2-like form. By this means learners can 'outperform their competence', or seem to know more of the language than they actually do. Speech production, which is a normally an unconscious process, is considered by Krashen (1977) to be the product

of acquired knowledge. Optimally, in normal conversation the Monitor is not used much. The Monitor is used when the speaker is concentrating upon producing grammatical speech. As it is generally assumed that the rules within the Monitor are grammatical (Flanigan 1991; Green and Hecht 1992), use of the Monitor should result in more grammatical speech, but, due to the amount of time it takes to use the Monitor, at the cost of fluency (Krashen 1978, 1985).

More recently, Krashen (1985) has added an Output filter to his model. Based on information from Stevick's (1976) book, the proposal is that affective states, such as anxiety about speaking the L2, may result in the production of speech that is both less fluent and less grammatically accurate. In Krashen's version of this hypothesis, it is assumed that the knowledge has been acquired, but that speakers have not yet or are not able to "perform their competence" (Krashen 1985, pp. 45 and 64-65). Although the Output filter is used when speakers are concerned with the adequacy of their speech, it is distinct from the Monitor in that consciously-known rules are not accessed, and grammaticality as well as fluency is impaired. The relations between the Output filter and the Monitor — does it develop into or form the basis of the Monitor? — and the Affective filter (which inhibits the acquisition of knowledge from comprehensible input), are less clear.

Felix's position is closer to Seliger's than to Krashen's, although less fully developed than either. According to Felix (1985), child learners use a UG-like system, the Language-Specific Cognitive (LSC) system, to acquire language. It is the only system able to deal properly with the abstractness and complexity of language. Around the ages of 10-12, Piagetian Stage IV, the stage of Formal Operations, is attained (Rosansky 1975). At this point, children have a fully developed Problem-Solving Cognitive (PSC) system. Felix states that although "adults also use

the LSC-system where they proceed in the same way as children" (1985, p. 59), usually the older learner "is not able to consciously or even unconsciously suppress the operation of PSC-structures, if exposed to new linguistic data" (p. 57). As the PSC-system is inadequate for dealing with language, the older learner is thus incapable of achieving native-like knowledge of the L2. Felix claims that the structure of adult L2 speech "shows quite clearly that postpubertarian L2 learners partially use cognitive structures in the acquisition process that are essentially different from those used by children" (p. 64).

There is, however, evidence that not only accent but structural differences are detectible in the speech of those who began acquiring an L2 as early as age 7 (Swain and Lapkin 1986; Oyama 1976; Johnson and Newport 1989). However, if the PSC-system begins to develop at age seven, this evidence could be considered as support for Felix's hypothesis. Another area that remains undeveloped, in Krashen's as well as in Felix's theory, is how learning environment and attitudinal factors affect L2 acquisition (Gregg 1989, 1984).

The current trend within this framework is to base one's rejection of UG as a source of knowledge for older learners by using the principles and parameters approach itself (Clahsen 1984, 1990; Clahsen and Muysken 1986, 1989; Jordens 1988a, 1988b; Bley-Vroman 1989; Schachter 1989, 1988, and many others). There are distinguishing features within this group, such as whether one uses speech data — as do Clahsen, Muysken, and Jordens — or grammaticality judgement data (Schachter, Bley-Vroman) as the basis of one's arguments. Whatever the data source, the general conclusion is that UG is not available for L2 acquisition in the same way it is for L1 acquisition. Another difference is the degree to which UG is considered unavailable for L2 acquisition. Bley-Vroman (1989) and Jordens (1988a) tend to reject the availability of

UG altogether, while Schachter (1989, 1988), Clahsen (1990), and Muysken (Clahsen and Muysken 1986, 1989) have claimed that access is merely incomplete or unreliable. Unfortunately, with the exception of Bley-Vroman, most researchers in this area have concentrated upon disproving the accessibility of UG for L2 acquisition rather than develop their own account of how L2 acquisition does take place.

Bley-Vroman (1989) has called his position the Fundamental Difference hypothesis, as he believes adult L2 acquisition to be fundamentally different from L1 acquisition. As a domain-specific cognitive system such as UG or depth perception should not be able to be supplanted by a less domain-specific cognitive system (p. 61-62), the only conclusion remaining is that UG is not available for later language learning. Other evidence, such as fossilization, lack of native-like competence even if performance is native-like (p. 44), variability, etc., are also used as support for Bley-Vroman's position. In order to account for UG-like effects in L2 acquisition — which he believes to be theoretically more parsimonious than assuming that UG is fully operational — Bley-Vroman proposes that knowledge of features of language (e.g. structural units and discourse features) that are universal are consciously or unconsciously used to acquire the L2.

The adult foreign language learner constructs, therefore, a kind of surrogate for Universal Grammar from knowledge of the native language (Bley-Vroman 1989, p. 52).

One problem for this model, aside from its lack of detail, is that the parameters and principles model is itself incomplete, as Bley-Vroman (1989) and others (Gregg 1989; White 1990) have acknowledged. Not only do researchers disagree on the existing parameters (duPlessis et al. 1987; Tomaselli and Schwartz 1990), but the status of the principles (e.g. the

Subset Principle, Zobl 1988; Hirakawa 1990) and the role of Markedness (Liceras 1988; Sharwood-Smith 1988) are also undetermined and of major importance. Moreover, like other approaches to L2 acquisition, the Fundamental Difference hypothesis is also focused upon explaining the acquisition of L2 competence. Although Bley-Vroman discusses variability in performance, this is once again ascribed directly to lack of competence (p. 45).

1.4 Social and Psychological Factors

Not all researchers believe that neurophysiological or cognitive maturation are the main source of imperfect L2 acquisition or performance. Some have proposed that this is caused, rather, by affective factors and differences in how and how much one is motivated to learn the L2. The cognitivists tend to consider affective factors as contributing to poor L2 acquisition/performance, and the fact that affective factors need to be taken into account for L2 acquisition they see as support for their position (Krashen 1973, 1978, 1981, 1982, 1985; Bley-Vroman 1989). Those who work within this framework, however, tend to assume that L2 acquisition is for the most part the same as L1 acquisition.

This viewpoint was assumed by Gardner and Lambert in their seminal book, *Attitudes and Motivation in Second Language Learning* (1972). Their aim was "to construct the beginnings of a social psychological theory of second- or foreign-language learning" (p. 3). The specific factors they were interested in were instrumental vs. integrative motivation, familial attitudes towards speakers of the L2, and how these correlated with intelligence, linguistic aptitude, and academic achievement. Different social settings for L2 acquisition were studied; majority-language speakers learning a minority language (anglophone Americans learning French);

minority-language speakers learning the majority language (Franco-Americans learning English), and people learning a prestigious foreign language (Philipinos learning English).

The instrumental motivation was defined as learning a language in order to be able to make practical use of it, e.g. in the workplace. The integrative motivation was defined as learning the L2 in order to know more about its speakers and their culture, and to enable the learner to communicate with them. Gardner and Lambert studied not only the learners' attitudes towards and stereotypes of the L2 speakers and their culture, but also one's identification with one's own culture, ethnocentrism, and authoritarianism (p. 16-17). This approach was based in part upon the authors' belief that the L2 necessarily embodies the cultural values of the L2 speakers; thus L2 acquisition necessarily entails some degree of enculturation (p. 14). Ethnocentric learners, then, would be less likely to have an integrative motivation for studying the L2. A course which results in advanced knowledge and high proficiency in the L2, for such learners, entailing a high degree of enculturation, would cause the learners to draw back from the L2 experience (p. 6).

Gardner and Lambert did find correlation between L2 achievement, affective factors, and motivation independent of intelligence and linguistic aptitude — although to what extent remains unclear. For example, for foreign language learners who are highly motivated, have a good attitude and family support, L2 learning for instrumental reasons is as successful for typical North American learners who have an integrative motivation. The American anglophones, moreover, had some negative attitudes towards the French whose language they were studying. There were also smaller, more individual trends within each grouping.

Subsequent research has rarely been of such large scale. Guiora, Brannon and Dull (1972) found that high scores on a rating of 'empathy' corresponded with a more native-like accent. Brown (1981) had similar findings when examining 'permeability of ego boundaries'. Both of these factors, however, could be considered as part of or equivalent to Gardner and Lambert's integrative motivation and the consequences of a lack of ethnocentrism. Other researchers also studied different versions of what were largely the same factors as Gardner and Lambert (Gardner and Smyth 1975; Jakobovits 1970). Lack of consensus on the factors involved has, no doubt, contributed to the stasis within this field. Another contributing factor to the decline of these studies has been the change from large-scale studies such as Gardner and Lambert's to smaller ones such as Guiora et al.'s. The large-scale studies were intended to provide information upon which to base L2 teaching methodology. The smaller studies, concentrating on individual variation in affect and motivation, are not so applicable. As a result, interest in such studies has waned within the area of L2 acquisition research.

1.5 Performance and Foreign Accent

The theories and hypotheses described in the preceding pages have all attempted to ascribe foreign accent to problems or lacunae in L2 competence. However, one of the bases of recent linguistic theory has been the distinction between competence and performance. That most current linguistic theories and hypotheses are theories of competence does not deny the need for a theory of performance. Linguistic performance is not generally assumed to be a perfect reflection of linguistic competence; indeed, the discipline of sociolinguistics is based upon the idea that, even when people are speaking their mother tongue, performance varies. Yet it is only

recently that the phenomenon of foreign accent has been approached as an aspect of performance.

One of the first to hold this point of view, Bever (1981) proposed that there is a 'psychogrammar', a neuropsychological mechanism dedicated to matching performance to what has been perceived and internalized. This system remains intact until age 6, Bever claims, when, unless it is kept intact by being used to acquire an L2, it deteriorates. For most people, once the psychogrammar has deteriorated, the system cannot be wholly reactivated. As deterioration is seen as being a process, differences in level of accent can be attributed either to the level of deterioration of the psychogrammar when L2 learning begins, or to variability in the level of reactivation possible. Speech perception and ability to internalize the L2 remains intact, then, although the learners may never be able to produce phonemes that are within their competence.

In one of two experiments, Neufeld (1979) found that adults who had learned French as adults and were fluently bilingual obtained test scores comparable to those of Francophone controls when their speech production was assessed. The text that the Subjects used had been specially constructed for the experiment, and contained words and phonemes that Anglophones typically have trouble pronouncing well. In the other experiment, which was also presented in two earlier articles (1977, 1978), 20 adult Subjects were first given auditory discrimination exercises of, and later given vocal imitation exercises in Japanese, Chinese, and Inuktitut. After the instructional phase, each Subject recorded 10 utterances, which were judged by native speakers (Inuktitut was dropped due to a lack of native-speaking judges). Eleven of the Subjects were judged to be native speakers of Japanese; 9 were for Chinese. The utterances were not merely simple words; some of them were as long as 16 syllables, a length which surpasses general memory constraints (generally held to be 7 items, plus or minus two).

On the basis of the evidence from these experiments, Neufeld concluded that adults can produce native-like speech, independent of semantic and syntactic competence (the length of the phrases uttered would require phonological processing). In another experiment (1980, 1988), Neufeld found that highly fluent, bilingual Subjects received native-like results on several tests of perception, but most were still identified as non-native speakers on a speech production task.

Swain and Lapkin (1986) compiled the results from several experiments comparing the French proficiency of early and late immersion students with Francophone controls. The skills tested were listening comprehension, reading comprehension, writing, and speaking, and both formal and informal registers were covered. While all the Anglophones were weaker in grammatical knowledge than the Francophones — only reaching the 30th percentile on the grammar test (p. 4) — the immersion students were able to match them in the social appropriateness of the register used for both the writing and speaking tasks. On measures of listening comprehension and speaking, the early immersion students performed better than the late immersion students. More interesting, though, is the authors' assertion that "based on available comparisons with francophones, it appears that immersion students' receptive skills [competence] are stronger than their productive skills" (p. 6), and that "... speaking is the weakest of the four skill areas for immersion students" (ibid.). Native-like competence, then, does not guarantee native-like speech production.

Flege (1981, 1987, 1988) sees foreign accent as the result of interference, but at the phonological level. The new phonological representations can be learned, even by adults, but both the internalization as well as the production process can be hampered by similar sounds occurring in the L1 (1987, 1988). He has found (1981, 1988) that L2 sounds that are completely

unlike any found in the L1 are produced more correctly than those that are more similar to L1 sounds. Those who had more exposure to the L2 were better at discriminating the correct sounds, even if they could not produce them well. During speech production, the actual sounds produced by learners fall phonetically between L1 and L2 variants of the phone (1981, p. 449-451). Flege (1987, 1988) proposes that this means either that the learner has constructed a performance routine that is intermediate to the underlying L1 and L2 representations, or that the attempt to perform the underlying L2 phone is interfered with by the underlying L1 phone. The superior performance of young children (he gives a cut-off age of 3 years, 1988, p. 448) he likewise attributes to two possible reasons; either children do not process the second language through their L1 phonology — they learn it 'directly' — or, because young children are still processing their L1, they are more readily able to create entirely new phonological representations. Flege also proposes that the amount and quality of exposure to the L2 may affect both internalization and performance ability, as may also motivation, affective, and social factors (1987).

Two others who see foreign accent as the result of interference at the phonological level are Eckman (1977, 1981) and Tarone (1978, 1980). Eckman's Markedness Differential Hypothesis holds that L2 learners will have greater difficulty learning phones (he does not distinguish between internalization and performance) that are more marked than they will those that are unmarked. He does expect there to be some influence from the L1, such that there is minimal interference from the L1 when the L2 sound is unmarked, while there should be considerable difficulty in learning a marked phone when the L1 variety is less marked. It is difficult to evaluate this position, as the relative markedness of phones is not well known.

Moreover languages can have different phonetic realizations of the same underlying phoneme phone (e.g., in English, word-initial voiceless stops — p, t, and k — are usually aspirated, while in French they are not).

Noting that some L2 learners of English inserted epenthetic vowels into consonant clusters, and that speakers of different L1's broke the English consonant clusters in different places, Tarone (1978) first proposed that the learners were using the L1 syllabification template. This was later refined into the Open Syllable Structure hypothesis (Tarone 1980), which states that L2 learners revert to the universally unmarked CV syllable structure, although where they will break up consonant clusters may be transferred from the L1. This hypothesis received some support from an experiment described in the article. There is contradictory evidence (Sato 1984), however, and there seem to be no further studies in this area.

There is much room for further study in this relatively new theoretical domain. This is especially true of the more purely performance approaches (Neufeld, Flege), as opposed to those that intersect syntactic theories (universals, Eckman and Tarone) or neurophysiology (Bever). It is hoped that the experiment described in this study may add to this area of research.

1.6 Summary

Of these five methods of accounting for non-native-like L2 acquisition and/or performance, I have tried to show that three — the neurophysiological, CA, and the cognitive position — deal with the nature and acquisition of L2 competence, rather than performance. The fourth, which emphasizes the role of affective factors, does not usually involve distinctions between competence and performance at all — although it could do so. Those who use

competence theories have criticized the affective factor approach, claiming it is impossible to account for how affective factors could interfere with the acquisition of specific rules (Gregg 1984, 1989; Bley-Vroman 1989). The same could be said from a performance perspective, namely, how can affective factors prevent use of existing competence? The fifth approach, the performance theory approach, is relatively new. I have attempted to introduce performance theory as a necessary complement to a theory of competence, and have presented several of the approaches taken within this area. Being a fairly new domain, there is relatively little evidence either for or against it. Another problem with this domain is that few researchers take variability of performance into account.

One of the more recent trends in L2 research is the combining of approaches to account for the existence of foreign accent. Major (1987a), for example, combines transfer with affective factors, and Hecht and Mulford (1982) and Goodluck and Birch (1988) propose that both transfer and universal developmental factors can account for non-native-like performance. A purely performance-based account for foreign accent need not invalidate a competence-based hypothesis, as a performance theory is a necessary complement to a theory of competence. This may be one reason why competence-based accounts for foreign accent tend not to mention performance as a variable — they may take the idea for granted. Yet by not consciously maintaining this distinction, much effort may be expended trying to account, within a competence-based theory, for variability in speech production that should properly belong to a performance theory.

A theory of L2 speech performance or production could actually support a theory of L2 competence. However, before, one can properly evaluate a theory of speech performance in the

L2, a review of models and theories of speech performance in an L1 is necessary. Thus the next chapter will consist of a review of models of speech production processing, for both L1 and L2 speech.

Chapter Two

Speech Production Models

2.1 L1 Speech Production Models

As Fromkin (1973) points out in the introduction to her book, speech error data has not always been used for linguistic purposes only. Yet since the 1960's, errors have often been used to investigate the structural elements of speech and how they are assembled to produce speech. Although some researchers make use of aphasic speech as data (e.g., Garrett 1984; Butterworth 1989), it is unclear to what extent aphasic speech can be considered analogous to normal speech. The same criticism could be directed at speech error data as to aphasic speech data — that breakdowns in the speech production process might not give an accurate picture of the correctly functioning processor.

Nevertheless, when non-aphasic speech is the source of error data, one is at least assured that the subjects' linguistic competence and performance abilities are intact. Thus if an error occurring to a distinct processing unit passes into speech, it is generally considered to support the reality of such a processing unit as part of the speech processor (Fromkin 1971; Laubstein 1988; Lapointe and Dell 1989). Several of the more widely referenced models of speech production which I review in this chapter are based entirely upon errors from native English speakers — most of whom are likely unilingual, but this variable is not usually controlled or discussed. The earlier models generally depict speech production as the output of several serially-ordered linguistic modules, corresponding to the division of linguistic processes into semantics, syntax, phonology, and sometimes morphology. Some recent models using a more

neurologically-based spreading-activation structure (Stemberger 1985; Dell 1986, 1989) are easily able to include editing and self-correction mechanisms, which were previously accounted for by means of feedback loops (Boomer and Laver 1968; MacKay 1970; Motley et al. 1983).

Most scholars who work in the area of second language speech production would seem to agree that speech production is basically the same in L1 and L2 (Wiese 1984, pg. 21; Crookes 1991). What work has been done on errors in languages other than English (see Berg 1987) has usually been compared to English-based models. This is a necessary precursor to a description of models of L2 speech production.

2.1.1 Fromkin's Model

Fromkin (1973) was the first collection of articles on speech errors and their significance for linguistic theory. Through it the model presented in her 1971 paper became highly influential. Earlier models (e.g., Fry's and Laver's) were among the papers collected in the book and were referred to by Fromkin in her paper. The model presented in Fromkin (1971) — a more complete version of the model described in an earlier paper (1968) — has each stage of processing accounted for and supported by examples of speech errors. The stages are as follows:

1. At Stage One, the meaning of a sentence is generated. This would seem to consist of choosing referents and semantic features (1971, p. 239; 1973, p. 37). Errors in semantic features, Fromkin claimed, may later result in the wrong words being chosen from the lexicon, e.g., 'hate' for 'love' (1971, p. 239). The referents or substantives may themselves be bundles of semantic features. (1971, pp. 235-236)

2. The syntactic form is generated at Stage Two, by associating syntactic features to the previous semantic elements. The syntactic functions and structural relations of the semantic elements are specified here, with misattachments causing errors such as 'Rosa always date shranks' and 'I disregard this as precise' (1973, pp. 32 and 33, respectively). In the first example, the past tense marker was affixed to the incorrect lexical item, coincidentally also a verb in its underived form. In the second example, the negative indicator is attached to the verb rather than to the verb phrase. As the words have not been retrieved from the lexicon yet, the syntactic features are appended to the semantic feature groupings that fill the lexical 'slots'. Words can also be marked for emphasis at this stage, which will later affect the intonation contour of the sentence, as in 'he threw a báll', as opposed to 'he threw a bat', 1971, p. 239).
3. The next stage is when the intonation contour is generated, for phrases and for the whole sentence. These stresses are independent of word stress, but are attached to the appropriate word slot. If, later, the wrong word is put into one of these slots, the word will still receive the correct level of phrasal and sentential stress, even if the word itself is misshapen.
4. At Stage Four the lexical items are retrieved. Only stems or content words are accessed, via their semantic features. The words are retrieved syllable by syllable, which can lead to errors within syllable as well as whole syllable errors. The wrong word may also be accessed at this point, due

either to semantic or phonological similarity. An error in reading the semantic feature list could cause a 'love-hate' substitution, while the substitution of 'present' for 'pressure' could be due to retrieving a highly similar set of syllables.

5. The last stage is when all morphophonological processing is done. The error 'Rosa always date shranks' (see Stage 2) could possibly arise at this stage also, if the past tense marker is misinterpreted as attaching to 'shrink', before it is zero derived from verb to noun and pluralized later during this stage.

The output of the last stage is a string of phonetic segments, which go through "automatic phonetic and phonological rules, converting the sequences of segments into actual neuro-motor commands" (1971, p. 241). Whether further errors could be committed during this conversion process is a topic Fromkin did not address.

While Fromkin's model does have semantic and syntactic stages, it is most detailed when describing the phonological stages, and much time was spent establishing the phonological elements (e.g. the phoneme, the syllable, and the distinctive feature). Fromkin was, however, well aware of the limitations of the model, noting that "[p]arallel processing is continual which is not revealed in the diagram" of the model (1973, p. 43). Moreover, although Fromkin mentions monitoring or editing (1971, 1973) and there is brief discussion of feedback loops as the editing mechanism (1973, p. 43), this matter was never examined to any considerable extent. The model, once formed, was not substantially altered.

2.1.2 Garrett's Model

Garrett's model, originally similar to that of Fromkin, has evolved into something quite distinct (1975, 1976, 1980a, b, and 1984; Garrett and Kean 1980). Garrett's focus has been to examine the levels of syntactic processing, as denoted by speech errors that are syntactically or phonetically accommodated to the linguistic environment. The stages are:

1. **Message Source/Level:** In the earlier versions of this model, other than to state that the source of the communicative intention is not necessarily linguistic in nature, this level was not described in much detail. It also inputs directly into the semantic processor which would construct the functional representation of the utterance. In Garrett (1984), the Message Level is presented as the result of inferential or general conceptual processes, some of which may be linguistic in nature. Although not entirely linguistic, it nonetheless controls the next three stages which comprise the Sentence level.
2. **Functional Level:** The output of this level, created largely by means of semantic features, was said to resemble deep structure (1975, 1976) in that the basic lexical items had been chosen, the content words, and their phrasal relations indicated. It was later (1980b) stated that this was a predicate-argument type of representation, and that the content words were assigned roles and arguments here. The errors that could arise during this stage were substitutions and blends, including those that result from competing plans (1980b); word and phrase exchanges also resulted from

errors at this level. Errors at this level usually occurred to whole words, which might be accommodated by bound morphemes or agreement which are processed at level 3 and 4. Later (1984), the Functional Level was described as the outcome of logical and syntactic processing. Although the output did not differ significantly, lexical items are not seen as being accessed via the semantic component, but via the message component directly. Simultaneously, functional structures are accessed, and combine with the lexical items to form the functional representation.

3. **Positional Level:** The output of this level was described as being analogous to the surface structure (1975, 1976). Here the syntactic component of the grammar was accessed, and all grammatical functors and morphemes indicated and marked, and their slots created. When grammatical morphemes or functors are misplaced or 'shifted', they may accommodate to their new environment, although this occurs less than for errors that arise at the preceding level. These errors often involve adjacent lexical items. Phrasal and emphatic stress are also marked at this level, independently of the lexical item that will later fill the slot. In Garrett (1984), during the syntactic and phonological processing, the phonological form of the content words are retrieved, and phonologically-based substitutions, sound exchanges and shifts occur here.
4. **Instructions to articulators:** The remaining grammatical formatives are retrieved, and all automatic phonological processing is done, resulting in

a complete, phonetically spelled-out utterance. Earlier errors are accommodated here, and haplogogies also occur here. This stage, named the Phonetic Level in Garrett (1984), described as where regular phonological processing is performed, was not materially changed over the years.

5. **Articulatory System:** Originally (1975, 1976) not described, beyond stating that tongue twisters affected this level of processing, in (1980b) Garrett did postulate that a form of monitoring could occur at this level, based on permissible sound combinations. Later still (1984), the motor coding/control processing at the Articulatory Level is described as translating the spelled-out utterance into motor commands.

Although Garrett's model is fairly detailed as to the nature of the syntactic processes involved at the functional, positional, and sound level, the semantic component has been entirely removed. How semantic 'formatives' might affect the Message and Functional Levels is not discussed in Garrett (1984). In one of Garrett's earlier papers (1976), the functional representation was described as being analogous to a case-grammar structure. It is possible that the semantic component is being downplayed due to Garrett's adopting a more Chomskyan, syntactically projected model (1984). Nonetheless, the model has been criticized by Lapointe and Dell (1989) as being neutral with respect to the type of computations performed.

More criticism, and expansions of Garrett's model, have centred on his not assigning the syllable a role as a unit of production (Crompton 1981; Shattuck-Hufnagel 1983; Laubstein 1988). This is also one area where Garrett differs significantly from Fromkin. Garrett also

discusses whether or not the processing levels are strictly isomorphic (1980b, 1984; Garrett and Kean 1980) — i.e., that information from one processing unit (e.g. semantic features) does not affect processing on another level (e.g. the syntactic level). This is despite the fact that other authors (Fromkin 1971; Shattuck-Hufnagel 1979, 1983) noted that substitutions and blends were often of the same grammatical category, were similar phonologically, and had the same number of syllables. Yet Garrett concludes that when items are similar in meaning, they are not structurally similar (syntactically or phonologically), and vice versa (1976, pp. 244-249; 1980b, pp. 207-209). Others have since found more evidence that an error may be affected by more than one processing level (Harley 1984). As Garrett's later model is beginning to include elements of parallel processing, it remains to be seen whether this will be used to allow more than just one type of linguistic feature to affect error creation.

2.1.3 Expansions upon Garrett's Model

As mentioned above, many of the criticisms or expansions upon Garrett's model have focussed largely upon his lack of discussion of the role of lexical accessing and retrieval in error production. Shattuck-Hufnagel's work, individually (1979, 1983) and in collaboration with Klatt (1980), has treated how lexical items are inserted into their slots. Basically, she proposes (1979) that the phonological representation of a word is accessed as a whole first — presumably at Garrett's Functional Level as the selection or integration of grammatical morphemes is not mentioned (Laubstein 1988, p. 147). These whole phonological representations are stored in a memory buffer, in the order determined by syntactic processing. Errors of word selection and word ordering presumably take place before and at this step, respectively. The segments that

make up these words are then copied by a scan-copier into the sentence slots, phoneme by phoneme. The source of these phonemes is the phonological component of the language itself, not the memory buffer. Each phoneme is checked off by a monitor as it is used. Anticipations, perseveration, exchanges, additions and deletions are considered to be the result of malfunction in either the selection of phoneme and/or the monitor. Another monitor, an error monitor, scans the sentence later (whether immediately after scan-copying or when the word or sentence is complete is not clearly stated) and looks for incorrect or unusual sequences (1979, pp. 310-311). It can also create errors by editing out segments or sequences that should repeat.

The number of syllables, and the individual phonemes' syllable positions are represented in "a framework of slots provided by canonical representation in terms of larger units, perhaps syllables and feet" (Shattuck-Hufnagel 1983, p. 131-132). It is not clear whether Shattuck-Hufnagel believes that this information is present in the lexicon (Laubstein 1988, pp. 146-152). It is also not clear how the slots are marked so that the correct segment will be chosen for it. As Laubstein (1988) also points out, it seems unusual that Shattuck-Hufnagel assigns the responsibility for creating syllable-part errors to a scan-copier that has no access to the representation of the syllable, and that copies segments regardless of their syllable or word position. The role of lexical stress in errors is discussed more as a static part of the linguistic environment rather than an actively assigned element — which it is for Shattuck-Hufnagel, as it is placed in the lexicon, and is a characteristic of the slot-representation (1983).

Shattuck-Hufnagel limits herself to discussing lexical retrieval. There is little discussion of the accessing and retrieval of grammatical morphemes, except to note that phones that are grammatical functors (e.g. plural -s and -z) behave differently from non-grammatical phones

(1979, 1983). Although some mention is made of how the activation of competing words can lead to blends, and how blends and substitutions and many errors are often phonologically and semantically similar (1979, 1983), in direct contrast to Garrett's denial of such possibilities, this does not constitute a significant part of Shattuck-Hufnagel's model. Once again, it is not clear how such information, residing in the slot-representation or at earlier stages, could affect the actions of the scan-copier. What is most problematic, however, is the initial separation of the segment from all other parts of the grammar (Laubstein 1988, p. 155). While it might seem that the segment is treated as if it were a motor code (Shattuck-Hufnagel 1979), which would justify its separation from all other things linguistic, it is also implied that these scan-copied segments must undergo further phonological readjustment.

For Crompton (1981), the outcome of the retrieval of segments is an articulatory programme — although the comment just made about Shattuck-Hufnagel's model also holds here. The slots the retrieved items are to fit into are marked for everything — grammatical information, syllables, phones, and the features comprising the "classical phonemic representation" (Crompton 1981, p. 706). The articulatory programme, the only thing lacking, is retrieved in chunks, one syllable at a time. Thus, according to Crompton, all speech errors, no matter how small, are syllable errors. For example, when a word such as 'dog', is retrieved, all syllables beginning with /d/ are activated, all syllables with /a/ as the nucleus are activated, and likewise all syllables ending with /g/. The syllable /dag/, then, should end up receiving the most activation, and thus end up being scan-copied into an incorporator, which orders the selected syllables and performs all and any further linguistic processing (1981, p. 682). Errors can arise before the syllable is even activated, e.g. word errors, but the wrong syllable

representation can also be searched, which is an error of 'addressing'. As more than one syllable or word is being processed at one time and many syllables are activated, the retrieval of incorrect syllables is often due to the large amount of 'noise' in the system. Competing words also add noise to the system.

Unlike Shattuck-Hufnagel Crompton does not separate the segment from the word, but there are other problems with this model. He proposes that grammatical affixes should be treated as lexical items, and accommodated to their environment by the incorporator. Part of Garrett's motivation for assigning the processing of content and functor morphemes to different levels was the distribution of errors, namely that for certain types of errors, content and grammatical morphemes did not interact. Crompton's model completely wipes out such distinctions, without proposing any means of achieving such an error distribution.

Lexical stress and its role in speech errors also receives little attention from Crompton. He does note that "[t]he fact that stress and syllabification are predictable and therefore not marked in lexical representations raises further problems for our hypothesis" (p. 674), and even cites works on stress, but he does not deal any further with this problem. He proposes (pp. 689-690) that reduced-vowel syllables are represented separately from unreduced syllables in the library of syllable routines. The comment quoted above, however, makes it unclear whether or not he assumes that stress is lexical or generated by rule.

2.1.4 Neurologically-based models

The neurologically-based models of speech production, most notably those of Stemberger (1985) and Dell (1986, 1989; Dell and Reich 1980, 1981), share many of the features of the

previously discussed models; it is how these elements interact that is the major difference. In these models, neural activation is seen as spreading from a semantic or pragmatic processor, where thoughts (non-linguistic) are converted to some type of linguistic representation — seen by some as semantic in nature (Stemberger 1985, p. 147) — to the syntactic processor and the lexicon and all other parts of the speech processor. The semantic and syntactic processors and the lexicon are all interconnected, and it is the combination of activation from many sources that results in a morpheme's receiving enough stimulation to be chosen.

Speech errors are seen as resulting from noise within the system. Some of this noise is caused by the spreading activation itself, as activation spreads not only to one or two specific lexical items, but also to other items that are related within each area of the linguistic processor (semantic, syntactic, and phonological). For example, when the set of semantic features for the word 'cat' is activated, other words with shared semantic features such as 'dog' or other small, furry, four-footed mammals are activated, and can possibly interfere with the selection of the word 'cat'. Noise in the system can also be the result of processing other phrases, or even be the result of non-linguistic stimulus, such as, to continue the example above, if a dog were to walk into one's line of sight (Dell, 1986). And while it might be thought that spreading activation would result in so much noise in the system that correct speech would be impossible, the effect of the spreading is mitigated by the simultaneous decay of the signal, such that only those items that receive much stimulus can compete for selection (Dell and Reich 1980; Dell, 1986).

The Motley et al. (1983) model, which is not truly a neurally-based model, includes editing as a distinct process at each stage of production, performed by an editor. Dell (1986) has

found in his computer simulations of two-word utterance production that if the activation spreads bottom-up as well as top-down, such a mechanism is not required (also proposed in Dell and Reich 1980, 1981). The spread of activation both ways, decay, and the fact that the morpheme must achieve a certain threshold of stimulus to be chosen acts as an editor in and of itself.

Neurally-based models are new, and not only has little work been done in this area but relatively little is known for certain about the makeup of neural networks in general, yet already there are differences in approach. Motley et al.'s and Stemberger's models both assume that more than one lexical item may be processed at a time — processing is not seen as following the same order as the final surface representation (Motley et al. 1983, p. 82; Stemberger 1985, p. 152). This is not a very radical claim, however, as even Garrett's model did not fill its slots according to the final surface order. For these models, the issue of how morphemes are integrated into their surface order remains a problem. Dell (1986, p. 287, and 1989, p. 160), however, proposes that morphemes are processed in what will be their surface order. While there is parallel processing in Dell's model, it is only that different items and/or levels of processing are being performed concurrently, rather than one or many items being simultaneously processed in all areas of linguistic processing. This outlook is possibly due to the fact that Dell deals mainly with phonological processing, and his simulation studies, like Motley et al.'s error production studies, produce only two-word items.

For all of these models, however, the issues of how large a chunk of code is processed at once, and the nature of the linguistic processors and their representations are still unclear. While there seems to be little difference in the steps in processing proposed in these two types of models, the radical difference in the conformation of these two types of models makes it

difficult to evaluate this point. Although some models are fairly detailed in one area or another, e.g. Garrett's in syntactic processing, Crompton's and Shattuck-Hufnagel's and Dell's in phonological retrieval, no model is highly detailed at all levels. This is largely due to the uncertainty as to the exact nature of the processing performed at each level — a situation that cannot, at this early stage of research, be remedied.

2.2 Models of Second Language Speech Production

There are few actual models of L2 speech production. As noted previously, this is a relatively new area of study, and is entirely based upon previous work on L1 speech production. It is generally assumed by the authors whose work I will present that L2 speech production uses the same general speech production processors as the L1 (Wiese 1984; Crookes, 1991). In these papers, the topic is rather, for each particular linguistic processor, how L2 processing might differ from L1, in the event that one finds evidence that it does.

One problem with this field, attributable to its youth, is the lack of material. Books or articles that might seem to deal with this subject matter turn out, on closer examination, to have little to do with speech production processing per se. One example of this is Berg's *A Cross-Linguistic Comparison of Slips of the Tongue* (1987), which does not deal with bilingualism at all. Some of his findings are nevertheless useful. The book contains a bibliography of speech error papers, books, and corpora for many different languages. The incidence and types of slips found in other languages are compared with those occurring in English, to evaluate the universality or language-specificity of commonly proposed features of speech processing, as attested by speech errors. Much of Berg's in-depth cross-linguistic examination of speech errors

refer to German errors, as Berg has his own corpus. The author finds no data that contradict current models of speech processing, and believes that the differences found were due to differences in the structures of the languages examined. One such difference (p. 19) is that a language that has few if any consonant clusters will not exhibit the same phonological error pattern as English, a language replete with consonant clusters.

Other papers (Sridhar 1989; Bates and Devescovi 1989) have examined universal cognitive, semantic and pragmatic features of language processing — but not bilingualism per se. Sridhar compared utterances from ten different languages of different types (e.g., Chinese, Hungarian). The focus on the expression of the perception of motion, temporal and positional order, agentivity, etc., in each language individually renders this article of little use to bilingual studies. The latter article, however, could be applied to a study of Italian-English bilingualism. It examines the comparative use and function of several complex syntactic structures in Italian and English. They found that the frequency with which a syntactic structure is used can vary between languages. In this case, it can be ascribed, in part, to the structures having "a wider range of meanings in Italian" (p. 252). Information of this type would be useful to studies of bilingualism.

The papers collected in *Second Language Production* (Dechert, Mohle and Raupach 1984) give few if any specific details that might distinguish L2 from L1 speech processing. This is because the data base used has been examined only for hesitation phenomena — e.g., pauses, hesitations, false starts and, for some authors, corrections. While these criteria do render accessible the larger structural units being processed, it does not give much information on the nature of processing being performed upon these units.

Raupach's study (1984) does contain some useful observations for evaluating L2 speech. He claims that "certain formulaic items and schemata occur with excessive frequency, so that even if their form is not erroneous, their stereotyped use often results in non-idiomatic performances" (p. 134). However his definition of the term 'formulae' — any string of connected speech bordered by pauses — is unusual, as the more common definition would be "prefabricated routine" (p. 115). While Raupach's definition fits in with the hesitation phenomena framework used throughout the book, unfortunately the term is sometimes used in a manner more consistent with its usual designation, which is confusing. While the study itself deals with German-French bilingualism, the formulae studied, as defined, shed but little light upon the process of speech production. The author finds that less proficient bilinguals vary greatly in the length and types of formulae used, whereas more proficient speakers are near-native-like in the location of pauses and hesitations. The author notes, however, that this has no connection to the Subjects' knowledge of vocabulary (p. 134), nor, presumably, semantics or syntax or accent.

Wiese's paper (1984) uses a fairly standard model of speech production, although he believes that feedback loops may be required. It is this author who includes corrections with the hesitation phenomena, yet there is no discussion of the types of errors corrected (phonological, syntactic, or lexical). The corpus on which Wiese's discussion is based is the transcript of a story-telling task administered to 16 German and 16 Anglophone Subjects. Half of the Subjects in each group spoke in their native language, and half in the second language. The only differences Wiese found he ascribes to qualitative differences in the use of linguistic processing modules/levels. This may be caused by a lack of automaticity, high processing load, and perhaps

also to more monitoring. Yet Wiese concludes that "no property of L2 production has been found which cannot also be found in L1 production (p. 21)."

The final paper in the book (Dechert 1984) summarizes the authors' findings. Dechert isolates six hypotheses — the Procedural Knowledge Hypothesis, the Competing Plans Hypothesis, the Hypothesis Testing Hypothesis, The Island Hypothesis, the Task Stress Hypothesis, and the Monitor Hypothesis. He proposes that they all, simultaneously and to varying degrees, affect and/or describe various aspects of L2 speech production. Dechert, however, adds a "feed forward link that regulates and controls in advance what is planned at the preutterance stage" (p. 226), as well as a feedback loop. He also proposes that "there may be interference from L1 on all levels of processing" (p. 227).

2.2.1 Performance and Second Language Speech Production

In the event that, as we have seen, the L2 speech production processing does not differ significantly from that of L1, what, then, can we find to account for the phenomenon of foreign accent? As in the opening discussion of explanations of foreign accent, once again performance factors seem to be the deciding factor, and have been overlooked by most of those who work in this area. However Neufeld's Pre- and Post-articulatory Verification model (PAV) attempts to fill this gap. While not a complete model of speech production processing — a claim that no one makes — it can serve to add more completeness to existing models, either the more traditional or the newer neurologically-based speech production models.

Neufeld's position, based on data from his own studies and from Swain and Lapkin (1986), among others (see Section 1.5 for discussion), consists of three main points. The first

is that older learners tend to concentrate more on the form of L2 utterances. Even those who do not learn an L2 in a classroom tend to take care when constructing a sentence in order to be comprehended. This focus on form, be it conscious or unconscious, results in the use of the knowledge — taught or untaught — that one has about the language (metalinguistic knowledge), as well as the naturally internalized syntactic representations. Neufeld believes that the use of this metalinguistic knowledge interferes with the normal method of utterance construction, which relies on natural linguistic representations.

While this does imply the use of a monitor and editor of some sort, they are quite different from the mechanisms proposed by Krashen (Section 1.3). Krashen's Monitor is composed of consciously known rules of grammar; Neufeld distinguishes between a natural internal editing function, similar to those proposed by Motley et al. or Dell, and that metalinguistic interference comes from both conscious and unconscious knowledge about the language. Krashen's Monitor is considered as rendering speech more grammatical, if less fluent, whereas Neufeld holds that the attempt to use metalinguistic knowledge interferes with both processing and editing and impairs grammaticality. Similar points of view are also held by Malakoff and Hakuta (1991), Sharwood-Smith (1991), and Snow et al. (1991). Empirical support for this point of view comes from Green and Hecht (1992). On an experiment testing the correspondence between grammatical correction and knowledge of rules, they found that only 43% of the corrections that were grammatical were given by Ss who knew the correct form of the rule.

As the second point, Neufeld proposes that there is another processor, which feeds this metalinguistic and other cognitive and perceptual data into the linguistic processing stream.

Although Piagetians would hold that such a processor could not develop until puberty, on the basis of Oyama's (1976) and similar studies, Neufeld posits that it begins to develop as of school age, at 5 or 6 years (Malakoff and Hakuta 1991; Snow et al. 1991; Watson 1991).

The third point is the effect that this processor has on pre- and post-articulatory verification. The interference from the 'pseudo-linguistic' processor slows down the more basic semantic and syntactic processing, due to the increase in material to be processed. In order to respond within the constraints of real-time processing, pre-articulatory verification — when final phonological rules and phonological fine-tuning are performed — is sacrificed. It is the non-meaningful part that is most affected, namely the phonological fine-tuning. Post-articulatory verification, where suprasegmentals (intonation, phrasal stress) are edited, suffers likewise.

According to this model, intermediate-level L2 speakers who are unsure of their ability to speak an L2 are likely to rely heavily on metalinguistic knowledge to verify the constructions that they have made with their underlying competence, and so cannot perform phonological editing that might better realize their underlying phonological competence in the L2. In other words, they are more likely to have a pronounced foreign accent when overloaded with metalinguistic processing. A more advanced student is more likely to exhibit foreign accent when using unfamiliar words or constructions that require more (metalinguistic) attention to proper form, or during a stressful situation (e.g., emotion, job interview in the L2).

Appel's (1984) study of the effect of instruction on story-telling ability has results that I believe support Neufeld's model. An advanced English student was tested before and after a five-month course on 'text-processing'. The author ascribed the Subject's improved performance on the story-telling task (criteria measured were story completeness and hesitation phenomena)

to increased metalinguistic processing (p. 205). Applying Neufeld's model, I would interpret the Subject's improvement as a result of having less metalinguistic interference with normal linguistic processing. By the end of the course, the Subject had learned appropriate ways to structure a narrative; there was no longer any interference from the (metalinguistic) need to consider multiple possible narrative structures. Further confirmation of my interpretation might have come from speech error data. If fewer errors were performed on the post-test, it could be interpreted as indicating that story-planning was no longer interfering with utterance planning. Unfortunately, such information was not tabulated.

Further support is provided by Flanigan's (1991) study on bilingual competence and the performance abilities of children, although I reinterpret the data considerably, as with Appel's (1984) study above. The tasks — a grammar (cloze) task, a reading task, an interview and a story-retelling task — studied the effect of formality on grammatical errors. The author claims that the data support what she calls the

Labov-Tarone-Ellis model: as speaking style shifts toward the vernacular and away from a formally constrained task, grammar variability increases; ...for non-native or IL speakers it means 'errors' are produced, i.e., forms which deviate from the standard target language norm for verb usage. (Flanigan 1991, p. 225)

Support for this model would consist of high error counts on the more 'vernacular' tasks, the interview and story-retelling tasks, and low error counts on the reading and grammar tasks.

One problem with the model is that the 'vernacular' of the L2, even for these immersion students, may not be the same as that of native speakers. Another problem is that Labov has characterized the vernacular as the most consistent and stable variety of speech, thus it should be the level of language least prone to error, although the speech rate and range of sentence

structures available could wipe out such an effect. The last problem I find with the model is that the 'standard target language norms for verb usage' are usually literary rather than vernacular norms — although this did not seem to affect the data.

The major problem with the model is that the data do not support the model. The author herself notes that the distribution of scores for only one of four measures support the data, and not for all tasks. Flanigan claims that the problem is that the Ss do not consciously know the rules (i.e., the model presupposes a Krashen-style Monitor). She then points out that on the vernacular tasks the Ss unconsciously and properly use forms that they committed errors on during the grammar task. The data presented show that overall, fewer grammatical errors are committed on the vernacular tasks. Moreover, when examining the distribution of scores by age, the older Ss' results seemed to show the effect of focus on form even on the vernacular tasks, as their scores on these tasks began to approach the scores on the formal tasks. I interpret these data as showing metalinguistic interference on the formal tasks, and some on the vernacular tasks for the older learners, supporting Neufeld's PAV model.

* * *

The present study is intended to address this gap in the research. Highly proficient Francophone bilinguals and Anglophones performed five tasks designed to overload the speech processor, by requesting both accuracy and speed, inducing speech errors. In order to evaluate the effect of age of acquisition on Neufeld's pseudo-linguistic processor, the Subjects varied in

the age at which they acquired English. In the event that tuition, with its emphasis on form and lack of natural input, could account for the existence of a pseudo-linguistic processor or affect its use, Ss who had acquired English in a non-academic environment were included.

Chapter Three

Method

The primary aim of this study was to examine how metalinguistic processing may interfere with speech production processing, using the criterion measures of response time, speech errors and foreign accent. Age of acquisition and acquisition context (formal/natural) were also varied in order to examine their effect on the existence or strength of the pseudo-linguistic processor, as proposed by Neufeld.

The chief assumption was that non-native-sounding speech frequently results from problems that arise during the implementation of linguistic knowledge — due to interference from metalinguistic knowledge — rather than from lacunae in competence. A secondary goal was to obtain data that may contribute to further study of the PAV model discussed earlier. If results are obtained which indicate a relationship between type and frequency of production errors, language acquisition context and/or age of acquisition, the idea that L2 competence may contain both 'real' and 'meta' structures will be supported. It is hoped that the data from this study will also contribute to our understanding of speech production processing in general.

3.1 Hypotheses

3.1.1 Working:

1. Linguistic knowledge acquired during the first five years of life, L1 or L2, will be comprised primarily of real structures.

2. Competence acquired in L2 after the age of five will be comprised of both real and 'meta' structures, the proportions of which will depend upon the 'naturalness' of language data.
3. Metastructures slow down the encoding process because of the need to translate these 'pseudo' rules into real rules.
4. Bizarre, anomalous or otherwise malformed structures in utterances or print material are disruptive and require additional processing time.
5. Disruptive elements in stimulus material will generate speech errors in paced oral production or reproduction tasks, systematically, according to the processing level they affect, which may affect response-time negatively.

3.1.2 Tested:

1. Anglophone controls and 'balanced' (defined below) French-English bilingual Subjects who learned L2 prior to six years will not differ significantly in type or frequency of experimentally-induced speech errors or response-time in English.
2. Significant differences with respect to Controls will occur with balanced French-English bilinguals who learn L2 after age five in a natural context.

3. Response-time and errors will be even greater and more varied for bilingual older learners who acquire the L2 in a primarily formal (school) context.

3.2 Subjects

3.2.1 Control Group

Five native Anglophones between 18 and 22 years of age were the controls for this project. Baseline data from these controls was necessary for theoretically relevant inferences about the errors and response times of the Francophone Ss to be made. First, information was obtained about the length of stimulus presentation time (threshold) below which errors occurred with 'normal' sentences (Task 2, Section 3.3.2). Second, data on type and extent of disruption resulting from morphological, syntactic and semantic anomalies was gathered. These particular Ss were selected from a larger group of anglophones. Those not considered for inclusion in this group were those who did not perform within one standard deviation of response time on Task 1, and who did not perform adequately on Task 2.

3.2.2 Experimental Group

20 adult French-English bilinguals were divided into four groups of five. Subgroup 1 was comprised of persons who acquired English as a second language before age six (hereinafter referred to as the Bilinguals). Subjects in Subgroup 2 learned English between the ages of six and ten in a primarily natural (street) setting (referred to as SC, for street context). Subjects in

Subgroup 3 will also have learned English between six and ten, but in a largely formal (school) context (referred to as FC, for formal context). Subgroup 4 consisted of people who learned English after the age of 10 (referred to as 10+ group). All Ss were selected according to the following criteria:

1. Over 18 years of age.
2. French as dominant language for both parents.
3. No high proficiency in a third language.
4. Highly proficient readers of English (i.e., performed within one standard deviation of the response time of the Control group on Task 1, and evaluated the sentences as the Control group did).
5. It was believed that these Ss would be 'balanced' in fluency in French and English (any difference in response time on Tasks 3 and 4 similar for both the Experimental groups and the Control group, and, if truly balanced, little difference in type-token ratio on Tasks 4 and 5).
6. No obvious accent in English when first contacted.

3.3 General Procedure

A PC-type computer was used to perform all experimental tasks, the stimulus items being displayed on the screen. Stimulus items were presented individually for varying amounts of time (see detailed task descriptions below), with approximately one second between items, except in Task 2. The testing was performed in two sessions, with the English tasks (Tasks 1-4) being

performed at the first session. Task 5, the French task, was performed at a later date by some of the subjects. The Control group did not perform Task 5.

Subjects responded to all items on Task 1, Task 4 and Task 5. For Task 2 and Task 3, Ss responded only to those items cued with a beep. The beep sounded approximately one quarter of a second after presentation of the stimulus item. The constraints for cuing items was as follows:

1. beeped items, which required verbal response (with or without modification, see task description below), were followed by at least one nonbeeped item to avoid cascading effects.
2. no more than three items of one type occurred consecutively (see below for stimulus types).
3. no more than four uncued items occurred consecutively.

The pattern of cued versus uncued items was fully randomized within each task and for each S. The order of presentation of items within tasks was also fully randomized for each subject. The order in which the tasks were performed was semi-random for each subject, as Task 1 was always performed before Task 3.

All stimulus items belonged to one of six categories, a maximum of 30 items per category. The items used in tasks 4 and 5 differed significantly, not using all six categories, and were constructed differently. The stimulus items for tasks 1, 2 and 3 were as follows (the stimulus items are given in Appendices A-E):

1. **Normal (N): S-V-O1-O2 (O1, direct object, O2, indirect object) statements with no structurally or semantically bizarre or anomalous elements.**
2. **Semantically Anomalous (SA): Structurally well-formed S-V-O1-O2 statements in which the relationship between subject and predicate is inappropriate.**
3. **Semantically Bizarre (SB): Structurally well-formed S-V-O1-O2 statements in which the relationship between subject and predicate is highly improbable, but not impossible.**
4. **Morphologically Malformed (MM): Semantically unambiguous and syntactically well-formed S-V-O1-O2 statements with incorrectly inflected verbs.**
5. **Syntactically Anomalous (StA): Semantically unambiguous S-V-O1-O2 statements with correctly inflected verbs, but an unacceptable VP structure.**
6. **Syntactically Bizarre (StB): Semantically unambiguous S-V-O1-O2 statements with correctly inflected verbs, but an improbable VP structure.**

To reduce the number of contingent variables arising from differences in structure, bizarre or anomalous elements were confined to verbs or verb phrases. All sentences were 8-9 syllables in length. While it was intended that the words used to create the stimulus items be in common use there were, nonetheless, some problems due to vocabulary, and also to a couple of

typographical errors (see Appendices). Each task was preceded by 12 sample items, consisting of an appropriate number of sentences from each category.

3.3.1 Task 1

Task 1 consisted of 180 sentences, 30 stimulus items for all six categories, as described above. As the sentences flashed upon the screen, the Subjects indicated the 'goodness' or 'badness' of the item by pressing the appropriately labelled button. Ss had up to 10 seconds to judge each item, after which the next item was presented. Items not judged in time were counted as 'bad' items. The items were presented in groups of sixty, with a two-minute break separating the groups of items.

Task 1 was always performed before Task 3. This allowed the response time and judgements from this task to appear on the Task 3 printout as 'perception time'. This task functioned not only as a measure of reading time, but also allowed us to verify the validity of the categories presented. It also served as a criterion measure of reading proficiency; the RT of all Ss chosen were within one standard deviation of the mean response time of all the Anglophones who had performed the task.

3.3.2 Task 2

Task 2 consisted of 100 normal sentences. For this task, Ss responded to cued items only (N=33). At the beginning of the task, sentences were presented for a maximum of 1.5 seconds. Presentation time of the stimulus items decreased by .1 seconds for each item, whether or not the

S was to respond to that item. Response consisted of reading the sentence aloud. All Ss were informed that they need not change or correct the sentences on this task.

As the next item was presented upon .25 seconds of silence following vocal activity, it was imperative that Ss respond as smoothly and fluently as possible, with no pauses or false-starts or pre-verbal vocal activity (pre-vocalized consonants, lip-smacking) before beginning the sentence. This would usually result in the computer's presenting the next item. This caveat also held for Tasks 3, 4, and 5. All Ss were requested to read as quickly and as smoothly as possible. On this task, if Ss missed five items in a row, the experiment ended automatically.

Task 2 served to verify that response times in Tasks 3 and 4 were not unduly affected by rate of speech. This task also allowed the comparison of speech errors caused by rate of speech alone with those caused by both quick response and sentence correction or sentence construction. Ss whose rate of speech was slow and/or who did not perform well on this task were not selected for further analysis. Ss who repeated from memory the item presented before the beep instead of the item presented with the beep were also eliminated.

3.3.3 Task 3

Task 3 consisted of 180 sentences, as in Task 1, of which 120 were repeated from the first task. These items were the filler items, as Ss responded to only to those items cued with a beep (N=60). The cued items in Task 3 were based closely upon actual Task 1 items.

The Ss were to read aloud the cued item as is if they thought it was a 'good' item, as on Task 1. The 'bad' items, however, were to be changed into syntactically and semantically acceptable sentences. Items could be corrected by adding to or omitting parts of the sentence,

substituting more appropriate words or phrases, or by reordering the words and phrases. Ss had up to 12 seconds to respond to stimulus items. As in Task 1, there was a two-minute break after sixty items had been performed.

In Task 3 the element of correction is added to the sentence reading process of Task 2, while Task 1 was a reading comprehension task. Both Task 1 judgements and response times were included on the Task 3 printout. If Task 1 gives the reading and evaluation time, and Task 2 gives the reading and speaking time, then any increase in processing time above the combination of Task 1 and Task 2 response times may be due to the increased processing load. The consistency in judging the 'bad' items can also be compared with the Task 1 responses. Moreover, errors triggered by the need to evaluate and correct sentences may be revealed by comparing errors on Task 3 with those of Task 2.

3.3.4 Task 4

Task 4 contained 120 stimulus items of 3 keywords (mostly N-V-N). The only categories used were normal, semantically bizarre, morphologically malformed, and syntactically anomalous. There were 30 items in each category. The semantically bizarre category was further divided into three sets of 10 stimulus items, in which the anomaly consisted of inappropriacy between subject and verb, verb and object, and subject and object. These subcategories are not analysed in this study. Syntactic anomaly was created by reordering the words in the stimulus item (N-N-V).

Subjects were to use the stimuli to create semantically and syntactically correct sentences. They did not have to use all three keywords, and were to correct any errors in the keywords

presented, if used. Ss had approximately 15 seconds to respond to items. There was a two-minute break halfway through the task.

The purpose of this task was increase the processing load by forcing the Ss to create more of the utterance. While the stimuli presented are not considered to be analogous to the type of underlying semantic frames actually used in the construction of a sentence — as all the verbs presented were inflected — it was hoped that the semantically bizarre and syntactically anomalous items would force decisions at an early level of sentence construction. Moreover, if age of acquisition would affect error type and rate, it was believed that it would most likely appear in the results from this task, as it is the most creative of the tasks performed by all Ss.

The semantically anomalous category was discarded for two reasons: the first being, in the event that the stimuli might resemble the Positional frame (Section 2.1.2), it was unlikely that anything so anomalous would ever be created, and presenting the Ss with such items might lead to skewed results. The second was that pilot testing showed that when these items were included, they were accepted and skewed the results. When this category was removed, the anomalous category was varied as explained above to keep the stimuli from being unduly limited. Moreover, as some of the verbs used could have more than one interpretation, it would be impossible to truly have only subject-verb anomaly, in any case.

Using only three words in the stimulus items made it impossible to induce syntactic bizarreness, and so that category was discarded. While it might be noted that the order of words in the syntactically anomalous category (see Appendix D) could be used to create perfectly good sentences with a left-dislocated object — e.g. 'man banana slipped' could be made into 'The man on the banana peel slipped' — it was felt that between the time limit and the usual

presentation bias of SVO, this interpretation was nigh impossible. This was found to be the case during both pilot testing and the experiment itself.

3.3.5 Task 5

Task 5 was essentially Task 4 translated into French, the stimuli being altered where necessary, and also to prevent boredom. All items were printed in uppercase letters, in order to avoid the problem of accents.

These last two tasks were the main criterion measures of 'balancedness'. If response time and type/token ratios varied significantly between the two tasks, the S was not considered a balanced bilingual.

3.5 Measures

Aside from response latency and speech time, speech errors and other behaviours were examined. While based upon the types of phenomena used in L1-based speech production models, other criteria have been added. The phenomena are: semantic errors, syntactic errors, phonological errors, dysfluencies, corrections, substitutions, additions, omissions, number of incomplete items, number of items to which no response (null response) was uttered, and number of items read as given on screen ('read as is').

A single error could be counted under more than one category, e.g. missing a verbal inflection was counted as both a syntactic error and an omission. Items 'read as is' are counted only on Task 3, as it was unlikely that the three-word stimuli of Tasks 4 and 5 would be read as given, although some were. This category did not apply to N items, and for MM items it

referred to the morphological form of the verb only. These items were not double-scored (i.e., counted as syntactic or semantic errors as well), as items may have been read as given due to the Ss being unable to correct items in time or deciding that the sentence is 'good'. This score will be contrasted with the category evaluation from Task 1.

The semantic errors were difficult to assess, as I attempted to avoid including anything resulting from the use of an inappropriate preposition, for example, which could be considered a syntactic error. Nonetheless, overlap could not be completely avoided. Although I include these 'errors' in my tables and discussion, due to a lack of confidence in this category, they will not be given much importance. Syntactic errors included missing or incorrect or redundant inflections, and missing or mismatched prepositions and determiners or gender inflections.² Phonological errors covered such classic speech errors as anticipations, perseverations, blends and malapropisms (choosing a phonologically related but different word, of the same syntactic category, although these errors could be the result of semantic error). Simplification of a name (e.g. Susie instead of Susan) was counted as a substitution, as were otherwise correct but unnecessary changes of preposition, determiner, tense, plurality, etc. All other changes were counted under one of the other headings.

Dysfluencies covered hesitations, stammers, slow or extended pronunciation of a word, syllable or phoneme, and word repetition not associated with correcting an error. Due to some inconsistency in transcription, non-linguistic fillers (ums and ers) could not be included. The correction category covered corrections to all errors. A correction was also counted when an S

² When problems arose during transcription, homorganic sequences were not counted as phonological or syntactic errors. When there was no agreement to confirm that a lexical item was plural, when the situation would normally require a (regular) plural in that position, these items were usually counted as phonological errors.

would 'correct' a substituted word back to the one given on the screen, even if the substituted item was not incorrect. For Tasks 4 and 5, due to the nature of the task, substitutions were almost always this type of change. Additions and omissions were counted for a single phoneme, a word (usually an adjective, preposition or noun), or a whole phrase (NP, superordinate or subordinate VP or an S). Errors within additions were counted. On Tasks 4 and 5, which had N-V-N stimuli, only additions or omissions of whole phrases (NPs, VPs, etc.) were counted as such, due to the nature of the task.

Assessing incomplete utterances was rarely difficult, as they were usually marked by dysfluency and an incomplete intonation curve, and the click or beep signalling the presentation of the next item was audible. While the 'no response' category would seem to be self-evident, responses of "I don't know" or "It makes no sense" were counted as null responses, even with respect to the type-token counts, below.

Type-token counts were performed on the responses to tasks 4 and 5. A token count is a count of all morphemes (meaning units) used, and thus contractions were counted as two (or more, as appropriate) words. Words repeated when the Ss were correcting errors, or as gap-fillers, were included. A type count was a count of all the individual words used; e.g. even if 'the' is used 70 times, it is only counted once. On both of these counts, differently inflected appearances of a verb did not count as separate words, nor did plural forms or gender variants. When a word was used in two different syntactic categories — whether due to inflection or derivation — e.g. as both a verb and a noun (French 'inviter' and 'invité'), two words were counted. As with English contractions, the French words 'du', 'des', 'au', etc., were counted as two words.

These word counts will be compared with the previous measures. The number of words counted, the tokens, may vary with the number of incomplete and null answers, additions and omissions. The number of words used may be affected by latency. The difference between the type and token counts may reflect the speakers' command of both vocabulary and structure, in that if a speaker can create many different kinds of sentences, a higher type count may result from using a wider variety of auxiliaries and prepositions, etc..

The statistical analysis performed on the response-time data was a modification of Analysis of Variance (ANOVA) called Quasi-F, intended for groups less than or equal to 12; Chi Square tests were performed on all other data. The Chi Square was performed only on those categories or groups that had at least one score of 10 or more. When a group or category was revealed to have a significant distribution of scores, the relevant scores were discovered by elimination, and confirmed with paired tests.

Chapter Four

Results

4.1 Task 1

All Ss, including the Controls, were chosen only if they were highly proficient readers. The mean overall response-time (RT) on Task 1 was calculated for all Ss, and they were considered highly proficient readers if their response time on this task fell within one standard deviation of the mean of anglophone performance. Thus RT on Task 1 was not significant between groups. While there is some variation in RT, as seen on Table 1, no significant differences were found between categories on this task as well as on Tasks 3, 4, and 5. While the scores of the SA and SB categories are on average higher than those of the other categories, it must be noted that there are similar high scores on the StA and StB categories. And while the FC group has the highest average RT, its scores are not significant, so the average would not be significant either. These results are merely a trend, at best. Nor does there seem to be any pattern to these high scores — the 10+ group on the SA category, the FC group on the SB, StA and StB categories.

The other measure produced by this task is category evaluation. As indicated on Table 2, Chi-square tests of the distribution of scores for each category reveal significant differences (.05 level minimum) on four categories — SA, SB, MM and StB. When the scores are examined by group, the SB and StB scores prove significant for all groups. The SA scores are significant for only the Controls and the 10+ group. In the SA category the scores of the Controls and the 10+ group are significant with respect to the Bilingual group. The 10+ group's score is also

Table 1: Task 1, Response-Time Latency

	N	SA	SB	MM	StA	StB	Avg.
Ctrl.	6683	6191	7454	4396	7489	6911	6521
Bil.	6994	6328	8607	5544	5911	5234	6436
SC	5104	5254	6450	4426	5303	5464	5334
FC	7735	7940	9272	7150	9135	9146	8396
10+	7865	10065	8400	6148	6088	6133	7450
Avg.	6876	7156	8037	5533	6785	6578	

(The groups are, as follows: Controls, Bilinguals, Street Context, Formal Context, and those who began learning English after the age of 10. The category headings are: Normal (N), Semantically Anomalous (SA), Semantically Bizarre (SB), Morphologically Malformed (MM), Syntactically Anomalous (StA), and Syntactically Bizarre (StB). The Avg. column contains the average latency of the group. Averages for each category appear at the bottom of each column. All other figures are means. The figures are in .5 millisecond units; a count of 2000 equals one second.)

significant with respect to the SC group. In the SB category, the 10+ group differs significantly with respect to the SC and FC groups. In the MM category, the Controls and FC group differ significantly only with respect to each other. In the StB category, the Controls and the FC group's scores differ significantly from all the other groups.

There is no discernable pattern to these results. Yet it can nonetheless be said that overall, the N, the M and the anomalous categories were accepted, and the bizarre ones rejected, as intended. There does not seem to be any correlation between RT, by group or by category, and category evaluation scores.

Table 2: Task 1, Category Evaluation

	N	SA*	SB*	MM*	StA	StB*
Ctrl.	135 (75)	117* (65)	80* (44)	142* (79)	128 (71)	44* (24)
Bil.	142 (79)	137 (76)	85* (47)	129 (72)	139 (77)	11 (6)
SC	140 (78)	135 (75)	70* (39)	134 (74)	147 (82)	6 (3)
FC	133 (74)	131 (73)	67* (37)	116* (64)	131 (73)	40* (22)
10+	137 (76)	115* (64)	97* (54)	132 (73)	141 (78)	10 (4)

(The columns contain total of items correctly evaluated, and, below in parentheses, the percentage this represents of the total possible score (n=180). Asterisks denote the statistically significant columns and scores.)

4.2 Task 2

For this Task, RT essentially corresponds to the number of items read. As can be seen on Table 3, the number of items responded to does not vary greatly between groups. All groups responded to a minimum of 75% of the items. Keeping in mind, however, that one had to overrun on 5 consecutive items before the task would be ended, it was possible for Ss to come close to completing the task while performing relatively few items within the time limit given.³

For both sets of scores — the number of items responded to and the number completed in time — the score of the Controls is significantly different with respect to the scores of all the

³ One could perform as few as six items, out of 33, and still 'complete' the task. The lowest number completed was 12, by an FC group member.

Table 3: Task 2 Data

	No. done	Percent	No. done in time	Percent (del)	Inc.
Ctrl's	149*	90	102*	68	1
Bil	125	76	82	66	4
SC	124	75	77	62	-
FC	125	76	75	60	6
10+	129	78	65	50	1

(The columns contain the total number of items performed by group; the percentage that figure constitutes of the total number possible (n=165); the total number of items completed within the time limit given; the percentage these items constitute of the number done; and the number of incomplete items done. Scores that are significantly different are marked with an asterisk.)

other groups. Otherwise, as regards the items completed in time, there is a steady decrease down the column from the Controls to the 10+ group. The numbers of incomplete items are quite low. The highest figure, six items, formed only 10% of the items read in time by the FC group, and thus only 50% of the items performed were actually read in time — a total almost as low as that of the 10+ group. The low score of the 10+ group, however, may be due to these Ss' lack of experience reading English. These data indicate that it takes these Ss at least 1.5 seconds (3000 units) to read and utter a seven- or eight-word sentence.

All groups made changes to the items, and no group performed entirely without error. Table 4 compiles the errors, corrections, and other phenomena that each group performed on Task 2. The four categories with scores high enough to be subjected to statistical analysis — phonological errors, corrections, dysfluencies and substitutions — all contained significant scores.

Table 4: Task 2, Errors, Corrections, Dysfluencies, Substitutions, Additions and Omissions

	sem.	syn.	phn.*	cor.*	corr. rate	dys.*	sub.*	add.	omit
Ctrl	-	4 (24)	10 (15)	3 (8)	21%	12* (12)	4 (19)	-	4 (21)
Bil.	-	2 (12)	8 (12)	3 (8)	30%	27 (27)	3 (14)	3 (27)	2 (11)
SC	-	1 (6)	11 (17)	5 (13)	45%	24 (24)	1 (5)	3 (27)	1 (5)
FC	1	9 (53)	23* (35)	26* (66)	79%	25 (25)	10* (48)	4 (36)	9 (47)
10+	-	1 (6)	14 (21)	2 (5)	14%	12* (12)	3 (14)	1 (9)	3 (16)
Total	1	17	66	39	(38%)	100	21	11	19

(The first three columns of figures contain the totals of semantic, syntactic, and phonological errors; the next two columns contain the number of corrections performed and in parentheses the correction rate (percentage of errors corrected); the last four columns of figures contain the numbers of dysfluencies, substitutions, additions and omissions performed. The numbers in parentheses below the column figures are the percentage of the total errors, etc., that the column figure represents. In the totals row, the figure in the correction rate column is the average correction rate.)

The FC group, as well as having the only semantic error noted and performing the most syntactic errors, had the only significant score on both phonological errors and corrections. The 10+ group's phonological error score was the only one, for both categories, that the FC group

did not prove significant to. Nevertheless, the FC group also corrected more errors than any other group, having a correction rate of 79%.⁴ All the other groups performed 5 corrections or less.

In the dysfluencies category, the Controls and the 10+ groups have significantly low scores with respect to the other three groups. In the substitutions category, the FC group is once again significant with respect to all but the Controls. They also add and omit more than the other group. It may be noted that the numbers in the omissions column are, except for the 10+ group, identical to the numbers in the syntactic error column. This is because many of the syntactic errors involved the omission of (usually sentence-initial) determiners and prepositions. These parts of speech were most often substituted and added, as well.

4.3 Task 3

Although the RTs are not statistically significant, as noted above, the highest scores occur in the SA and SB categories, and for the FC group (see Table 5), as on Task 1. Except for these categories and this group, the RTs have decreased in comparison with Task 1. This result is not as expected, as on this task Ss must not only read and evaluate the sentence, but correct it before or as they begin to speak.

The numbers of incomplete and null answers are so low as to be untestable except for the null answers in the SA category. The score of the Bilingual group is significant with respect to the Controls, SC and 10+ groups. The score of the FC group is also significant, but with respect to the 10+ group only. For the column totals of incomplete answers, the SA score is significant

⁴ Although I have no figures to support this statement, many of the corrections involved reverting from a substituted word to the one given on screen.

Table 5: Task 3, Response-Time Latency

	N	SA	SB	MM	StA	StB	Avg.
Ctrl.	4503 1-2	6426 4-2	7296 4-2	4055 1-2	4323 1-2	4750 1-0	5226 12-10
Bil.	6870 1-2	8051 0-13	8878 6-2	4033 4-2	5619 2-0	5456 0-1	6485 13-20
SC	5539 1-3	11333 4-4	8791 3-5	5136 0-0	4220 0-0	4866 1-1	6648 9-13
FC	7502 2-1	11380 2-8	12450 4-4	6991 3-2	6003 1-3	6849 1-2	8529 3-20
10+	6220 2-1	7771 6-0	7356 2-3	4610 1-2	3812 1-3	4557 3-3	5721 15-12
Avg.	6127 7-9	8992 16-27	8954 19-16	4965 9-8	4795 5-8	5296 6-7	

(The figures underneath the means are, respectively, the number of items that received incomplete and null answers, separated by a dash. Averages for each category appear at the bottom of each column. The Avg. column contains the average latency, and, below, the total number of incomplete and null answers for the group. All other figures are means. All significant categories and figures are boldfaced.)

with respect to the StA score only, while the SB score is also significant with respect to the N and StB scores. The result for total null answers is fairly evident, with the SA score significant with respect to the MM, StA and StB scores. When tested by group, the total numbers of incomplete and null answers are not significant.

As shown on Table 6, only three of the error categories proved to have significant score distributions; the 'read as is' category, phonological errors, and dysfluencies. The status of the read as is category poses a problem. While for most groups these responses would be syntactic errors committed on the SA, SB or MM items, the FC group scored these errors on StA and StB

Table 6: Task 3, Errors, Corrections, Dysfluencies, Substitutions, Additions and Omissions

	cats.	sem.	syn.	read as is*	phon.*	corr.	corr. rate	dys.*	subst.	add.	omit
Ctrl.	N	1	3	-	2	1		13	14	1	4
	SA	1	3	13	3	-		14	-	5	2
	SB	-	-	16	4	2		13	-	2	3
	MM	-	1	-	3	1		4	3	3	2
	StA	-	-	-	1	1		7	7	2	-
	StB	-	5	1	3	-		14	10	-	4
	totL	2	12	30	16	5	(17%)	65	34	13	15
Bil.	N	-	1	-	3	1		14	10	4	3
	SA	-	3	-	1	-		7	-	2	6
	SB	1	4	9	4	4		18	1	3	12
	MM	-	2	1	5	2		9	2	3	2
	StA	1	2	-	6	2		9	4	1	2
	StB	-	-	-	4	1		20	15	2	-
	totL	2	12	10*	23	10	(27%)	77	32	15	23
SC	N	-	2	-	6	2		21	11	2	3
	SA	-	5	-	1	1		12	-	5	2
	SB	1	4	3	3	2		28	2	3	7
	MM	1	-	-	5	2		13	5	-	-
	StA	-	1	-	5	2		21	3	-	-
	StB	1	3	-	9	5		20	7	1	1
	totL	3	15	3*	29	14	(30%)	115*	28	11	13
FC	N	-	4	-	10	4		13	12	3	4
	SA	2	2	5	2	-		12	2	3	3
	SB	1	7	8	10	5		18	1	2	15
	MM	-	7	2	6	3		13	4	3	1
	StA	1	1	3	9	1		11	1	1	1
	StB	-	2	4	11	1		15	3	2	1
	totL	4	23	22	48*	14	(19%)	82	23	14	25
10+	N	-	3	-	6	1		4	12	7	8
	SA	-	5	11	2	-		9	1	5	3
	SB	-	1	17	3	-		14	-	2	10
	MM	-	1	1	6	2		12	3	-	2
	StA	-	3	-	9	1		8	2	5	1
	StB	-	1	-	8	2		11	11	1	1
	totL	0	14	29	34	6	(13%)	58	29	20	25

items as well. This pattern of response is rendered all the more peculiar when one considers that on Task 2 this group performed the most changes to the stimulus items, although the groups performed similarly in category evaluation on Task 1. However, both the Controls and the 10+ group had higher totals of items read as is. Yet it is the Bilingual and SC groups that are significant with respect to all the other groups, though not to each other. As for phonological errors, the score of the FC group is significant to all scores but that of the 10+ group, which is also significant with respect to the Controls' score.

Overall, the FC group has committed the most errors, even if one does not include the read as is count. All the groups scored more syntactic errors and more items read as is on the semantic categories, while there were more phonological errors on the syntactic categories. The Bilinguals, SC and FC groups corrected more than the Controls did, the 10+ group less. The SC group has the highest correction rate, although they corrected as many errors as the FC group. The average correction rate is only 21%, down from 38% on Task 2. This may be the result of task demands, which are greater on this task than on Task 2, and leave little time for correction even if the monitor assigns enough priority to already committed errors.

Corresponding to the rejection of the bizarre items on Task 1, we see few 'read as is' items for the StB items. For the SB items, the 10+ and Control groups read somewhat more of the items as is than they accepted on Task 1. The Bilinguals reject these items and corrected most of them. The FC group read one item less than the Bilinguals — less than they accepted on Task 1 — and the SC group read only three items. The high scores of the Controls and the 10+ groups on the SA items are also of note, as they had judged these items 'bad' on Task 1.

It is the SC group's score that is the significant score for dysfluencies, and it is significant with respect to the scores of all other groups. Otherwise, the groups perform quite similarly as regards substitutions, additions and omissions, although it must be noted that the groups with the least exposure to English, the FC and 10+ groups, perform the most omissions. All groups performed more additions and omissions on the SA and SB items, while more substitutions were performed on the syntactic categories. No one category seemed to cause dysfluencies, with the Controls and the 10+ group scoring high on the semantic categories, and the others on the syntactic categories.

Some notable results are the high scores on the Normal items for the SC, FC and 10+ groups on phonological errors, as well as on dysfluencies and substitutions, and by the 10+ group on additions and omissions. It was originally thought that the Bizarre items would prove acceptable. Together with the Normal items, then, fully half of the items presented would be acceptable as given and not need correcting. On Task 1 the Bizarre items were largely rejected, reducing the 'good' items to approximately one in six. The greater amount of correcting required greatly increases the processing load, monitoring, and use of the pseudolinguistic processor. As the N items are so rare, there is more likelihood that these items would automatically be processed like the other items, which may account for much of the error and dysfluency noted on the N category.

4.4 Task 4

As can be seen on Table 7, the RTs are lower overall on this task than on Tasks 1 and 3. As the stimuli consist of 3 keywords, less decoding is required, and the sentences constructed

Table 7: Task 4, Response-Time Latency

	N	SB	MM	StA	Avg.
Controls	2599 2-1	4181 5-8	3263 3-4	3122 2-6	3291 *12-19*
Biling.	3937 4-6	6954 *15-10	4702 3-8	4687 6-10	5070 28-34
SC	3879 3-16*	6128 6-14	5502 3-15	4698 2-12	5052 *14-57*
FC	5884 0-3	10466 10-13	6432 9-11	6485 7-10	7317 26-37
10+	7894 2-7	9640 6-9	7589 *11-8	7088 4-10	8053 23-34
Avg.	4839 *11-33	7474 *42-54	5498 29-46	5216 21-48	

can contain one less object clause than the items in Task 3. Although not significant, once again the highest RTs occur on the SB category, but the group with the highest average RT is the 10+ group. While the FC group's RT has decreased somewhat in comparison with the Task 3 results, while that of the 10+ group has increased dramatically. One possible reason for the dramatic difference in the performance of all the groups except the 10+ group is that this group did have to plan more, and took the time to do so. Some of the other groups have rather high numbers of incomplete utterances, perhaps caused by beginning too soon. Some Ss may also have begun by speaking very slowly, or using an utterance-initial non-linguistic filler, which would result in artificially low RTs.

As is consistent with both the larger number of responses and the difficulty of this task, there were many more null and incomplete answers on this task than on Task 3, and more null than incomplete responses. When the total scores of the stimulus categories are compared, only the incomplete responses are significant, and both the low and the high scores — N and SB respectively — are significantly different from the MM and StA scores. In the SB category, the Bilingual group's score differs significantly from the scores of the Control, SC and 10+ groups. For the MM category, the score of the 10+ group proved to differ significant from all the others. The null answers of the N items were testable, and the SC group differed significantly with respect to all other groups.

The responses of the group totals differ more, as there are significant differences for both incomplete and null responses. For incomplete responses, the Control group differs significantly with respect to the Bilingual and FC groups, the SC group only with respect to the Bilinguals. For null answers, the SC group is significant with respect to all the other scores, but the Control group is also significant with respect to the FC group's score.

Overall, there are more null than incomplete answers, but the score distributions of the incomplete answers were more often significant. The numbers of null answers are routinely high, but the highest score, by category, is as usual on the SB items. It is also interesting to note that group performance differed significantly more than category distributions.

The number of errors and other phenomena is also greater on Task 4 than on the previous tasks. This can be ascribed to two factors. The first is that there are thirty items per category (total possible, 120) as opposed to ten per category on Task 3 (total possible, 60). The second is that while RT on Tasks 3, 4 and 5 remained fairly constant, more processing was required to

Table 8: Task 4, Errors, Corrections, Dysfluencies, Substitutions, Additions and Omissions

	cats.	sem.	syn.*	phon.*	corr.*	corr. rate	dys.*	subst.*	add.*	omit*
Ctrl.	N	-	9	3	1	8%	24	2	8	8
	SB	3	6	7	5	31%	56	2	11	5
	MM	-	24	6	4	13%	32	4	11	17
	StA	2	19	10	9	29%	38	4	7	17
	tot.	5	58*	26	19	(20%)	150*	12	37*	47*
Bil.	N	-	15	8	4	17%	48	-	15	8
	SB	2	11	9	1	5%	82	-	13	6
	MM	2	26	9	6	16%	63	1	13	18
	StA	2	22	4	1	4%	70	5	12	17
	tot.	6	74	30	12*	(11%)	263*	6	53	49*
SC	N	1	20	12	5	15%	39	-	14	17
	SB	1	26	14	7	17%	76	4	30	21
	MM	-	33	9	7	17%	46	2	17	24
	StA	1	25	7	5	15%	47	-	8	21
	tot.	3	104	42	24	(16%)	208	6	69	83
FC	N	1	40	25	7	11%	37	1	13	34
	SB	2	34	21	7	12%	65	5	19	30
	MM	-	41	27	6	9%	37	-	18	27
	StA	2	27	24	7	13%	50	3	17	23
	tot.	5	142	97*	27	(11%)	189	9	67	114
10+	N	-	17	17	3	9%	34	-	28	18
	SB	1	22	18	6	15%	65	1	28	20
	MM	1	48	18	3	4%	59	-	29	31
	StA	-	31	18	3	6%	51	-	26	29
	tot.	2	118	71*	15	(8%)	209	1*	111*	98

create the utterances for the latter two tasks. The uncertainty factor, mentioned in the previous section, if taken as a matter of the category of the stimulus, changes to one item in four being usable without major changes as opposed to one in six. But due to the nature of the stimuli for these tasks, and the amount of creative work required, the uncertainty about whether or not the stimulus item was 'good' was likely a negligible contributor to error on these tasks. Being decreased, it may even have helped lower the RTs.

All types of errors and other phenomena (except for the semantic error category, which was not testable) had significant score distributions. On the syntactic error category, the Controls differed significantly with respect to the SC, FC, and 10+ groups, the Bilingual group with respect to the FC and 10+, and the SCs with respect to the FC group. For phonological errors, the FC group's score differed significantly from all other scores, and the 10+ group also differed significantly with respect to the SC, Bilingual and Control groups. For corrections, the Bilingual group differed significantly only with respect to the SC and FC groups. Once again the FC group committed the most errors overall, and performed the most corrections. Yet the Controls have the highest correction rate, 21%, as the number of errors they performed was low and the number of corrections relatively high. While the number of semantic errors has become notable, scores are not clustered in any particular category — a finding which may reflect the overall increase in processing required by this task. Syntactic errors, however, are highest on the MM items. While not shown on Table 8, a considerable portion of these errors were due to Ss using the malformed verbs as given. The Controls, Bilinguals, and the SC group used, respectively, 4, 5, and 8 verbs as is, which constitute 17%, 19% and 24% of their syntactic errors. The FC and 10+ groups performed many more of these errors, 15 (31%) and 19 (46%), respectively.

There is no discernable pattern in the distribution of syntactic errors by category, nor for the phonological errors, corrections, or correction rate. However, when the correction rate is averaged, one finds that it falls from 38% on Task 2 to 21% on Task 3, and is only 13% on Task 4. I ascribe this change to task demand.

For dysfluencies, the Bilinguals' and the Controls scores were significant with respect to all other scores. A similar situation prevails for additions, with the 10+ group's score differing significantly with respect to all other groups, and the Controls' with respect to the FC and SC groups only. On substitutions, however, the 10+ group was significant to only the FC and Control groups, and on omissions, the Control and Bilingual groups were significant to all other groups.

As with Task 3, there is no discernible pattern to the distribution of scores on dysfluencies and other phenomena, nor if the significant scores are the high scores, the low ones, or at both extremes. All the Francophone groups are more dysfluent than the Controls, but the Bilinguals have the highest score. Given my definition of this term (Section 3.5), it is interesting to note that the Controls performed the fewest additions, although as native speakers the 'cost' to add a clause, in terms of task demand, should be the lowest for all Ss tested. All the Francophone groups performed more additions than the Controls, but it is the group with the least exposure to and practice of English, the 10+ group, that has been most creative and performed the most additions, as on Tasks 2 and 3. Nevertheless, the Controls perform the most substitutions, although the number is low. I believe this too is due to how I have defined the term (Section 3.5), and how self-correction of this type conflicts with task demand. The Controls also performed the fewest omissions; being native speakers, they should have the least need to simplify the structure. On this measure, the Bilinguals performed comparably to the Controls.

The distribution of type-token scores (Table 9) did not prove to be statistically significant, whether tested by group or by stimulus category. Comparing the groups against each other, the Controls perform fairly consistently across categories despite differing numbers of incomplete and

Table 9: Task 4, Type-Token Data

	Ctrl.	Bil.	SC	FC	10+	Avg.
N						
token	779	797	734	809	801	784
type	513	510	492	522	504	508
diff.	266	287	242	287	297	276
SB						
token	817	750	854	798	843	812
type	479	448	514	465	498	481
diff.	338	302	340	333	345	332
MM						
token	760	771	750	759	779	764
type	499	494	495	485	518	498
diff.	261	277	255	274	261	266
StA						
token	786	768	749	767	808	776
type	510	488	489	476	516	496
diff.	276	280	260	291	292	280
Avg.						
token	786	772	772	783	808	
type	500	485	498	487	509	
diff.	285	287	274	296	299	

(The scores listed are: the total number of tokens (morphemes); total number of types used (individual lexical items); and the difference between the preceding two scores (tokens-types). Chi-square tests of the type and token counts revealed no significant differences.)

null items. Corresponding to the number of additions performed, the 10+ group also has the highest overall type and token counts, and differences, of any group. The SC and FC groups fit between the Bilinguals and the 10+ group on types, tokens, and difference, with the SCs using fewer tokens, but on average more types, while the FCs repeat much more, as is evident in the difference scores. This may be due, in part, to the two groups having high numbers of

incomplete and null answers. It is possible that some Ss 'compensated' for their null and incomplete answers by creating long utterances when they got the opportunity, keeping overall type and token scores fairly similar. Unfortunately I have not found a measure that would clearly show this effect. There is little correlation between RT and the Type-Token data. While the SB category, which had the highest RT, had the highest token and lowest type counts, the group with the highest RT, the 10+ group, had the highest token and difference counts.

4.5 Task 5

Even though Task 5 was performed in the Francophone Ss' native language, none of the groups has as low a latency on this task (see Table 10) as the Controls did on Task 4. As this may be due to differences in language structure rather than to bilingual ability, this discrepancy cannot be evaluated. The average RT of the 10+ group decreased dramatically from its Task 4 score, and although the FC group's RT also decreased, they had the highest RT. The Bilinguals and the SC groups were both slower on Task 5. The SB category once again has the highest RT.

The numbers of incomplete and null answers has dropped in comparison with Task 4. Null answers decreased the most. Category-wise, there was least change in the N counts, with incomplete answers actually increasing by 4. Not only were there fewer testable categories than there were for Task 4, of those tested, only the null answers for the Normal items and the SC group's null answer scores proved to have significant results. In the N category, the 10+ group's high score was significant with respect to the SC group only. Interestingly, it is the score on the Normal items that is the significant score among the SC group's null answers, with respect to the StA and MM scores.

Table 10: Task 5, Response-Time Latency

	N	SB	MM	StA	Avg.
Biling.	5418 5-4	6810 8-10	5747 4-7	5582 5-4	5889 22-25
SC	5568 3-2*	8147 5-6	6501 4-11	5873 2-10	6522 14-29*
FC	6333 4-4	8969 2-5	5994 1-3	6023 3-7	6830 10-19
10+	4376 3-10*	5850 5-6	5002 3-8	4699 3-7	4982 14-31
Avg.	5424 15-20*	7444 20-27	5811 12-29	5544 13-28	

When Chi-square tests were performed on each testable pair of Task 4 and Task 5 null and incomplete answer scores, five of the 13 sets showed significant differences between the tasks. The two groups are the SC and FC groups, rather than the FC and 10+ groups as the model would predict. The SC group performed significantly different on the N and SB items, while the FC group's null answers were significant for the SB and MM items, and the number of incomplete items also differed significantly on the SB items.

The results for errors and other phenomena are quite different from those on Task 4, and there seems to be no discernable pattern to these differences, as seen on Table 11. Comparing the error distribution of the groups to that of the Controls on Task 4, they perform similarly as regards the number of semantic errors, fewer syntactic errors overall (although the Bilinguals'

Table 11: Task 5, Errors, Corrections, Dysfluencies, Substitutions, Additions and Omissions

	cats.	sem.	syn.*	phon.	corr.*	corr. rate	dys.*	subst.	add.*	omit
Bil.	N	1	13	9	9	39%	58	9	24	10
	SB	-	11	8	5	26%	78	-	20	3
	MM	1	10	6	5	29%	59	3	19	4
	StA	2	14	11	6	22%	55	1	31	7
	tot.	4	48	34	25	(29%)	250*	13	94	24
SC	N	2	8	8	6	33%	37	6	26	7
	SB	-	11	7	12	67%	59	4	24	8
	MM	2	7	9	9	50%	43	3	16	1
	StA	-	10	14	7	29%	50	1	24	6
	tot.	4	36	38	34	(44%)	189	14	90	22
FC	N	-	7	5	4	33%	19	1	26	3
	SB	1	1	11	5	38%	37	5	45	2
	MM	-	2	9	2	18%	45	1	34	2
	StA	-	6	6	2	16%	32	1	27	5
	tot.	1	16*	31	13*	(27%)	133*	8	132*	12
10+	N	-	8	7	4	27%	51	3	37	4
	SB	3	8	7	3	16%	48	1	32	3
	MM	-	6	9	5	33%	58	2	38	3
	StA	1	7	8	4	25%	45	1	37	3
	tot.	4	29*	31	16*	(25%)	202	7	144*	14

score is close), and more phonological errors. To what extent the differences in syntactic⁵ and phonological errors are due to the nature of the French language can unfortunately not be determined.

Overall, there are fewer significant categories than on Task 4. Those that are not significant are the phonological errors, substitutions, and omissions. The categories with significant differences are syntactic errors, corrections, dysfluencies, and additions. The extreme

⁵ Even though the verbal inflections for person and number are spelled differently in French, many of them sound the same. Thus it is less likely that verbal inflection errors will be made.

decrease in the number of phonological errors (the score of the FC group fell approximately 60%) and omissions (the smallest decrease, that of the Bilinguals, is approximately 50%), and the increase in the number of substitutions, strongly suggests that these Ss are not balanced bilinguals. The exceptions, namely the Bilinguals' performing four more phonological errors on this task, the small decrease in those of the SC group, and the FC group's decrease (by one) in the number of substitutions cannot mitigate the more dramatic changes of the other groups.

The most striking result is the change in performance by the FC group. On Tasks 2, 3, and 4, this group consistently produced the most errors. On this task, they perform the fewest. Nevertheless, the group's scores on syntactic errors and corrections are still significant. For the syntactic error category, the FC group's low score was significant with respect to all other groups, and the 10+ group's score was also significant with respect to the Bilingual group. On corrections, the FC group's low score is significant with respect to the Bilingual and SC groups, while that of the 10+ group is significant with respect to the SC group only.

When the scores of the groups were compared statistically against the Task 4 scores, all four groups performed significantly differently in the number of syntactic errors committed. On phonological errors, the FC and 10+ groups differed significantly between tasks, while only the FC group's score was significantly different for corrections. Nonetheless, the SC group corrected approximately 40% more, while the 10+ group corrected only one more of their errors on this task than on Task 4. The number of corrections performed by the FC group fell by approximately 50%.

The correction rate has also increased with respect to both the groups' own scores and those of the Controls on Task 4. The average correction rate climbs to 31%, from 13% on Task

4. Correction rates vary from 25% to 44%, whereas on Task 4 they were in the teens. The number of corrections performed, however, is not appreciably different — the correction rate has climbed only because fewer errors were committed.

The results for the dysfluencies category show that the FC and Bilingual groups differ significantly from all the other groups. For additions, the FC and 10+ groups' scores were significant with respect to those of the Bilingual and the SC groups, or vice versa. Overall, the groups added more and omitted less compared with both their own scores and the Controls' scores on Task 4. Once again, the 10+ group was the most creative, performing 31% of all additions on this task. The change in the number of omissions performed is outstanding. The Bilinguals and the SC group performed half as many as the Controls did on Task 4, the FC and 10+ groups performed just over 10% as many.

The consistently greater number of additions made by the Francophone groups may be the result of French syntax — it is a commonplace that it takes more words to say something in French than in English, as is often noted with translated text. This same tendency, however, may have helped keep the dysfluency counts high. This is also seen when the Task 4 and Task 5 scores are compared, as only the FC group performs significantly differently on dysfluencies between the two tasks. For substitutions, only the SC and 10+ groups differ significantly, while on additions and omissions, all groups do.

The type-token data seem still less conclusive than they were for Task 4. When comparing these data (Table 12) with the latencies for this task, the SB category has a cluster of high token scores and low type scores, and thus all the highest differences scores. This corresponds to the number of incomplete items, although the number of null answers remains

Table 12: Task 5, Type-Token Data

	Bil.	SC	FC	10+	Avg.
N					
token	885	911	864	881	885
type	546	565	533	568	553
diff.	339	346	331	313	332
SB					
token	878	902	942	905	907
type	480	498	544	522	511
diff.	398	404	398	383	396
MM					
token	869	835	912	912	882
type	535	513	550	555	538
diff.	334	322	362	357	344
StA					
token	886	893	842	898	880
type	490	526	503	538	514
diff	396	367	339	360	366
Avg.					
token	880	885	890	899	
type	513	526	533	546	
diff.	367	360	358	353	

fairly constant on all categories except N. The SC group has the highest differences for both Tasks 4 and 5 on this category, as well as the group's lowest type count; the FC group has its highest token count for both tasks on this category, and has its biggest difference, but the scores of these groups are otherwise matched by other groups on other categories.

As on Task 4, the high type counts are clustered in the N category, although those of the MM category are almost as high. Due to fewer incomplete and null answers, and more additions and omissions, all counts are generally higher than on Task 4. They are also higher than those

of the Controls on Task 4. To what extent this may be due to the change in language, as opposed to the Ss speaking their L1 instead of their L2, is unfortunately not known. The differences between type and token scores are also higher, but this may reflect French vocabulary, or merely be a side effect from the greater number of sentences completed and words used. For example, if Ss averaged 2 repetitions per utterance, the difference, if there were no null answers, would be 300. On Task 4, the groups surpassed this number only on the SB category. On Task 5, all the groups had differences of over 300.

Comparing the groups against each other, the group with the overall highest counts for both types and tokens is the 10+ group, as on Task 4, despite having the highest number of null answers. This is due partly to additions, which lengthened the utterances performed, but creatively, with relatively little repetition. The Bilinguals have the lowest type and token counts, also as on Task 4, and the most incomplete utterances.

Chapter Five

Discussion and Conclusions

5.1 Discussion

When presenting the working and tested hypotheses (Section 3.1), it was proposed that: 1) the Bilinguals' performance would be similar to the Controls; (2) that older learners' (the SC, FC and 10+ groups) performance would be more negatively affected by metalinguistic interference than the Bilinguals'; and (3), that those whose initial and most significant contact with the L2 was in a formal (school) context — the FC group — would have more problems with L2 speech production than those who learned in an informal (street) context. While some of the results seem to partially support these hypotheses, there are many problems.

Hypothesis 1 is only partially supported. It is true that overall the Bilinguals did perform more like the Controls than any other group. Yet they consistently performed more phonological errors on Tasks 3 and 4, they were consistently more dysfluent than the Controls, and there were differences on other categories as well. Some of these differences were significant, counter to tested hypothesis 1. Moreover the differences in response time between Tasks 4 and 5, even if not significant, as well as between null and incomplete answers, errors and other phenomena also indicate that these Ss, as well as all the others, were not balanced bilinguals (Sections 3.1 and 3.2.2).

The lack of balancedness can be ascribed in part to the fact that the Francophone Ss received their primary and secondary education in French, and spoke French at home. Exposure to English in a formal, school context and the nature of the experiment may have induced

sufficient focus on form to partly counterbalance the effect of early acquisition. It is also possible that the conditions of the experiment revealed language dominance that was not evident during the initial interview and subject evaluation session. The experiment, then, is not as complete a test of the model as was intended, although the data found are still useful.

Hypothesis 2, that older learners' performance would suffer from metalinguistic interference, also receives some support. The SC and FC groups usually produced more errors, corrections, dysfluencies, etc., consistent with metalinguistic processing. This can be seen even on Task 2, when the Ss had been told that changes and corrections were not required.

The performance of the 10+ group, however, does not follow the model's predictions well. There are many phonological errors and omissions on Task 3, for example, but they are not all the highest scores. On Task 2, many of this group's scores were similar to those of the Controls, more so than those of the SC and FC groups. This would seem to indicate that use of the pseudo-linguistic processor does not increase with age. The lack of monitoring and metalinguistic interference is evident in the number of dysfluencies and low correction rate, and many substitutions and additions. Only on Task 4, with its heavy processing load, did the 10+ group perform more like the SC and FC groups.

Tested Hypothesis 3, that Ss who learned their L2 in a formal (school) context would have more errors and dysfluencies than those who learned in a street context, was the most clearly supported of all the hypotheses. The performance of the 10+ group is relevant here, as the Ss in this group had little or no formal exposure to English, and thus they likely had less metalinguistic interference than the other two groups of older learners. The FC group consistently responded more slowly, usually committed more errors and corrected more than the

other groups, as predicted by the model. Although it was proposed that the errors would be more varied, the data was not analyzed in enough detail to show such differences.

Other experimental evidence that is relevant to the PAV model is Postma et al.'s (1990) study on the relations between dysfluency, error, and repairs — even though it is L1 data. The experiment involved reading normal sentences and tongue twisters under several different conditions — time-pressure and accuracy being the categories relevant to this discussion. They found that when time-pressure was applied, 57% of the sentences were completed in the low accuracy condition, and only 51% in the high accuracy condition. These numbers are comparable to the completion rates on Task 2. They also found that the distribution of dysfluency counts resembled the repair counts rather than the error counts, and propose that dysfluency is associated with monitoring and repair. This conclusion is also shared by other researchers (Sharwood-Smith 1991; Snow et al. 1991).

They also claim that the high accuracy condition has relatively low error rates. Although the Ss were more accurate, they were also more dysfluent, and corrected more. This corresponds, I believe, to the performance of the SC group on Tasks 3 and 4. While they did not perform as many errors as the FC group, they were highly dysfluent, had high correction rates, and had low numbers of incomplete and null answers on Task 3, and few incomplete items on Task 4. The SC group might seem, then, to have more focus on form than the FC's. However, according to Neufeld's model, the difference in acquisition context might account for this discrepancy. The SC group, although equally focused on form, experiences less metalinguistic interference, not only committing fewer errors, but also performing greater numbers of successful repairs.

In an experiment designed to test Neufeld's PAV model, Xu (1992) focused upon the effect that the introduction of structural linguistic knowledge would have upon the ability of adult Ss to mimic Chinese phonemes and tone. As predicted by Neufeld's model, the Ss could reproduce tone and the new phonemes accurately, some even approaching native performance. Also as predicted by the model, when the Ss used their lexical and syntactic knowledge to create utterances, articulatory and tonal quality deteriorated, and English intonation contours were transferred. Sentence type (question, negation) and length (complexity) also negatively affected production.

5.2 Problems with Experimental Methodology and Design

The experiments discussed above not only offer more support for the PAV model, they also indicate areas in which the experiment could be improved. One of the more important problems was that the Ss were not 'balanced' bilinguals according to the definition given, and so the experiment was not as rigorous a test of the model as was desired. Choosing Ss who have had a more 'balanced' exposure to English might correct this problem, although it also might result in a shift from French- to English-dominant bilingual Ss. Another possibility is to add a separate measure of bilingual proficiency and require that Ss be 'balanced' on that task. A task of a more vernacular nature, such as story-(re)telling (Raupach 1984; Flanigan 1991), or taping and assessing the initial interview and subject evaluation might suffice.

Another unusual result was the lack of significance of RT variation. Although this is due, in part, to the small number of Ss, within each group and overall, transcription could have been more rigorous. Non-linguistic fillers, especially at the beginning of an utterance, could account

in part for the unusually low RTs on Tasks 3 and 4, as could utterance-initial extensions, or items read as is on Task 3. These phenomena can and sometimes were deliberately used as by Ss to keep from committing too many incomplete and null items.

Further analysis of the data should include measuring and categorizing the errors and other phenomena according to location. A count should also be kept of the items on Task 1 that the Ss did not respond to in time. This would indicate long RT items which could receive more detailed follow-up on the succeeding tasks. Another aspect that could be explored further is that of whether or not the Ss learned how to respond to the stimuli, especially for Tasks 4 and 5, where there were 30 items per category. If the Ss used their metalinguistic knowledge to create a response strategy, it may help account for the low response times.

A more difficult problem to solve is the failure of the syntactically 'bad' items (on Task 3) to induce metalinguistic processing, as they were so easy to correct. Ss only had to invert the order of the objects to produce a good utterance. This is not a great problem, but it is currently difficult to evaluate the relative effect of semantic versus syntactic processing on the changes involving substitution (for Tasks 2 and 3) or change of lexical item (Tasks 4 and 5). Syntactic structures with differing ranges of use in French and English could be targeted (Bates and Devescovi 1984). Problem structures revealed by traditional grammar tests could be used, as could discovering what structures bilinguals believe they have problems with, and are thus likely to pay more (metalinguistic) attention to (Hulstijn and Hulstijn 1984; Green and Hecht 1992). However the problem of reliably inducing a semantically-based focus on form will remain.

It is possible that evidence of syntactic processing could be found by analyzing further the data already collected. An in-depth examination of the sentence structures created — tense

changes, embedded clauses, added modifiers and internal arguments, etc. — and their correlation, if any, with error and dysfluency could be of use. A more traditional analysis of the data, counting anticipations, perseverations, etc., and identifying the source, target, and processing level (semantic, syntactic) of the errors could also contribute to a more complete model of L2 speech production processing and metalinguistic interference.

Detailed analysis might also indicate whether or not the FC group, or any of the older learners, performed more different varieties of speech errors than did the Controls or Bilinguals, although my impression, when counting errors, was that this was not the case. On Tasks 2 and 3, the sentence structures were provided, the sentences could be quite short on Tasks 4 and 5, and all the structures were very simple. The findings of Wiese (1984) and Berg (1987) — that L2 speech differs in the number of errors made rather than the types of errors performed — agree with my impressions rather than supporting the hypothesis on this point.

Some means of counting words per utterance might also be useful, as the type and token counts do not take null and incomplete answers in to account very well. A few Ss uttered very long sentences when they got the chance, as if to make up for their null and incomplete answers, but I currently have no way of showing such occurrences. However, in order to draw conclusions of any theoretical scope, data from more Ss is required.

5.3 Conclusions

The effect of metalinguistic information on bilingual speech processing is a relatively new area of research. Previously, the occurrence of foreign accent was attributed to a lack of competence in the second language, due to factors such as cognitive or neurological maturation.

Yet more recent research has shown that not only are adults able to produce L2 speech without foreign accent, but that native-like competence does not guarantee accentless speech.

Existing models of speech production, based on L1 speech error data, barely acknowledge that metalinguistic input can affect speech production. L2 speech production is not considered to be significantly different from L1 speech production, and so the models are quite similar. Discussion centres upon the requirements of processing more than one set of lexical, syntactic and phonological data. While the role of metalinguistic data is assigned somewhat more prominence in these models than it is with the L1-based models, researchers have but little data upon which to base their proposals.

The model that is the object of this thesis is Neufeld's Pre- and Post-Articulatory Verification (PAV) model. Neufeld proposes that when metalinguistic knowledge — the information we have about a language, both conscious and unconscious — is used during speech processing, this extra input can overload the processor. When the processor is overloaded, it slows down, and further processing may also be affected. Neufeld proposes that because phonological processing is less integral to message comprehension, it is this level, and its pre- and post-articulatory verification, that is most negatively affected by metalinguistically-induced overload, which results in foreign accent. He also claims that both older learners (over age 5) and those who learn an L2 in a formal (school) context are more likely to focus on form when speaking, and thus are more prone to have foreign accented speech.

Twenty Francophone and 5 Anglophone Ss performed an experiment designed to test Neufeld's model. All groups contained 5 Ss, and were divided as follows: Anglophone Controls; simultaneous Bilinguals (learned before age 5); two groups learners who began acquiring English

between age 5 and 10, one group having informal (SC) exposure to English, the others having learned it in school only (FC); and a group who began to learn English after age 10 (10+).

Focus on form was induced by including stimulus items that were morphologically malformed (MM), anomalous and bizarre both semantically (SA and SB) and syntactically (StA and StB), and also some normal (N) items. Ss were also asked to create grammatical and sensible utterances as quickly as possible. The five tasks were: Task 1, a reading and sentence evaluation task; Task 2, on which oral reading speed was timed; on Task 3, Ss read aloud the 'good' sentences, and uttered corrected versions of the 'bad' items; on Task 4, Ss used three keywords to create sentences, and Task 5 was a French version of Task 4. Speech error, dysfluency, and latency of response were among the measures used.

The data support the thesis that foreign accent is affected by performance factors. The task with the smallest processing load, Task 2, had the lowest numbers of dysfluencies and errors, and highest correction rates. Dysfluencies and errors increased and correction rates decreased on the tasks with higher processing loads, Tasks 3 and 4.

Neufeld's PAV model is also supported, with qualifications. The Ss did not meet the criteria set for balanced bilinguals, so the experiment was not a complete test of the model. Nevertheless the Ss were highly proficient L2 speakers. The highest number of syntactic errors performed by a group on any category was 48, by the 10+ group on the MM category on Task 4. On average, then, the 10+ Ss performed a syntactic error on only one utterance in three. Moreover, the lack of balancedness may be due to S selection or the experimental context.

Support is weakest for the theory that older learners have more focus on form than younger learners. Although both the SC and FC groups had more dysfluent and error-full speech

than the Controls or Bilinguals, the 10+ group was not more adversely affected. Moreover the Bilinguals did not perform as much like the Controls as predicted by the model.

Context of acquisition — school versus street — had a much more visible effect on performance. The SC and 10+ groups, older learners who acquired English primarily in informal settings, usually committed fewer errors and more dysfluencies than the FC group did. The 10+ group had the least formal exposure to English, and tended to correct much less than both the SC and FC groups. While the number of dysfluencies would seem to indicate more focus on form than attested to by their correction rate (Postma et al. 1990), this group was also the most creative, consistently performing many additions and substitutions, and this may have affected their fluency. The result of the focus on form for the SC group was a controlled number of errors, and metalinguistic knowledge that did not interfere with correction, although it may have impaired fluency. The FC group, despite their focus on form (generally high dysfluency and high numbers of corrections) still performed the most errors on all tasks, Task 5 excepted. In their case, metalinguistic knowledge seemed to interfere with both utterance creation and repair.

Although the data demonstrate the negative effects of metalinguistic knowledge, there is also considerable evidence that L2 learners who are metalinguistically aware tend to be better learners (Hulstijn and Hulstijn 1984; Green and Hecht 1991; Snow et al. 1991). It must also be noted that metalinguistic knowledge also affects L1 speech production, as shown by the increase in the number of errors performed by the Controls on Tasks 2, 3 and 4, and the performance of the Francophone groups on Task 5.

I believe that data should be considered as support for the processors effects on L2 speech production, and that this does not weaken Neufeld's PAV model. This experiment has proved

to be a useful addition to the study of speech production processing, and the richness of the data elicited may contribute yet more material for investigation for years to come.

APPENDIX A

Task One

(Note: All items in Task One were responded to: Ss pressed a button to indicate whether the sentence was acceptable or not. The items preceded by an asterisk were used as filler items in Task Three. Typographical errors are underlined.)

Sample Items	(N)	Jason backed the car into a tree.
	(MM)	The kids swung on the tire in the tree.
	(SB)	The wind picked the blooms off the bush.
	(StA)	Sally drained in the pot the pasta.
	(SA)	Decoys copy ducks for hunters.
	(MM)	Alan stucked a bow on the present.
	(StB)	The tractor dragged from the mud the car.
	(SA)	Woodpeckers hammer holes in trees.
	(StB)	The thief threatened with a knife the man.
	(N)	Elsa pricked her finger while sewing.
	(StA)	The plane landed on time at Paris.
	(SB)	The storm leaned the tree over the stream.
Normal	*	The man begged for money in the street.
	*	Martina used bleach on the stains.
		The dog licked the cat on its paw.
	*	Tim nibbled the candy in the dish.
	*	Justin composes songs in minutes.
		The guard stopped cars at the border.
	*	Andrew meets his friends at the mall.
	*	Ephrem brought gifts to the children.
		Jan types letters on the computer.
	*	Linda put the caller on hold.
	*	St. Nick filled the stocking with candy.
		The sun melted the ice in the glass.
	*	Eric stacks the dishes in the sink.
	*	Stephanie loaded the gun with blanks.
		Don comforted Dana with a hug.
	*	Ambrose cheered up Martin with jokes.
	*	Nora piles laundry in the hamper.
		The drill bored some holes in the wall.
	*	Dermot offered Monique a job.
	*	The Jones' joined the Smiths for dinner.
	Jed spoke to Martha for a moment.	
*	Jack tied the toboggan to his car.	

**Semantically
Anomalous**

- * Bea dried the child's tears on her blouse.
- * The girl snatched the doll from the dog.
- * The clerk billed the firm for service.
- * Jezebel rides her bike to school.
- * The boy grabbed the ball on the shelf.
- * Maureen opened the door for the men.
- * Allen painted the walls in that house.
- * Simon pushed the cat out of his lap.
- * The turnip told the kids a story.
- * The horses bought houses for themselves.
- * The glass stroked the dog on the head.
- * The pizza paid the clerk for the books.
- * Cashews sign autographs for their fans.
- * The cube punched Colin in the nose.
- * The desk stared at a tree in the park.
- * The cat barked at the dog in the street.
- * The eggs forgot the wine in the fridge.
- * The tree desired the bird out of the nest.
- * Sharks charge divers with trespassing.
- * The ball returned the shoes to the store.
- * The water grabbed the house on the shore.
- * The germs built the house by themselves.
- * Nails do their shopping on Fridays.
- * The grass drank the milk in the saucer.
- * The cup asked the teacher a question
- * The pens searched the desk for paper.
- * Some corn rolled the log over the stone.
- * A tape chased the raccoon out of doors.
- * A mouse designed the house on the hill.
- * The garbage baited its hook with meat.
- * The desk covered its face with its hands.
- * The shovel loaned its neighbour a hose.
- * The axe drew circles on some paper.
- * Ants enjoy the weather in Spain.
- * A grape bet four dollars on the horse.
- * The sauce flew from France to New York.
- * The rope squeaked the chalk on the blackboard.
- * A knife dreamed in its bed by the door.
- * Bees build themselves houses out of wax.
- * The puck scored the goal for the team.
- * The cat locked Fay out of the house.
- * The sun returned to the sky on time.
- * The dog put the squirrel in the tree.
- * The bag trapped the fly inside itself.

**Semantically
Bizarre**

**Morphologically
Malformed**

- * The truck drinks gallons of gas each day.
- * The horse paid its owners by winning.
- * A beaver closed the road with a tree.
- * The ocean left the boat on the shore.
- * Plants suck water out of the ground.
- * Some tree killed a man with a branch.
- * The clouds dropped snow on the city.
- * The reef wrecks the boats far from the shore.
- * The chain pulled the plug from the bathtub.
- * The mouse chased the cat out of the room.
- * The spider wrapped the fly in a web.
- * The sun allowed the moon an eclipse.
- * The sun colours the sky at sunset.
- * The sheep answered her lamb with a bleat.
- * The fish planned the run on Friday.
- * The horse bothered the flies with its tail.
- * Nuts leap ffom (from) the tree to the ground.
- * The sheep hoped for shearing in the fall.
- * The waves flirted with plants on the shore.
- * Goats ask for attention from farmers.
- * Trains insist on leaving on time.
- * Birds compose songs before singing.
- * Cows create milk out of their feed.
- * The plant threw its seeds on the ground.
- * Jim forgotteed the soup on the stove.
- * Jason rided the horse with the spots.
- * Mark beated the dust from the carpet.
- * Don Juan killen Carmen with a knife
- * Jess winned the race for her mother.
- * The frog caughted a fly with its tongue.
- * Edith tores a hole in her nylons.
- * Amanda taked the subway home.
- * Dan holded the apple in his hand.
- * David teached the children music.
- * I lits the candles with a match.
- * Bob maded pancakes for breakfast.
- * Eileen readeed stories to the kids.
- * The horses runned the race without riders.
- * Ed haves pictures in museums
- * Rod doved from the board into the pool.
- * Doreen buyed her brother a present.
- * Natalie gotted clothes for Christmas.
- * Gord hitted the wall with his foot.
- * Frances sticked the stamps on the box.

**Syntactically
Anomalous**

- Margot flew planes for a living.
 * Patrick telled lies to his parents.
 * Sean foughted the champ for the title.
 Tyrone repaided Jane for dinner.
 * Marise hungs ornaments on the tree.
 * Gene growed vegetables in the garden.
 Gerrard ringed the bell for the church.
 * Zach shutted the windows on his hand.
 * James pinnen the note on the corkboard.
 Nell satten on the couch in the hall.
 * He begged for mercy the audience.
 * Jean put on the heat the saucepan.
 His grandmother gave the ring him.
 * Consultants charge for their work a lot.
 * Henry paid for his aunt the bill.
 Burt crossed with roses carnations.
 * The coaches compared to athletes dancers.
 * The rector thanked for coming the crowd.
 Jack blames for the accident Jill.
 * Faust signed with his blood the contract.
 * The chef stirred in the pot the soup.
 Kay suspects of murder the butler.
 * The cop guarded of the crime the scene.
 * Doug winked on the corner at the girl.
 Jan told about the party her friends.
 * Renee melts on apple pie cheese.
 * Mona forgot of her friend the name.
 Danny pulled on his hands mittens.
 * Moira donates to shelters food.
 * Maureen glued to the ring the stone.
 Jessica read on the wall the sign.
 * The parents punished for lying their kids.
 * Vito broke while boxing his nose.
 Lee drew of the mountain a picture.
 * Anne designs for a living clothing.
 * Veronica kissed on the nose Dave.
 Ted accused of theft Valerie.
 * Tom thought on vacation of his friend.
 * The man reminded of someone Kay.
 Hal knows with long hair a student.
 * Barry goes at eleven to bed.
 * Jane handed to him the cup.
 Librarians put on shelves books.
 * Clark lay on the floor the carpet.

**Syntactically
Bizarre**

* Graham held on course the sailboat.
The police charged with murder the maid.
* Sarah scratched with her foot her knee.
* The baby ate for breakfast strained pears.
Diana cooked for dinner some fish.
* Laurel placed on the table the vase.
* Kids rub out with erasers mistakes.
Bess cut into pieces the paper.
* Erika washed in the tub the clothes.
* Sandra wrapped for her kids the gifts.
Horst looked up in the book the number.
* The thieves tied to a chair Penny.
* Hans plugged with his finger the hole.
Gilbert stuck in the socket the plug.
* Brian drilled in the wall a hole.
* Michael saved from fire his family.
Robin changed on her car the tires.
* The cook stirred into the soup the cream.
* Victor scraped on the mat his boots.
Murna soaked in the sink the towels.
* Bruce hurt playing football his wrist.
* Audra offered to Laura a job.
Garth concealed in his sleeve the card.
* Sue smothered in dressing the salad.
* The cow licked on the shoulder her calf.
Bill ignored on the sign the message.

APPENDIX B

Task Two

(Note: Items to be read aloud marked with an asterisk.)

Sample Items

- Hospitals test all blood for Aids.
- Fred cooked some chicken for dinner.
- * Nancy scrubbed the ring from the bathtub.
- Martha scratched the finish on her car.
- Georgette hurt her foot on the stairs.
- * Aladdin rubbed the lamp with his hand.
- Gardeners plant bulbs in the fall.
- Ross found the end of the rainbow.
- * Mark sharpened the pencil with a knife.
- The teacher praised the student's report.
- Joseph waved goodbye to his mother.
- * Chloe owns a business in Toronto.

Test Items

- Women carry baskets on their heads.
- Police consult the psychic on crimes.
- * Benny left the keys in the lock.
- The wind pushed the plant off the ledge.
- Karl basted the turkey with butter.
- * Bertha tucked the sheet under the mattress.
- Tim shovelled the car out of the snow.
- Albert treated the kids to some cake.
- * Vera loaned Cameron the money.
- The wind blew the wash off the line.
- Carol wrote her mother a letter.
- * Christopher fed the lions their meat.
- Caroline gave the baby its milk.
- Phil hates the song on the radio.
- * Irene slid the note under the door.
- Paula chipped the polish on her nails.
- The movers lay the rug on the floor.
- * Francine drives Josephine to work.
- John puts fruit on his cereal.
- Paul held out his arm for the doctor.
- * Karl surprised Angela in the park.
- Students protested the war last week.
- Jocelyne saw the bird in the tree.
- * Brenda packed the boxes with toys.
- Ralph stuffed the letter in the mailbag.

- Teresa orders drinks with ice.
- * The bird pecked at the seeds on the ground.
Barb knits sweaters for her children.
Roger brushes his hair with a comb.
- * Georgia mixed the filling for the pies.
Burt combed his hair with his fingers.
Don washes the dishes after dinner.
- * Lauren pinned the bows in her hair.
Mary cut the cloth into pieces.
Dave drank the beer from the bottle.
- * Karen fixed the sink with a wrench.
The waiter served the cakes for breakfast.
Beth gathered flowers for her mother.
- * Gladys grows carrots in her garden.
Jeffry called his parents on the phone.
Carol teased the cat with a feather.
- * Tina takes vitamins every day.
Bernadette stroked the cat with her foot.
Alda printed her name on the form.
- * The fire man saved the cat in the tree.
The cat stared at the bird in the tree.
Joe stapled the poster to the wall.
- * The pitcher threw the ball to third base.
John bought himself an apple pie.
Cary owns a car with a black top.
- * Grandmother baked cookies for us.
Cars stop at the light on the corner.
Todd shut the door on his finger.
- * Jake went to the movies with a friend.
Barb stepped on some gum on the sidewalk.
The horse enjoys the oats in its feed.
- * Charles protected the dog from the cat.
The kids broke the vase with the toy plane.
The dog barked at the sheep on the hill.
- * Liz prevented the boat from tipping.
The office received a load of pens.
Some people eat hot dogs with ketchup.
- * The spider caught a fly in its web.
The women heard the knock at the door.
The genii refused the man his wish.
- * I like coconut from Jamaica.
Nell showed the pictures to her sister.
The judge fined Josh fifty dollars.
- * Jan got a sunburn on her back.

- The teacher excused students from class.
Babe Ruth hit the ball with the bat.
- * They poured the wine into the decanter.
The guide offered the tour for free.
Matthew polished the wood with oil.
- * The cop found the parent of the boy.
Joanne patted the dog on the head.
The waiter handed Elise the bill.
- * Larry frightened the kids with noises.
Pete returned the book to the store.
People blamed the blackout on the storm.
- * Dot looked under the bed for her shoes.
Faith supplied the wine for the party.
Arthur pulled the sword from the stone.
- * Morris mows the lawn on Sundays.
Raymond starts his new job on Monday.
Margaret trimmed the fat off the meat.
- * The wolf cornered the sheep in the pen.
The fans cheered their team to victory.
George paints pictures of animals.
- * Mark dried his hands on his shirt.
Bernard earns fifteen dollars an hour.
The dog chased the cat into the room.
- * Jenny finishes work at five.
The Queen shakes hands with many people.
Kathy decorates cakes for weddings.
- * Watchdogs growl at strangers at the door.
Cooks fill doughnuts with cream or jam.
The plane flies from New York to Paris.
- * Becky relied on her son for help.
The dog lead Jeremy to safety.

APPENDIX C

Task Three Items

(Note: This is a list of the cued items, the items responded to, only. For the filler sentences, see Appendix A. Typographical errors are underlined, and the intended word follows in parentheses.)

Sample Items	(N)	Tony wiped the dirt off the car.
	(N)	Nicole mixed the batter for the cake.
	(SA)	The pin threw the peels in the garbage.
	(StB)	The guide showed the tourists the city.
	(MM)	Ben cutten his finger on the page.
	(SB)	Birds hold funerals for their friends.
	(N)	The cat licked the milk off the floor.
	(MM)	Natalie woren socks to bed.
	(StB)	Victor wet in the puddle his feet.
	(SB)	The sheep followed the shepherd home.
	(SA)	The hat drowned itself in the river.
(StA)	Amy sprained while jogging her ankle.	
Normal		The cat licked the cut on its tail.
		The cops stopped the bus at the station.
		Ken writes letters on his typewriter.
		The heat melted the cheese in the dish.
		Mother comforted Jean with a kiss.
		The bit bored a hole through the rock.
		Janet spoke to Bob for an hour.
		The men snatched the child from the school.
		The child grabbed the stick on the ground.
		The girl pushed the buggy out the door.
		The jar stroked the cat on the back.
Semantically Anomalous		The hill punched a hole in the rug.
		The cake forgot the bread on the shelf.
		The square returned the can to the shop.
		Horns do their washing on Sundays.
		The blocks searched the floor for circles.
		A rat designed the box in the hall.
		The painter loaned its printer some ink.
		An apple bet ten cents on the mare.
		A shirt dreamed in its chest by the door.
		The dog locked John out of the shed.
		The box trapped the bee inside itself.
Semantically Bizarre		The horse closed the road with a wagon.

**Morphologically
Malformed**

Some wind killed the bord (bird) with a blast.
 The fishline pulled the weed from the mud.
 The net wrapped the fish in its meshes.
 The mice planned the game for Monday.
 The cow hoped for milking each evening.
 Planes insist on landing on time.
 The tree threw its leaves in the air.
 Don beated the eggs with a wire whisk.
 The boy caught a bug in his hand.
 Bob holded the baby in his arms.
 Ed maded supper for his guests.
 Dan haves paintings in his house.
 Valerie gotted toys for Christmas.
 Tyrone flewed kites for the children.
 Margot repaided Bess for the meal.
 Susan ringed the chimes for prayers.
 Anne satten on the sofa outside.
 Her grandfather gave the key her.
 Jack crosses with lemons oranges.
 Burt blames for the damage Olga.
 Jill suspects of assault her boyfriend.
 Kay told about the movies the guys.
 Janet pulled on her head the hat.
 Diana read in the book the notes.
 Ted drew on the building a mural.
 Lee accused of rape Abraham.
 Garth knows from France a professor.
 Secretaries put in stacks files.
 The detective charged with theft the man.
 Jessica cooked for dessert a pie.
 Alan cut into bits the tousers (trousers).
 Ned looked up in the index the name.
 Robin stuck in the hole the stopper.
 Gilbert changed on the bike the wheels.
 Bella soaked in some bleach the clothing.
 Hal concealed with his hand a smile.
 Bess ignored on the tape the warning.

**Syntactically
Anomalous**

**Syntactically
Bizarre**

APPENDIX D

Task Four Items

(Note: Items marked with an asterisk have structures inappropriate to their category.)

Sample Items	(N)	child scribbled paper
	(SO)	fly struggled frog
	(MM)	Betty droved home
	(StA)	man plant collided
	(MM)	boy sangs choir
	(N)	Leslie choked sandwich
	(N)	otters play water
	(SV)	birds sing supper
	(StA)	gunman weapon lay
	(VO)	cat switched radio
	(N)	Bruce stumbled stairs
	(MM)	Santa comed chimney

Normal		bird flew window
		picture hangs wall
		Mary ran plane
		animals live woods
		doctor operated patient
		crickets sing night
		Ellen made story
		cat stares birds
		leaves fell tree
		Cora winks Alfred
		Elmo worried test
		Vincent paid lunch
		child begged present
		Dora rested sofa
		girl asked help
		rider leaped saddle
		car smashed truck
		cat sits fridge
		hunter looked deer
		Derek flirts Anna
		Marion slid ice
		Paul listens music
		cowboy searched water

		<p> baby cries mother Joel pulled socks ships sail ocean child bled cuts Jerry shops market mouse squeezed hole boy whistled teeth blood trickled bandage fish leaped boats wolves howl rabbits chicken pecked blanket duck landed tree fish lives forest rain drips clouds vine climbed rock grapes ripen fridge moon rises clouds dog wound string log swam river frogs glide air horse hurried work sheep curled bed cow snores sleep lizard shook dry tree wilted heat crayons melted sunshine rain blanketed ground bull flirted cow fleas perform circus goat complains farmer spider tied fly sheep pursued shepherd tree threw leaves plant begs water chick burst eggshell fish pawed glass dust seeped cracks cats jumps table father readed baby bird crashen window Frank hanged coat Clara sitted tree boxers fighted gloves winner helded trophy </p>
Semantically (SV)		
Bizarre		
	(VO)	
	(SO)	
Morphologically		
Malformed		

**Syntactically
Anomalous**

Elise getted bed
 stripper teared clothes
 Paul wrotes mother
 Irma choppen tree
 Eddie dranks bottle
 canoe sinked swamp
 Nigel callen help
 girl goed store
 Andrew stoles shop
 terrorist blewed bridge
 cat spattd dog
 dog digged bones
 Steven throwed papers
 Picasso drewed paper
 guests sleeped floor
 relatives hearded accident
 thief broked houses
 Andrew standed hands
 robbers hided police
 Wendy knews party
 roses growed garden
 thread catched button
 chauffeur drivd door
 man banana slipped
 Max restaurant works
 kids bed bounce
 boy clown laughs
 teenager car wished
 Gus sports gambles
 Keith door knocked
 chef wok cooks
 lovers beach met
 driver brakes slammed
 Judy package ripped
 cop gun reached
 artist canvas paints
 Eugene room walked
 priest floor kneeled
 beaver lodge swam
 Drew diner paid
 Bryce himself read
 lion bone chewed
 Sue bargains looks
 lawyer client pleaded

farmers weather grumble
police crowd pushed
explorers jungle hacked
mouse puddle drowned
Ben apartment lives
chicken barn sleeps
cowboys cattle rounded
boy hoped friend
carpenter wood sawed

*

APPENDIX E

Task Five Items

(Note: Items marked with an asterisk have structures inappropriate to their category. All items are in upper case due to inability to use the correct lower case accents.)

Sample Items	(N)	FEMME GRIBOULLAIT ENVELOPE
	(SO)	PAPILLON LUTTAIT CRAPAUD
	(MM)	JEANNE CONDUISSAIT MAISON
	(StA)	GARCON FLEUR HEURTAIT
	(MM)	FILLE JOUENT ORCHESTRE
	(VO)	GINETTE ETRANGLAIT PAIN
	(N)	PHOQUE AMUSE EAU
	(N)	GRENOUILLE CHANTE SOUPER
	(StA)	SOLDAT ARME PLACAIT
	(SV)	LION FERMAIT RADIO
	(N)	GEORGE GLISSAIT ESCALIER
	(MM)	CHARPENTIER DESCENDENT TOIT

Normal		PIGEON VOLE PORTE
		TABLEAU ACCROCHE PORTE
		JOSEE COURT VOITURE
		SOURIS VIT CAVE
		MEDECIN EXAMINAIT PATIENT
		GRENOUILLE CHANTE SOIR
		LISE RACONTE HISTOIRES
		CHIEN REGARDE CHATS
		BRANCHE TOMBE ARBRE
		CATRINE PENSE ROBERT
		JACQUES INQUIETE REUNION
		VICTOR N'AVAIT REPAS
		FILLE RECLAMAIT CADEAU
		DORA ETEND FAUTEUIL
		HOMME DEMANDAIT AIDE
		PASSAGER SAUTAIT TRAIN
		AUTO ECRASAIT BOITE
		OISEAU DORT COMPTOIR
		CHASSEUR CHERCHAIT LIONS
		DANIEL POURSUIT GINETTE
		MARIE GLISSAIT NEIGE
		RICHARD ECOUTE CHANSONS

	CHEVALIERS FOUILLAIENT BUISSONS
	BEBE PLEURE PERE
	JOANNE TIRAIT CHAUSSURES
	BATEAU TRAVERSE MER
	ADULTE SOUFFRAIT BLESSURE
	THOMAS MAGASINAIT BOUTIQUES
	RAT RONGE MUR
	FEMME SIFFLE LEVRES
	CHAIR ABSORBE SANG
Semantically (SV)	POISSONS SAUTAIENT CASSEROLE
Bizarre	LOUPS HURLENT ARBRES
	OURS ARRACHE VETEMENTS
	CANARD ATTERRISSAIT TROU
	POISSON VIT ARBRES
	GRELE TOMBE BRUME
	FLEUR MONTE ROCHER
	FROMAGE FOND FRIGO
	SOLEIL TRAVERSE NUAGES
(VO)	SINGE ENROULAIT CORDE
	PLANCHE SUIVAIT RIVIERE
	ECUREUIL FLOTTE AIR
	CHIEN DEPECHE ECOLE
	OIE ALONGE MATELAS
	TAUREAU RONFLE ENDORMI
	CROCODILE SONGAIT MONDE
	PLANTE EPANOUISSAIT SOLEIL
	FEMME FOND DOUCHE
(SO)	FEUILLES CAMOUFLAIENT TERRE
	VACHE TAQUINAIT TAUREAU
	VENTE PARLAIT FORT
	POULES PLAIGNENT FERMIER
	MOUCHES ATTIRAIENT ARAIGNEES
	MOUTONS CHASSENT CHIENS
	FEUILLES DANCENT VENT
	BUISSONS VEULENT OMBRE
	DINDON RONGE GRAIN
	TORTUE REMUAIT EAUX
	CHANDAIL IGNORE BRISE
Morphologically	CHAT SAUTENT CHAISE
Malformed	MERES LISAIT ENFANT
	COQ POUSSAIENT CRI
	PAUL SERRAIS LIVRES
	CLARA DORMISSAIT CHAISE
	POLICIERS LUTTE BATONS

GAGNANT TENAIS PRIX
 ELISE COUCHENT LIT
 DANSEUSES ENLEVE VETEMENTS
 ROBERT ECRIVE PERE
 VERONIQUE TAILLISSAIT BOIS
 CHALOUPE CHAVIRAIENT LAC
 NORA INVITAIS VOISINS
 GARCONS VA MAGASIN
 JULES VOLAIENT BOUTIQUE
 VOLEURS DECHIRE ARGENT
 CHAT CRACHAIS CHIEN
 TIGRE ENTERRISSAIT OS
 STEPHANE LANCENT PIERRES
 PICASSO DESSINAIS PAPIERS
 INVITEES MANGE SALON
 PARENTS DISCUTE BLESSURE
 OUVRIERS PENETRE MAISON
 ANDRE DANSENT MARIE
 *
 RODRIGUE ECHAPPAIT POLICIERE
 WENDY CONNAISSE JEAN
 ROSE POUSSERENT JARDIN
 FICELLE ACCROCHENT BOUTONS
 CHAUFFEUR ENDOMMAGENT PORTIERE
 VERRE TOMBAIENT PLANCHER
 FEMME BANANE GLISSAIT
 MAX GOUVERNEMENT TRAVAILLE
 ENFANT DIVAN SAUTE
 FILLETTE CLOWN PLEURE
 ADOLESCENT VETEMENTS DESIRAIT
 GUSTAVE COURSE GAGE
 JEANNE FENETRE FRAPPAIT
 CHEF CASSEROLE CUISINE
 AMANTS CHAMBRE RENCONTRENT
 CHAUFFEUR ROUES BRAQUE
 CATRINE COLIS OUVRE
 HOMME FUSIL CHERCHAIT
 ARTISTE VIOLON JOUE
 LUCIE CHAMBRE COURRAIT
 PRETRE ECHELLE GRIMPE
 CASTOR ARBRE DESCEND
 ARTHUR MEUBLES ACHETE
 ROGER JOURNAL LISAIT
 SINGE RIDEAUX MANGE
 *
 SUSANE ACHATS FAISAIT

**Syntactically
 Anomalous**

AVOCAT CLIENT GRONDAIT
FERMIER PLUIE DESIRE
SOLDAT FOULE BOUSCULE
PIONNIERS CABANES CONSTRUISENT
SOURIS EAU TOMBAIT
EDOUARD APPARTEMENT LOUE
POULE GRANGE CAVALE
FEMMES JEU GAGNAIENT
GARCON COPAIN ESPERAIT
AMIS BOIS TAILLAIENT

References

- Appel, G. (1984) "Improving Second Language Production." In Dechert et al. (eds.), 186-210.
- Asher, J. and R. Garcia (1969) "The Optimal Age to Learn a Foreign Language." *Modern Language Journal* 53, 3-12.
- Bates, E. and A. Devescovi (1984) "Crosslinguistic Studies of Sentence Production." In *The Crosslinguistic Study of Sentence Processing*, B. MacWhinney and E. Bates (eds.). Cambridge University Press: New York, 225-253.
- Bailey, N., C. Madden and S. Krashen (1974) "Is there a 'Natural Sequence' in Adult Second Language Learning?" *Language Learning* 24, 235-243.
- Berg, T. (1987) *A Cross-Linguistic Comparison of Slips of the Tongue*. Indiana University Linguistics Club: Bloomington, IN.
- Best, C. (1987) "The Emergence of Cerebral Asymmetries in Early Human Development: A Literature Review and a Neuro-embryological Model." In *Brain Lateralization in Children*, D. Molfese and S. Segalowitz (eds.). New York: The Guilford Press, 5-34.
- Bever, T. (1981) "Normal Acquisition Processes Explain the Critical Period for Language Learning." In *Individual Differences and Universals in Language Learning Aptitude*, K. Diller (ed.). Rowley MA: Newbury House, 176-196.
- Bialystok, E. (ed.) (1991) *Language Processing in Bilingual Children*. New York: Cambridge University Press.
- Bley-Vroman, R. (1989) "What is the Logical Problem of Foreign Language Learning?" In S. Gass and J. Schachter (eds.), 41-68.
- Boomer, D. and J. Laver (1968) "Slips of the Tongue." *British Journal of Disorders of Communication* 3, 2-12.
- Broselow, E. (1988) "Prosodic Phonology and the Acquisition of a Second Language." In S. Flynn and W. O'Neil (eds.), 295-308.
- Broselow, E. (1984) "An Investigation of Transfer in Second Language Phonology." *IRAL* 22, 253-269.
- Broselow, E. (1983) "Non-Obvious Transfer: On Predicting Epenthesis Errors." In *Language Transfer in Language Learning*, S. Gass and L. Selinker (eds.). Rowley MA: Newbury House, 269-280.

- Brown, H. (1981) "Affective Factors in Second Language Learning." In *The Second Language Classroom*, J. Alatis, H. Altman and P. Alatis (eds.). New York: Oxford University Press, 111-129.
- Butterworth, B. (1989) "Lexical Access in Speech Production." In *Lexical Representation and Process*, M. Marslen-Wilson (ed.). MIT Press: Cambridge, 108-135.
- Clahsen, H. (1990) "The Comparative Study of First and Second Language Development." *Studies in Second Language Acquisition* 12, 135-153.
- Clahsen, H. (1984) "The Acquisition of German Word Order: A Test Case for Cognitive Approaches to L2 Development." In *Second Languages: A Cross-Linguistic Perspective*, R. Andersen (ed.). Rowley, MA: Newbury House, 219-242.
- Clahsen, H. and P. Muysken (1989) "The UG Paradox in L2 Acquisition." *Second Language Research* 5, 1-29.
- Clahsen, H. and P. Muysken (1986) "The Availability of Universal Grammar to Adult and Child Learners — A Study of the Acquisition of German Word Order." *Second Language Research* 2, 93-119.
- Courchene, R. and J. Bagheera (1985) "A Theoretical Framework for the Development of Performance Tests." In *Second Language Performance Testing*, P. Hauptman, R. LeBlanc and M. Bingham-Wesche (eds.). University of Ottawa Press: Ottawa, 45-58.
- Crompton, A. (1981) "Syllables and Segments in Speech Production." *Linguistics* 19, 663-716.
- Crookes, G. (1991) "Second Language Speech Production Research: A Methodologically Oriented Review." *Studies in Second Language Acquisition* 13, 113-132.
- Dechert, H. (1984) "Second Language Production: Six Hypotheses." In Dechert et al. (eds.), 211-230.
- Dechert, H., D. Mohle and M. Raupach (eds.) (1984) *Second Language Productions*. Gunter Narr Verlag: Tubingen.
- Dell, G. (1989) "The Retrieval of Phonological Forms in Production Tests of Predictions from a Connectionist Model." In *Lexical Representation and Process*, W. Marslen-Wilson (ed.). The MIT Press: Cambridge, MA, 136-165.
- Dell, G. (1986) "A Spreading-Activation Theory of Retrieval in Sentence Production." *Psychological Review* 93, 283-321.

- Dell, G. and P. Reich (1981) "Stages in Sentence Production: An Analysis of Speech Error Data." *Journal of Verbal Learning and Verbal Behaviour* 20, 611-629.
- Dell, G. and P. Reich (1980) "Toward a Unified Model of Slips of the Tongue." In Fromkin (ed.), 273-286.
- Dulay, H. and M. Burt (1974) "Natural Sequences in Child Second Language Acquisition." *Language Learning* 24, 37-53.
- duPlessis, J., D. Solin, L. Travis and L. White (1987) "UG or not UG, That is the Question: A Reply to Clahsen and Muysken." *Second Language Research* 3, 56-75.
- Eckman, F. (1981) "On the Naturalness of Interlanguage Phonological Rules." *Language Learning* 31, 195-216.
- Eckman, F. (1977) "Markedness and the Contrastive Analysis Hypothesis." *Language Learning* 27, 315-330.
- Felix, S. (1985) "More Evidence on Competing Cognitive Structures." *Journal of Second Language Research* 1, 47-72.
- Flanigan, B. (1991) "Variable Competence and Performance in Child Second Language Acquisition." *Second Language Research* 7, 220-232.
- Flege, J. (1988) "Factors Affecting Degree of Perceived Foreign Accent in English Sentences." *Journal of the Acoustical Society of America* 84, 70-79.
- Flege, J. (1987) "A Critical Period for Learning to Pronounce Foreign Languages?" *Applied Linguistics* 8, 162-177.
- Flege, J. (1981) "The Phonological Basis of Foreign Accent: A Hypothesis." *TESOL Quarterly* 15, 443-455.
- Flynn, S. (1989) "The Role of the Head-Initial/Head-Final Parameter in the Acquisition of English Relative Clauses by Adult Spanish and Japanese Speakers." In S. Gass and J. Schachter (eds.), 89-108.
- Flynn, S. (1987) *A Parameter-Setting Model of L2 Acquisition*. Dordrecht: Kluwer Academic.
- Flynn, S. (1984) "A Universal in L2 Acquisition Based on a PBD Typology." In *Universals of Second Language Acquisition*, F. Eckman, L. Bell and D. Nelson (eds.). Rowley MA: Newbury House, 75-87.

- Flynn, S. and W. O'Neil (eds.) (1988) *Linguistic Theory in Second Language Acquisition*. Dordrecht: Kluwer Academic.
- Fromkin, V. (ed.) (1980) *Errors in Linguistic Performance*. New York: Academic Press.
- Fromkin, V. (ed.) (1973) *Speech Errors as Linguistic Evidence*. The Hague: Mouton.
- Fromkin, V. (1971) "The Non-Anomalous Nature of Anomalous Utterances." *Language* 47, 27-52.
- Fromkin, V. (1968) "Speculations on Performance Models." *Journal of Linguistics* 4, 47-68.
- Gardner, R. and W. Lambert (1972) *Attitudes and Motivation in Second Language Learning*. Rowley, MA: Newbury House.
- Gardner, R. and P. Smythe (1975) "Motivation and Second-Language Acquisition." *CMLR* 31, 218-230.
- Garrett, M. (1984) "The Organization of Processing Structures." In *Biological Perspectives on Language*, D. Caplan, A. Lecours and A. Smith (eds.). Cambridge, MA: MIT Press, 172-193.
- Garrett, M. (1980a) "Levels of Processing in Sentence Production." In *Language Production*, B. Butterworth (ed.). New York: Academic Press, 177-220.
- Garrett, M. (1980b) "The Limits of Accommodation: Arguments for Independent Processing Levels in Sentence Production." In Fromkin (ed.), 263-271.
- Garrett, M. (1976) "Syntactic Processes in Sentence Production." In *New Approaches to Language Mechanisms*, E. Walker and R. Wales (eds.). New York: North Holland, 231-256.
- Garrett, M. (1975) "The Analysis of Sentence Production." In *The Psychology of Learning and Motivation*. New York: Academic Press, 133-177.
- Garrett, M. and M.-L. Kean (1980) "Levels of Representation and the Analysis of Speech." In *Juncture*, M. Aronoff and M.-L. Kean (eds.). Saratoga, CA: Anma Libri, 79-89.
- Gass, S. and J. Schachter (eds.) (1989) *Linguistic Perspectives on Second Language Acquisition*. Cambridge: Cambridge University Press.
- Geffner, D. and I. Hochberg (1971) "Ear Laterality Performance of Children from Low and Middle Socioeconomic Levels in a Verbal Dichotic Listening Task." *Cortex* 7, 193-203.

- Goodluck, H. and B. Birch (1988) "Late-Learned Rules in First and Second Language Acquisition." In J. Pankhurst et al. (eds.), 94-115.
- Green, P. and K. Hecht (1992) "Implicit and Explicit Grammar: An Empirical Study." *Applied Linguistics* 13, 168-184.
- Gregg, K. (1989) "Second Language Acquisition Theory: The Case for a Generative Perspective." In S. Gass and J. Schachter (eds.), 15-40.
- Gregg, K. (1984) "Krashen's Monitor and Occam's Razor." *Applied Linguistics* 5, 79-100.
- Guiora, A., R. Brannon and C. Dull (1972) "Empathy and Second Language Learning." *Language Learning* 22, 111-130.
- Harley, T. (1984) "A Critique of Top-Down Interdependent Models of Speech Production: Evidence from Non-Plan-Internal Speech Errors." *Cognitive Science* 8, 191-219.
- Hecht, B. and R. Mulford (1987) "The Acquisition of a Second Language: Interaction of Transfer and Developmental Factors." In *Interlanguage Phonology*, G. Ioup and S. Weinberger (eds.). Rowley MA: Newbury House, 213-228.
- Hirakawa, M. (1990) "A Study of the L2 Acquisition of English Reflexives." *Second Language Research* 6, 60-85.
- Hulstijn, J. and W. Hulstijn (1984) "Grammatical Errors as a Function of Processing Constraints and Explicit Knowledge." *Language Learning* 34, 23-43.
- Jakobovits, L. (1970) *Foreign Language Learning*. Rowley MA: Newbury House.
- James, C. (1971) "The Exculpation of Contrastive Linguistics." In *Papers in Contrastive Analysis*, G. Nickel (ed.). Cambridge: Cambridge University Press, 53-68.
- Johnson, J. and E. Newport (1989) "Critical Period Effects in Second Language Learning: The Influence of Maturational State on the Acquisition of English as a Second Language." *Cognitive Psychology* 21, 60-99.
- Jordens, P. (1988a) "The Acquisition of Word Order in Dutch and German as L1 and L2." *Second Language Research* 4, 41-65.
- Jordens, P. (1988b) "The Acquisition of Verb Categories and Word Order in Dutch and German: Evidence from First and Second Language Acquisition." In J. Pankhurst et al. (eds.), 132-166.

- Kimura, D. (1963) "Cerebral Dominance and the Perception of Verbal Stimuli." *Canadian Journal of Psychology* 15, 166-171.
- Kinsbourne, M. (1971) "The Minor Cerebral Hemisphere as a Source of Aphasic Speech." *Archives of Neurology* 25, 302-306.
- Knox, C. and D. Kimura (1970) "Cerebral Processing of Nonverbal Sounds in Boys and Girls." *Neuropsychologia* 8, 227-237.
- Krashen, S. (1985) *The Input Hypothesis*. Longman: London.
- Krashen, S. (1982) "Accounting for Child-Adult Differences in Second Language Rate and Attainment." In *Child-Adult Differences in Second Language Acquisition*, S. Krashen, R. Scarcella and M. Long (eds.). Rowley MA: Newbury House, 202-226.
- Krashen, S. (1981) "Aptitude and Attitude in Relation to Second Language Acquisition and Learning." In *Individual Differences and Universals in Language Learning Aptitude*, K. Diller (ed.). Rowley MA: Newbury House, 155-175.
- Krashen, S. (1979) "A Response to McLaughlin, 'The Monitor Model: Some Methodological Considerations.'" *Language Learning* 29, 151-167.
- Krashen, S. (1978) "Individual Variation in the Use of the Monitor." In *Principles of Second Language Learning*, W. Ritchie (ed.). New York: Academic Press, 175-183.
- Krashen, S. (1977) "Some Issues Relating to the Monitor Model." In *On TESOL '77*, H. Brown, C. Yorio and R. Crymes (eds.). Washington: TESOL, 144-158.
- Krashen, S. (1973) "Lateralization, Language Learning, and the Critical Period: Some New Evidence." *Language Learning* 23, 63-74.
- Ladac, R. (1957) *Linguistics Across Cultures*. Ann Arbor, MA: University of Michigan Press.
- Lapointe, S. and G. Dell (1989) "A Synthesis of some Recent Work in Sentence Production." In *Linguistic Structure in Language Processing*, G. Carlson and M. Tanenhaus (eds.). Dordrecht: Kluwer Academic, 107-156.
- Laubstein, A. (1988) *The Nature of the 'Production Grammar' Syllable*. Bloomington: Indiana University Linguistics Club.
- Lenneberg, E. (1967) *Biological Foundations of Language*. New York: John Wiley and Sons.
- Lennon, P. (1991) "Error Elimination and Error Fossilisation: A Study of an Advanced Learner in the L2 Community." *I.T.L.* 93-94, 129-151.

- Lennon, P. (1990) "The Advanced Learner at Large in the L2 Community: Developments in Spoken Performance." *IRAL* 28, 309-324.
- Liceras, J. (1988) "L2 Learnability: Delimiting the Domain of Core Grammar as Distinct from the Marked Periphery." In S. Flynn and W. O'Neil (eds.), 199-224.
- MacKay, D. (1970) "Spoonerisms: The Structure on Errors in the Serial Order of Speech." *Neuropsychologia* 8, 323-350.
- Major, R. (1987a) "A Model for Interlanguage Phonology." In *Interlanguage Phonology*, G. Ioup and S. Weinberger (eds.). Rowley MA: Newbury House, 101-124.
- Major, R. (1987b) "Foreign Accent: Recent Research and Theory." *IRAL* 25, 185-202.
- Malakoff, M. and K. Hakuta (1991) "Translation Skill and Metalinguistic Awareness in Bilinguals." In E. Bialystok (ed.), 141-166.
- Motley, M., B. Baars and C. Camden (1983) "Experimental Verbal Slip Studies: A Review and an Editing Model of Language Encoding." *Communication Monographs* 50, 79-101.
- Neufeld, G. (1988) "Phonological Asymmetry in Second Language Learning and Performance." *Language Learning* 38, 531-559.
- Neufeld, G. (1980) "On the Adult's Ability to Acquire Phonology." *TESOL Quarterly* 25, 285-298.
- Neufeld, G. (1979) "Towards a Theory of Language Learning Ability." *Language Learning* 29, 23-35.
- Neufeld, G. (1978) "On the Acquisition of Prosodic and Articulatory Features in Adult Language Learning." *CMLR* 34, 163-174.
- Neufeld, G. (1977) "Language Learning Ability in Adults: A Study on the Acquisition of Prosodic and Articulatory Features." *Working Papers on Bilingualism* 12, 45-60.
- Oyama, S. (1976) "A Sensitive Period for the Acquisition of Non-Native Phonological Systems." *Journal of Psycholinguistic Research* 5, 261-283.
- Pankhurst, J., M. Sharwood-Smith and P. Van Buren (eds.) (1988) *Learnability and Second Languages*. Dordrecht: Foris.
- Patkowski, M. (1990) "Age and Accent in a Second Language: A Reply to James Emil Flege." *Applied Linguistics* 11, 73-89.

- Penfield, W. and L. Roberts (1959) *Speech and Brain-Mechanisms*. Princeton: Princeton University Press.
- Postma, A., A. Kolk and D.-J. Povel (1990) "On the Relation among Speech Errors, Disfluencies, and Self-Repairs." *Language and Speech* 33, 19-29.
- Raupach, M. (1984) "Formulae in Second Language Speech Production." In Dechert et al. (eds.), 114-137.
- Richards, J. (1971) "A Non-Contrastive Approach to Error Analysis." *English Language Teaching Journal* 25, 204-219.
- Rosansky, E. (1975) "The Critical Period for the Acquisition of Language: Some Cognitive Developmental Considerations." *Working Papers on Bilingualism* 6, 93-102.
- Sato, C. (1984) "Phonological Processes in Second Language Acquisition: Another look at Interlanguage Syllable Structure." *Language Learning* 34-4, 43-57.
- Schachter, J. (1989) "Testing a Proposed Universal." In S. Gass and J. Schachter (eds.), 73-88.
- Schachter, J. (1988) "Second Language Acquisition and its Relationship to Universal Grammar." *Applied Linguistics* 9, 219-235.
- Scovel, T. (1988) *A Time to Speak*. New York: Newbury House.
- Scovel, T. (1969) "Foreign Accents, Language Acquisition and Cerebral Dominance." *Language Learning* 19, 245-253.
- Seliger, H. (1978) "Implications of a Multiple Critical Periods Hypothesis for Second Language Learning." In *Second Language Acquisition Research*, W. Ritchie (ed.). New York: Academic Press, 11-19.
- Selinker, L. (1972) "Interlanguage." *IRAL* 10, 209-231.
- Sharwood-Smith, M. (1991) "Language Modules in Bilingual Processing." In E. Bialystok (ed.), 10-24.
- Sharwood-Smith, M. (1988) "On the Role of Linguistic Theory in Explanations of Second Language Developmental Grammars." In S. Flynn and W. O'Neil (eds.), 173-198.
- Shattuck-Hufnagel, S. (1983) "Sublexical Units and Suprasegmental Structure in Speech Production Planning." In *The Production of Speech*, P. MacNeilage (ed.). New York: Springer Verlag, 109-136.

- Shattuck-Hufnagel, S. (1979) "Speech Errors as Evidence for a Serial-Ordering Mechanism in Sentence Production." In *Sentence Processing*, W. Cooper and E. Walker (eds.). Hillsdale NJ: Lawrence Erlbaum 295-342.
- Shattuck-Hufnagel, S. and D. Klatt (1980) "How Single Phoneme Error Data Rule Out Two Models of Error Generation." In Fromkin (1980), 35-46.
- Snow, C., H. Cancino, J. De Temple and S. Schley (1991) "Giving Formal Definitions: A Linguistic or Metalinguistic Skill?" In E. Bialystok (ed.), 90-112.
- Springer, S. and G. Deutsch (1989) *Left Brain, Right Brain*. New York: W.H. Freeman.
- Sridhar, S. (1989) "Cognitive Structures in Language Production: A Crosslinguistic Study." In *The Crosslinguistic Study of Sentence Processing*, B. MacWhinney and E. Bates (eds.). Cambridge University Press: New York, 209-224.
- Stemberger, J. (1985) "An Interactive Activation Model of Language Production." In *Progress in the Psychology of Language: Speech and Talk*, Vol. 1, A. Ellis (ed.). London: Lawrence Erlbaum, 143-186.
- Stevick, E. (1976) *Memory, Meaning and Method*. Rowley, MA: Newbury House.
- Swain, M. and S. Lapkin (1986) "Immersion French in Secondary Schools: the Goods and the Bads." *Contact* 5, 2-9.
- Tarone, E. (1980) "Communication Strategies, Foreigner Talk, and Repair in Interlanguage." *Language Learning* 30, 417-431.
- Tarone, E. (1978) "The Phonology of Interlanguage." In *Understanding Second and Foreign Language Learning: Issues and Approaches*, J. Richards (ed.). Rowley MA: Newbury House, 15-33.
- Tomaselli, A. and B. Schwartz (1990) "Analyzing the Acquisition Stages of Negation in L2 German: Support for UG in Adult Second Language Acquisition." *Second Language Research* 6, 1-38.
- Wardhaugh, R. (1970) "The Contrastive Analysis Hypothesis." *TESOL Quarterly* 4, 123-130.
- Watson, I. (1991) "Phonological Processing in Two Languages." In Bialystok (ed.), 25-48.
- Wiese, R. (1984) "Language Production in Foreign and Native Languages: Same or Different?" In Dechert et al. (eds.), 11-25.

- White, L. (1990) "Second Language Acquisition and Universal Grammar." *Studies in Second Language Acquisition* 12, 121-133.
- White, L. (1988) "Universal Grammar and Language Transfer." In J. Pankhurst et al. (eds.), 36-60.
- White, L. (1986) "Implications of Parametric Variation for Adult Second Language Acquisition: An Investigation of the Pro-Drop Parameter." In *Experimental Approaches to Second Language Learning*, V. Cook (ed.). Oxford: Pergamon, 55-74.
- White, L. (1985) "The "Pro-Drop" Parameter in Adult Second Language Acquisition." *Language Learning* 35. 47-61.
- Xu, Qiming (1991) *Processing Constraints on Segmental and Suprasegmental Production in a Foreign Language*. Unpublished M.A. thesis, University of Ottawa.
- Zobl, H. (1988) "Configurationality and the Subset Principle: The Acquisition of V1 by Japanese Learners of English." In J. Pankhurst et al. (eds.), 116-131.