INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI®

**Focus on form from the inside: The significance of
grammatical sensitivity for L2 learning in
communicative ESL classrooms**

Leila Ranta

A Thesis in a Special
Individualized
Program

Presented in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy at
Concordia University
Montréal, Québec, Canada

September 1998

Your file Votre référence

Our file Notre référence

0-612-43571-7

Canada

ii

ABSTRACT

Focus on form from the inside: The significance of grammatical sensitivity for L2 learning in communicative ESL classrooms

Leila Ranta, Ph.D.
Concordia University, 1998

Are analytic learners at an advantage in second language (L2) learning? Early research on aptitude (Carroll, 1962) established the significance of grammatical sensitivity when learners are in structure-based instructional settings. This study addressed the question of whether grammatical sensitivity is associated with L2 learning outcomes when instruction is communicative in nature. The participants were five classes of francophone children studying in an intensive ESL program at the grade 6 level. Learners' grammatical sensitivity was defined as 'the ability to focus on form' and was operationalized as the ability to correct errors in written French (L1). A variety of L2 proficiency measures were administered over the course of one academic year. Three types of analysis were performed on the data: a principal components analysis, a cluster analysis, and an interlanguage analysis. The results of these analyses indicate that grammatical sensitivity is associated with success in L2 learning to some degree since:

- a 'Grammatical Sensitivity' factor emerged as one of three factors in a principal components analysis;
- performance on a measure of grammatical sensitivity distinguished strong learners from weak learners;
- higher levels of grammatical sensitivity were associated with higher stages of grammatical development in oral production.

However, it is also clear that grammatical sensitivity is not sufficient to account for differences among weak and strong learners in this population. It is argued that other learner abilities such as working memory capacity and phonemic coding ability may account for the observed patterns of learner performance. Overall, the information-processing model of SLA proposed by Skehan (1998) proved to be a useful framework for the interpretation of the findings of this study.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

One of the characteristics of a communicative approach to second language (L2) teaching is 'learner-centredness'. This is understood to mean that instruction focusses on learners' needs, gives some control to the learner, and enhances learners' sense of competence and self-worth (Brown, 1994, p. 80). In addition, modern methodology textbooks recommend that teachers take individual learning styles into account and select classroom activities so that all students will benefit. While such statements at teachers' meetings may cause heads to nod in agreement, the matter of how to make such decisions is far from clear. For example, most people would believe that the analytical learner who enjoys talking about grammar should be matched with a grammar-based approach while the non-analytical learner should be matched with communication-oriented instruction (see Wesche, 1981 for report on such an approach). And yet, each of these matches may be inappropriate. Politzer (1970) argues that analytical learners may prefer to discover the L2 system themselves rather than have it served up to them on a platter. On the other hand, Skehan (1992) argues that the non-analytical learner may need help in imposing structure on the input and so would benefit from some kind of focus on form. (See Tudor, 1996 for a comprehensive discussion of the meaning of learner-centredness in L2 teaching.)

The teacher who looks to second language acquisition (SLA) research for guidance on how to match different learner types with appropriate learning activities is bound to be disappointed by the lack of hard evidence. R. Ellis (1994, p. 469) states that SLA researchers seeks to answer four questions concerning learner variables:

1. In what way do language learners differ?
2. What effects do these differences have on learning outcomes?
3. How do learner differences affect the process of acquisition?
4. How do individual learner factors interact with instruction in determining learning outcomes?

1

At present, Questions 1 and 2 can be answered with some confidence. Many learner variables have been identified and shown to be correlated with different measures of L2 learning. The difficulty of investigating the last two questions can account for the scarcity of studies addressing such issues. The researcher would need to first specify a model of the process of acquisition, then choose learner variables that are plausibly relevant to the process, and find ways of measuring these variables; the research design should be longitudinal and include the collection of different types of test and language data from a sufficiently large number of learners in contexts where the instructional content and teaching methodology are known. A tall order, indeed. To a large extent, theory and research methodology in SLA were not sufficiently developed to attempt such a program of research in the past.

This study is an attempt at addressing the 'how' question. A particular learner variable, grammatical sensitivity, was chosen on the basis of its relevance to the process of acquisition according to the latest developments in SLA theory. A variety of measures were administered over a period of one year to five classes of francophone learners in a communicative ESL program in an elementary school in the province of Quebec. The overall aim is to relate learner variation in performance in this context to the larger picture of how individual differences fit into a model of SLA.

Plan of the Dissertation

Chapter 1 begins with a general discussion of individual differences research in SLA and an overview of how specific models treat learner traits. This provides the background for the present study. Chapter 2 attempts to clarify what is meant by the ability to focus on form by examining the concepts of aptitude, learner style, and metalinguistic skill. This is followed, in Chapter 3, by a review of the classroom and laboratory studies which have investigated how learners' ability to focus on form (as aptitude, learner style or metalinguistic skill) relates to L2 learning outcomes. The discussion of the salient features

of these studies provides the rationale for the specific research questions and hypotheses addressed by this study. Chapter 4 describes the context of the research, the instruments used to collect language data, and the data analysis procedures. The results of the data analysis are presented in Chapter 5 which is organized around the three research questions. Finally, in Chapter 6, the results of the analyses are interpreted with respect to theory and research presented in earlier chapters. Specifically, the findings are discussed in the light of the information-processing model proposed by Skehan (1998).

CHAPTER 1

INDIVIDUAL DIFFERENCES AND SLA THEORY

This chapter begins with a general discussion of issues relating to the study of individual differences in L2 learning. Several models of SLA are described and the role of individual differences in each model is examined. The chapter ends with a discussion of recent developments in SLA, highlighting the need to include cognitive variables in any model of L2 learning which aspires to explanatory adequacy.

Research on Individual Differences in SLA

It is one of the few universally acknowledged truths in SLA that learners differ in how quickly they learn and in how far they get in their learning of a L2. In attempting to account for variation between learners in rate of learning and in ultimate attainment, different researchers have chosen to focus on different aspects of the many ways that people vary. As a result, the state-of-the-art in this area of L2 research is a bewildering assortment of taxonomies listing traits, attributes, tendencies, and states. A reading of extensive reviews by Larsen-Freeman and Long (1991) and by R. Ellis (1994) gives us the following list of factors that appear to influence L2 learning in some way:

- age
- anxiety (trait, state, and situation-specific)
- aptitude
- attitudes
- awareness
- beliefs about language learning
- cognitive style: field dependence/field independence, broad/ narrow categorizer, reflective/impulsive, analytic/ gestalt
- interest
- learning styles: visual, auditory, kinesthetic, tactile, concrete, analytical, communicative, authority-centered
- memory
- motivation: integrative, instrumental, resultative
- personality: introversion, extroversion, risk-taking, tolerance of ambiguity, empathy, self-esteem, and inhibition
- sex
- personal will

Interpretation of this research is hampered by limitations inherent in most of the studies (Larsen-Freeman & Long, 1991; Lightbown & Spada, 1993). First, the typical research design is correlational, which means that causal inferences cannot be made. Second, the vast majority of these studies suffer from inadequate measures of the learner trait being examined (the predictor variable) or the L2 learning outcomes (the criterion variables), or both. Even when such difficulties have been overcome (i.e., through careful validation studies and the use of statistical procedures such as causal modelling), explanatory power may still be limited by the 'black box' approach towards the mechanisms of L2 acquisition. This is true, for example, of the sophisticated Socio-Educational model developed by Gardner and his colleagues (see Gardner & MacIntyre, 1992, 1993 for a review).

In response to this state of affairs, Larsen-Freeman and Long (1991) recommended that research efforts should rather focus on explaining SLA processes. But, as a few scholars over the years have argued, to explain SLA we must take individual variation into account -- the baby of SLA theory needs to be interpreted in the bathwater of individual differences. Thus we find Selinker (1972) noting in the now classic "Interlanguage" article that "a theory of second language learning that does not provide a central place for individual differences among learners cannot be considered acceptable" (1972, p. 213, fn.8, cited in Larsen-Freeman & Long, 1991, p.154). Over ten years later, Skehan suggested that

> A model of learning which does not have anything to say about the context of
> learning is lacking, and similarly a model of learning which ignores characteristics
> of the learner is equally deficient. (1986b, p. 189)

And in this decade, R. Ellis (1994) observes,

> It is useful to distinguish two branches of enquiry within SLA research. One has as
> its focus learning and the other the learner. In the case of the former the emphasis is
> on identifying the universal characteristics of L2 acquisition. In the case of the latter

5

the aim is to account for differences in ways in which individual learners learn an L2. These two branches have tended to work independently of each other. A complete theory of SLA must account for and interrelate the findings from both branches. (1994, p.38).

Despite the acknowledged desirability of integrating the two branches of enquiry in SLA, there are clearly blocks to doing so. Why is this so?

The Two Research Traditions in SLA

The separate treatment of universal acquisition theories and of individual differences in SLA textbooks (see for example, R. Ellis, 1994; Larsen-Freeman & Long, 1991; Lightbown & Spada, 1993) reflects the present independence of two distinct research paradigms (Upshur, 1983). The 'measurement of individual differences paradigm' is associated with large sample sizes, data drawn from paper-and-pencil tests, statistical procedures like multiple regression and factor analysis, and theoretical constructs from psychology. In contrast, the 'universals of L2 acquisition paradigm' has been associated with two strands of research activity: the first involves longitudinal process-oriented case studies in which learner language is analyzed qualitatively, while the second strand is characterized by experimental designs in which hypotheses based on linguistic theory are tested through the use of grammaticality judgment data. Small wonder that scholars from each branch (and perhaps even within branches!) have carried on with their work as if in 'two solitudes'. Yet the solitudes must be bridged if the field is to attain explanatory adequacy in accounting for the diverse phenomena of L2 acquisition. Skehan (1998) offers an example of how this 'bridging' can be done. He links the cognitive learner trait called 'aptitude' with the cognitive processes which underlie L2 learning.

6

Individual Differences and SLA Models

It is not the case that SLA theorists have completely neglected to mention how individual differences fit into their theories. In the following section, we will examine a number of influential models in this regard. Rather than attempting to offer comprehensive coverage of all the theoretical approaches found in SLA, discussion here will be limited to models which aim to account for the on-line process of language acquisition from input to output.[1] We will focus on the models proposed by the following scholars: Krashen (1982, 1985), Bialystok (1978); Gass (1988, 1991; Gass & Selinker, 1994), the cognitive theorists (McLaughlin, 1987, 1990; Anderson, 1983, 1990; Ackerman, 1989; O'Malley & Chamot 1990, 1993), and Carroll (1962).

Krashen's Monitor Model.

In the Monitor Model, Krashen (1982, 1985) proposes that there are two kinds of knowledge that a L2 learner can internalize: 'acquired' knowledge and 'learned' knowledge. Each is a product of a different process which he terms 'acquisition' vs 'learning'. The comprehension of input is seen as what drives 'acquisition' (see Figure 1.1). When language input is comprehensible, it becomes available for processing by the Language Acquisition Device (LAD) and thus becomes part of the acquired system. However, the individual must also be in a positive affective state or else this input, despite its comprehensibility, will be filtered out. Krashen refers to this cumulative effect of the individual's attitudes, motivation and stress levels as the 'affective filter'. Only acquired knowledge is responsible for spontaneous language production. According to Krashen, output practice does not affect the development of the internal grammar.

The process of 'learning' is more general in nature than 'acquisition'. Information about language such as that typically provided in formal grammar lessons is handled by the same general cognitive processor that handles other kinds of encyclopedic information such as facts about history or mathematics. 'Learned' knowledge is only available for use when

tasks demand a focus on form. Krashen has always adamantly maintained that there is no interface between the two types of knowledge -- what is learned cannot become part of the acquired grammar by virtue of having been learned; it must be acquired by the separate process of language acquisition.

Two types of learner variables are of direct relevance to the Monitor Model: affective and cognitive variables. In Krashen's view, affective factors influence 'acquisition' because they have direct bearing on the quantity and/or quality of the input that feeds into the process which results in the learner's grammar of the target language. Attitudes, motivation and personality characteristics will influence 'acquisition', but cognitive traits such as IQ or aptitude will not; their influence is restricted to the domain of 'learning' (Krashen, 1981). In Krashen's view, cognitive traits do not act as a filter on input processing for language acquisition. They are, however, relevant to general learning mechanisms. As long as instruction focusses on promoting 'acquisition' through the provision of comprehensible input and a positive affective environment, aptitude differences among learners will not predict success in the L2. In Krashen's model, the role of the cognitive variables is constrained by the fact that there is no interface between explicit or 'learned' knowledge and implicit or 'acquired' knowledge.

Operation of the "Affective Filter"

filter

Input ⟶ ----- | *Language Acquisition Device* | ⟶ *acquired competence*
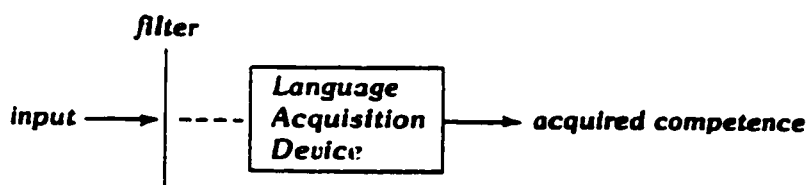
The affective filter acts to prevent input from being used for language acquisition. Acquirers with optimal attitudes (see text) are hypothesized to have a low affective filter. Classrooms that encourage low filters are those that promote low anxiety among students, that keep students off the defensive.[26]

Figure 1.1 Krashen's Monitor Model (Krashen & Terrell, 1983, p. 39)

8
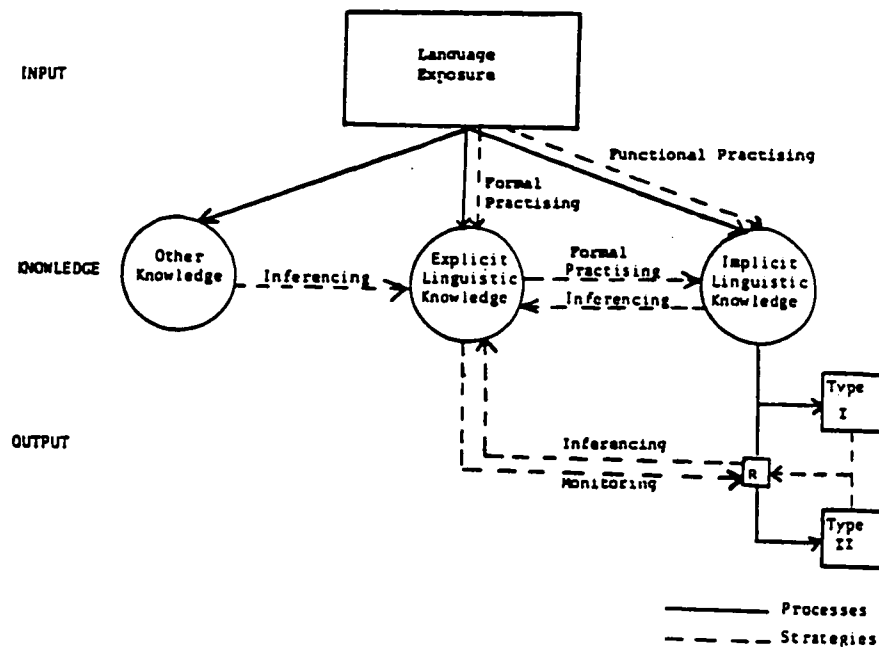
Figure 1.2 Bialystok's model of L2 learning (Bialystok, 1978)

Bialystok (1978)

Bialystok's 1978 model describes L2 learning and use in terms of three unique stages: input, knowledge, and output (see Figure 1.2). What connects the various stages of the model are obligatory processes and optional strategies. Language exposure feeds directly into either 'implicit linguistic knowledge', 'explicit linguistic knowledge' or 'other knowledge'. The type of input will determine which type of knowledge is affected. For example traditional grammar-oriented instruction would lead to explicit linguistic knowledge, while an immersion approach to instruction would lead to implicit knowledge. Different types of learner 'strategies' can facilitate the development of the different types of knowledge: functional practicing, which means using the L2 for communicative purposes,

9

facilitates the development of implicit knowledge, while formal practicing, which involves focussing on the language code itself, facilitates the development of explicit knowledge.

Output is the third stage of Bialystok's model. Here we find two types of language responses. Type I responses are spontaneous and immediate, relying totally on implicit knowledge. Type II responses are deliberate and follow a delay; for this type of response, explicit knowledge can be utilized along with implicit knowledge. Bialystok (1979) provides empirical support for distinguishing between the two sources of linguistic knowledge (implicit vs. explicit) by manipulating experimentally the conditions under which subjects respond to stimuli.

This is a general model of L2 acquisition, proposed to account for the way all learners learn. According to Bialystok, differences in achievement in L2 learning among individuals are attributable "to differences in the efficiency with which the model operates for different people" (1978, p.70). This applies to both acquisition processes and learning strategies. Learner differences are assumed to have a quantitative effect on learning. A study by Bialystok and Fröhlich (1978) revealed, however, that different learner variables have different roles to play in the process. They examined the relationship between cognitive variables (i.e., L2 aptitude and field independence), attitudes, strategy use, and L2 achievement in high school learners of French. It was found that aptitude and strategy use were significantly related to performance on the achievement tests. Attitudes were significantly related to strategy use but not to achievement directly. (More will be said about this study in Chapter 3.) Later developments in Bialystok's thinking have led to a different model of language use (Bialystok & Ryan, 1985a, 1985b) which will be discussed in Chapter 2, Part III.

10
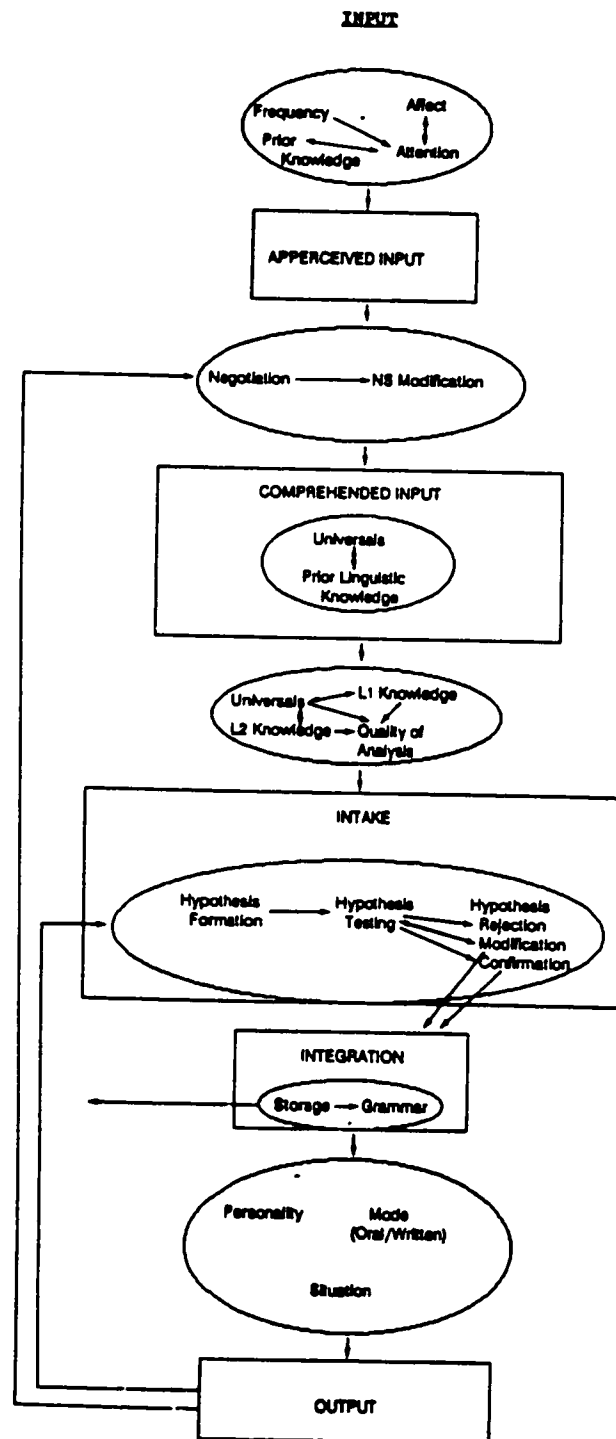
A model of second language acquisition (Source: Gass, 1988a).

Figure 1.3 Gass' model of L2 acquisition (Gass & Selinker, 1994)

Gass (1988).

Gass (1988, 1991; Gass & Selinker, 1994) proposes a model that describes the different stages in the process by which the learner converts input to output (see Figure 1.3). The model has five stages: (1) apperceived input, (2) comprehended input, (3) intake, (4) integration, and (5) output. The first stage involves 'apperceiving' or 'noticing' some feature of the target language in the ambient speech. In the next step, the feature which has been noticed is comprehended at some level. Here, comprehension of the input refers to a continuum of possibilities that range from understanding the general message to performing a mini-structural analysis. When the comprehended item has been 'assimilated', it is considered to have become intake which is the fourth stage in the model. Gass defines intake as a process which mediates between the input and the learner's internalized grammar. During the stage of integration, features that have been apperceived, comprehended and then become part of intake are either integrated into the L2 system or are stored for later integration. Finally, the language feature that has become integrated is available for output, although factors such as level of confidence, and the strength of knowledge representation will determine what the output will look like.

The interesting stage from the perspective of the present discussion is the stage where the input is noticed. Gass suggests that the factors that influence what gets noticed are: frequency in the input (either very frequent or very infrequent), affective variables such as motivation, attitudes, prior knowlege (of the L1, of other languages, of the world), and selective attention. While explicit language knowledge does not have a place in the chain, Gass (1991) proposes that explicit grammatical teaching can influence acquisition at the apperception phase. Learners can make use of such information in order to selectively attend to the target feature in the input. Thus the effects of explicit instruction will not be evident immediately, but will influence the process that is responsible for the acquisition of structures through meaningful input. This selective attention role for explicit knowledge about the target language thus represents an alternative to the 'no interface' (i.e., Krashen)

12

and 'full interface' (i.e., Bialystok) positions concerning the effect of instruction (see R. Ellis, 1994, pp. 652-660 for discussion).

Gass' model is clearly influenced by the earlier approaches to SLA that placed a priority on input-processing as the mechanism that drives acquisition. Just as in Krashen's Monitor Model, affective variables are explicitly integrated into the model as factors that can influence what aspects of the input are available for further processing, which may eventually lead to integration into the learner's grammar. As for cognitive variables, Gass suggests that "one might, for example, exclude such factors as memory, IQ, and analytic ability which relate to a learner's success. These factors are more likely to have a quantitative rather than a qualitative effect on the nature of learning" (1988, fn. 2, p. 214). She does not elaborate further on this point, so it is not clear why motivation is not also considered to be quantitative in its effect on learning. Moreover, it seems very likely that the notion of 'selective attention', which Gass considers to be a factor that directly influences the apperceived input stage, may be related to or influenced by cognitive variables such as learner aptitude.

Cognitive Theory in SLA.

From models that have developed from within the field of SLA, we now turn to an approach that comes from cognitive psychology. McLaughlin (1987, 1990) introduced the information-processing model into SLA under the rubric Cognitive Theory; it is based on the pioneering work of Shiffrin and Schneider (1977). The fundamental assumption of an information-processing model is that the development of skill in all domains, whether tennis or typing or a foreign language, involves the same processes. Performance is viewed in light of the limitations on the amount of effort humans can allocate to any particular cognitive task. Two separate modes of processing have been proposed: controlled processing which is effortful and demands attentional resources, and automatic processing which is less effortful and does not make demands on attention.

Initially, language can only be handled using controlled processing. As all beginning learners can affirm, the effort required for speaking and listening to an unfamiliar language is extremely exhausting. Practice plays an important role in improving performance so that it becomes more rapid and stable (Segalowitz & Segalowitz, 1993). With practice, components of a skill become automatized, which liberates attentional resources to be applied to new learning.

Such a view of skill acquisition would predict language learning to be a matter of accretion. To account for observed phenomena such as apparent backsliding and U-shaped development, as well as apparent instantaneous acquisition (Lightbown, 1985), the model is supplemented by the concept of 'restructuring' which is

> a process in which the components of a task are coordinated, integrated, or reorganized into new units, thereby allowing the procedure involving old components to be replaced by a more efficient procedure involving new components. (McLaughlin, 1990, p.118)

Thus, in addition to the process of automatization through practice, restructuring of internal representations can lead to qualitative changes in learners' performance, which may sometimes appear to involve decrements in skill.

A more elaborated information-processing model has been developed by Anderson (1983; 1990). The skill-learning aspect of Anderson's ACT* (Adaptive Control of Thought) model builds upon the distinction between declarative and procedural knowledge. "Declarative knowledge refers to knowledge about facts and things; procedural knowledge refers to knowledge about how to perform various cognitive activities." (Anderson, 1990b, 219). Declarative knowledge is encoded in a propositional network while procedural knowledge is encoded in what are termed 'productions' (also 'production systems') (Anderson, 1983); these are sets of condition-action pairs in which the clauses specified in the condition must be matched against information in declarative knowledge. An example

14

given by O'Malley & Chamot (1990, p. 74) is: "IF the goal is to say a memorized greeting formula, and the context is an informal one, THEN choose the appropriate language style".

In Anderson's model, skill learning involves the proceduralization of rule-bound declarative knowledge. This takes place in three stages. At the first stage, the cognitive stage, learners either receive instruction about how to do the task, observe an expert performing the task, or attempt to figure it out themselves. This involves conscious activity on the part of the learner and the knowledge gained is typically declarative and can be described verbally by the learner; performance at this stage is slow and full of errors. At the next stage, termed associative, declarative knowledge is turned into procedural knowledge but the declarative representation is not lost. Here performance is still slower and more errorful than expert performance. The final or autonomous stage is characterized by performance which is automatic, error-free, and with little demand on working memory or consciousness.

In this model, learners are usually consciously aware of rule applications during initial stages of acquisition and unaware of rules once proficiency has been achieved. Anderson (1990, p. 259) does suggest, however, that the two types of knowledge may coexist as in the case of the foreign language learner who can speak fluently while still remembering many rules of grammar. But it is the procedural, not the declarative, knowledge that governs the skilled performance.

Ackerman (1989).

How do individual differences fit into an information-processing model like Anderson's? Ackerman (1989) makes specific predictions about the influence of individual differences depending on the stage of learning attained. While Ackerman has not applied his model to the situation of L2 learning, Segalowitz (1997) has made a case for its usefulness in this domain. The information-processing model of learning that Ackerman uses is analogous to Anderson's described above. According to Ackerman, a given task

will make demands upon the individual's general cognitive abilities and specific abilities related to the content domain; in addition, speed of processing will also influence performance. The relative importance of the two will, however, change as the learner progresses from the initial effortful phase of learning to skilled performance. With novel tasks (Stage 1) performance will correlate highly with general cognitive abilities and with specific content abilities, but not with perceptual speed and psychomotor abilities. Segalowitz suggests that this is where L2 aptitude plays a role. With practice, the relationship between performance and general and content abilities will diminish and the relationship between performance and perceptual speed and psychomotor abilities will increase until, at high levels of performance, individual differences in psychomotor abilities will be the primary determinant of performance differences. Motivation is seen as a factor that influences the degree to which the individual exerts the effort to practice (Ackerman & Schneider, 1985) and therefore has an overall quantitative influence on learning.

O'Malley and Chamot (1990).

O'Malley and Chamot (1990, 1993) have explored the potential of the cognitive models of learning proposed by Anderson (1983) and by Gagné (1985) for the domain of SLA. L2 learning involves both declarative and procedural knowledge but declarative knowledge is learned best through building upon existing schemata, and linking new concepts to old ones; procedural knowledge, on the other hand, is learned most effectively through integrative practice of a skill in a meaningful and goal-oriented way. Three methods for learning procedural knowledge are:

1) to observe and imitate complete sequences of expert performance;

2) to identify manageable but meaningful and integrated components of the complete skill, gain partial mastery over the components by practicing them with feedback, and piece them together to make a complete action sequence that achieves a meaningful goal;

3) learn by analogy (i.e., draw parallels between familiar and unfamiliar language).

For O'Malley and Chamot (1993), learning strategies and motivation are the two major learner characteristics relevant to their theoretical framework. Their cognitive view of L2 acquisition places an emphasis on strategies, and they interpret the work of others in this light. Thus, in their discussion of aptitude, they conceptually redefine Carroll's (1962) four sub-components (i.e., phonetic coding ability, grammatical sensitivity, inductive language learning and rote memory) as different learning strategies. (We will examine the concept of aptitude in Chapter 2, Part I.)

Like Ackerman, O'Malley and Chamot consider motivation to be integral to L2 learning since motivational characteristics will determine how learners will respond to instruction and how much effort they will devote to learning. They see motivation as intricately related to learner strategies, and to metacognitive knowledge. When learners are able to match their learning strategies optimally to the learning task, they will see the connection between their own efforts and learning outcomes, and thus gain a "sense of self-efficacy" and this promotes further learning. Note that in all the cognitive models, the role of motivation offers a striking contrast to that of Krashen's since there is no dichotomy between the influence of affective and cognitive variables. As with other cognitive models, O'Malley and Chamot emphasize the intentional, conscious, and strategic behaviour of L2 learners.

Information-processing Models and SLA

Information-processing models such as those of Anderson or Ackerman have not as yet gained wide-scale acceptance in SLA circles. One reason for this may be the difficulty of reconciling the psychologist's view of learning, which is common to all the Cognitive models discussed here, to that of many SLA theorists. The former see learning as a progression from declarative to procedural knowledge mainly through output practice while the latter hold that acquisition proceeds mainly through the processing of meaningful input. Furthermore, it is not clear how these general information-processing models can account for some of the "accepted findings" in SLA such as:

- the fact that common patterns in development have been documented in different kinds of learners (i.e., adults and children, L1 and L2 learners) under diverse conditions of exposure (i.e., formal and informal contexts) (Long, 1990);

- the fact that learners in comprehension-based programs are able to produce L2 structures without having practiced production (Lightbown, 1992);

- the fact that practicing L2 forms does not necessarily lead to their acquisition (Lightbown, 1983)

In order to be viable as a model of SLA, a general cognitive model like Anderson's needs to deal with the issue that not all L2 learning begins with explicit knowledge of a rule that is practiced until it is automatized. Anderson suggests that in the cognitive stage of learning. learners gain declarative knowledge through instruction or through observation of an expert, implying that both methods lead to the same outcome. Yet studies in SLA have shown clear differences in L2 performance between learners in formal and those in informal settings (Lightbown, 1985; Long, 1988). Furthermore, as R. Ellis (1994) has pointed out, the nature of language practice needs to be clarified. There is, however, an information-processing model which provides a better fit to the findings of SLA. This model proposed by Skehan (1998) includes features from general information-processing models (e.g., the distinction between working memory and long-term memory) as well as

18

from the psycholinguistic processes involved in comprehension. Skehan's model will be discussed more fully later in this chapter.

Carroll (1962).

Finally, we turn to a model older than the others reviewed so far, one proposed by Carroll (1962). The Model of School Learning cannot be considered a model of SLA since the processes of learning are treated as a 'black box', but it does attempt to account for the interaction between individual differences and L2 learning in the classroom. The model allows predictions to be made about learning outcomes based on the interactions between the factors of time, quality of instruction and individual differences. The instructional variables include adequacy of presentation of learning tasks, and opportunities allowed for learning tasks; the individual difference variables include general intelligence, aptitude, and motivation. Aptitude is operationalized as the time required by the individual to learn a task, and motivation is the maximum amount of time that the individual is willing to devote to learning. In situations where ample time is available, a motivated person with low aptitude is able to attain success because despite requiring more time to learn, he/she is willing to devote the maximum amount of time to studying. In contrast, a learner with high aptitude will not require much time to learn, and thus can succeed even with low levels of motivation. Carroll also proposes that high quality instruction may make aptitude less significant.

While the model is appealing in how it predicts the interaction of motivation and aptitude, viewing them as off-setting forces in the learning process, there are some serious weaknesses with it. The main problem is that the language acquisition process is unspecified. Carroll's model was developed during the period when behaviourism was the dominant paradigm. Now, many unanswered questions come to mind: 'Motivation is time devoted to doing what exactly?', 'How do you determine the quality of instruction?' and

19

'What does it mean to be successful in L2 learning?' Despite these gaps, the Model of School Learning has as yet unexploited potential for use in classroom-based research.

Neglect of Aptitude

With respect to the issue of how individual differences fit into SLA models, the discussion so far has revealed the following points: Carroll's Model of School Learning deals neatly with the interaction between aptitude and motivation, but lacks an adequate account of L2 learning; information-processing models are limited because they do not adequately account for acquisition that arises from processing input for comprehension; and SLA models are clearer about the role of motivation than of aptitude. In the models proposed by Krashen, Bialystok, and Gass, motivation is understood as being a variable that influences acquisition in a quantitative way: by being more interested in language learning, one increases the amount of usable input. Surprisingly, although language acquisition is a cognitive process, a cognitive variable like aptitude is dismissed as irrelevant to the acquisition process itself.

This neglect of aptitude reflects the zeitgeist of the 1970's when the new field of SLA turned away from behaviourist models of language learning, and embraced the methods and goals of L1 acquisition research (Corder, 1968). Research focussed on the universal characteristics of language learning across languages and learners; similarities in the acquisition process were found between L1 and L2 learners, between child and adult L2 learners, and between learners from different L1 backgrounds (Dulay, Burt, & Krashen, 1982). It was proposed that the process is the same in all cases: the learner must process the input in such a way as to 'break the code' of the target language through hypothesis-testing and thus arrive at a grammar (i.e., internalized system) of the target language (Cook, 1991). Since the child acquiring an L1 does this intuitively, without conscious analysis, then the same must be true for the older L2 learners. The fact that the 'little linguist in the head' did not use the methods of 'the big linguist' (Sharwood Smith, 1994),

20

meant that the role of metalinguistic knowledge in L2 learning was seen as minimal. However, the present climate in L2 teaching and SLA research circles has favoured the renewal of interest in the role of formal instruction and the conscious involvement of the learner in the acquisition process.

The Role of Focus on Form

There is now a consensus in SLA that comprehensible input is not sufficient for attaining native-like ability in a L2, and that some attention to language form is necessary. This notion that learners need to have their attention drawn to formal aspects of the target language is discussed under different labels depending on the theoretical orientation of the researcher: negative evidence (White, 1989), consciousness-raising (Rutherford & Sharwood Smith, 1985), input enhancement (Sharwood Smith, 1991, 1993), focus on form (Long, 1990; Lightbown & Spada, 1993), corrective feedback (Lyster & Ranta, 1997), and noticing (Schmidt, 1990; Robinson, 1995). The hypothesis rests on many types of evidence. On theoretical grounds, White (1987) argues that it is <u>incomprehensible</u> rather than comprehensible input that pushes the learner to reevaluate hypotheses about the target language. Furthermore, she (White, 1989) argues that for some aspects of the target language, learners need 'negative evidence' or information about what is not grammatical in the L2. An example of this type of situation occurs when French native-speakers are learning the adverb placement rule in English . In French an adverb can be placed between the verb and the direct object (SVAO) but not between the subject and the verb (SAV); in English the reverse is true: adverbs go in the SAV position but cannot go in the SVAO position in simple sentences. For example,

> *Je toujours regarde la television
> Je regarde toujours la television
> I always watch TV
> *I watch always TV

Many French-speaking learners are observed to use the incorrect SVAO pattern in English. White argues that in order to master the rules of adverb placement, the

francophone learner has to notice that adverbs in English do <u>not</u> occur in the SVAO position, a situation which constitutes a considerable learning problem in the absence of negative evidence. The results of an experiment involving the teaching of the adverb placement rule provides some support for this notion: learners who received negative evidence concerning the SVAO position performed significantly better on such things as grammaticality judgment and sentence manipulation tasks as compared to groups of learners who did not receive such information (White, 1991). It is evident in this study that White assumes that explanations about the L2, which are by nature metalinguistic, will allow learners to restructure their grammar of the L2, and presumably then be able to access this knowledge in spontaneous language use. It is impossible, however, to confirm this with White's data because the tasks used to measure learners' knowledge and use of the rule were metalinguistic in nature.

Research findings from classroom contexts in which the instructional focus is on the communication of meaning also highlight the inadequacy of input-based approaches. For example, Lightbown and Spada (1993) conclude from their studies of learners in communicative, intensive ESL classes in Quebec that learners will be able to acquire many aspects of the L2 from the input, but need help in noticing certain problematic forms or structures. This conclusion is also supported by a study of the writing of francophone school children who had had five or six years (for about 30 minutes a day) of mainly comprehension-based ESL instruction in New Brunswick (Lightbown & J. White, 1995). Despite ample opportunity to read and hear English, learners were still unable to spell some of the most frequent words in the language.

In the context of French immersion, Harley and Swain (1984) and Swain (1985) report on the language development of grade 6 students in Toronto. Despite having seven years of comprehensible input, these students did not acquire many grammatical structures that were to be found in the input (eg., imparfait, conditional, etc.). Swain (1993) concludes that what L2 learners need in addition to input is output opportunities that push

them to 'notice the gap' between what they know and what they need to discover about the L2.

Schmidt (1990, 1994) draws upon the cognitive psychology literature investigating implicit vs explicit learning to argue that "the subjective experience of 'noticing' is the necessary and sufficient condition for the conversion of input to intake in SLA" (Schmidt, 1994, p. 176). Note that the term 'noticing' does not refer to conscious induction of rules or of hypothesis-testing, but simply to "the registration of the occurrence of a stimulus event in conscious awareness and the subsequent storage in long-term memory" (Schmidt, 1994, p. 179). Evidence in support of the role of noticing in SLA comes from Schmidt's diaries documenting his experiences learning Portuguese in Brazil (Schmidt & Frota, 1986). A careful analysis of his own retrospective accounts in comparison to analyses of the input he received and samples of his oral production, indicated a relationship between what Schmidt noticed in the input and what was found in his output (Schmidt, 1990).

Further evidence in support of the importance of focus on form is provided by studies that have compared explicit and implicit learning under controlled laboratory conditions. These studies make use of the experimental paradigm developed by Reber (see Reber 1989 for a review). In these experiments, subjects are exposed to strings generated by an artificial grammar (usually a finite-state grammar of meaningless strings). One group gets instructions to look for the rules (the explicit condition) while the other is given neutral instructions such as "Look at these strings" or "Memorize as many of these strings as you can" (the implicit condition). Shortly afterwards (usually about 10 minutes), all subjects are given a discrimination test in which they must judge strings for their well-formedness. SLA studies have used either artificial languages with meaningful referents (de Graaff, 1997; deKeyser, 1995), or natural languages (N. Ellis, 1993; Robinson, 1996). The overall finding from these lab studies is that the explicit learning conditions lead to superior results over implicit learning conditions on post-treatment grammaticality judgment tasks.

Schmidt (1990) proposes that several factors influence whether a form will be noticed; these include input characteristics, the provision of focussed instruction, task demands, and factors internal to the learner. In the last category, Schmidt includes individual differences in processing capacity. Skehan (1998) provides a visual schematicization of Schmidt's hypotheses concerning noticing and incorporates them into a model of language processing (see Figure 1.4). As Skehan's diagram illustrates, target language forms that are noticed become available for further processing in working memory and are then stored in long-term memory.
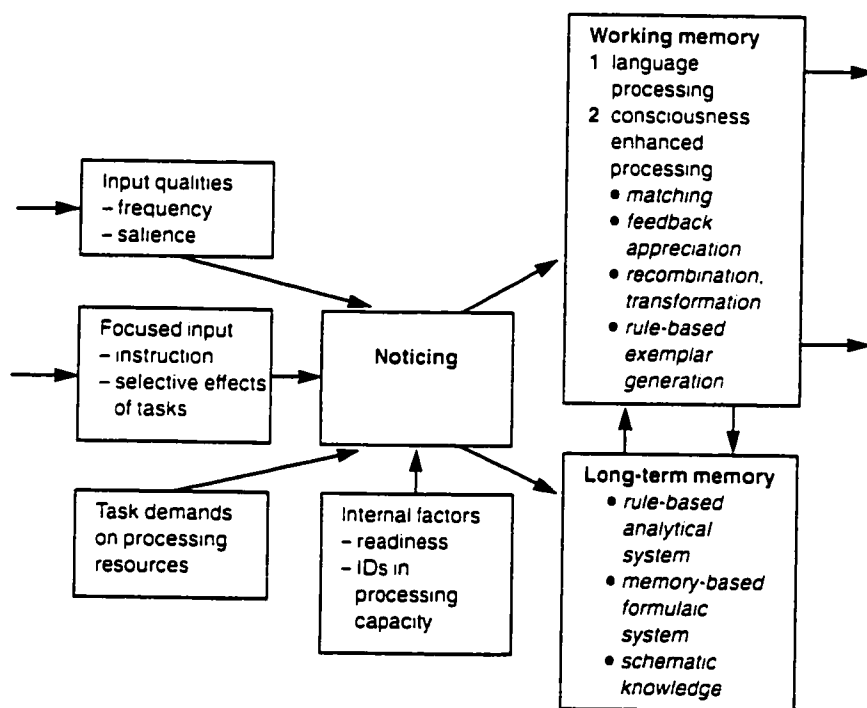
Figure 1.4 Skehan's information processing model (Skehan, 1998, p. 52)

A Role for Cognitive Variables in L2 Acquisition

This brief and selective review has highlighted the fact that various strands of theory and research in SLA lead to the same conclusion that L2 learners need to focus on the form of the target language in some way (see Spada, 1997 for a review of 30 classroom and lab studies; cf. Truscott, 1998 for a different interpretation of the research findings). However, teaching experiments have not always been able to demonstrate that form-focussed instruction leads to changes in learners' interlanguage knowledge or spontaneous communicative performance. One possible reason for this is that such instruction is often metalinguistic in nature and is therefore mediated by the learner's ability to handle this type of information. Establishing a place for focus on form in a model of SLA must necessarily entail a role for the learner's ability to focus on form. Skehan (1998) expands on the ideas in Schmidt (1990) and suggests that some individuals will "be more able to notice, for given input, new forms which may then be integrated into their language development" (p. 50). This hypothesis is addressed by the present study. The terms 'noticing' and 'focus on form' are relative newcomers to SLA theory, but they are conceptually related to others which have been extensively studied over the years. These concepts are reviewed in Chapter 2 under the headings of aptitude, learner style and metalinguistic skill.

NOTES

[1] Other influential theoretical models are omitted here because they operate at a different levels of explanation (see Marr, 1982). For example, scholars working with Universal Grammar (e.g., Cook, 1988; L.White, 1989) and connectionist models (e.g., Ellis & Schmidt, 1997) are mainly concerned with the issue of how linguistic knowledge is represented. The Competition model, on the other hand, is mainly concerned with the processing of input (e.g., MacWhinney, 1989; McLaughlin & Harrington, 1989).

CHAPTER 2

THE LEARNER'S ABILITY TO FOCUS ON FORM

In Chapter 1, it was suggested that models of L2 acquisition need to integrate individual differences into the description of universal processes. In particular, recent SLA research which has highlighted the role of what has been referred to as 'negative evidence', 'noticing' and 'focus on form' suggests the possibility that learner differences in the ability to focus on grammatical structure might influence the processes and outcomes of SLA. In this chapter, three different theoretical approaches to describing the ability to focus on language form are presented: the first is associated with Carroll's model of L2 aptitude, the second with the notion of learner style, and the third with models of metalinguistic skill. The chapter ends with a synthesis of the similarities and differences among the approaches.

Part I: The Ability to Focus on Form as L2 Aptitude

The ability to analyze grammatical form is one aspect of L2 aptitude as proposed by Carroll (1962). In Chapter 1, we considered Carroll's Model of School Learning in which aptitude and motivation were related to each other in terms of the time required for learning and the time the learner is willing to devote to learning, respectively. This model was an outgrowth of the development of the Modern Language Aptitude Test (MLAT).

Carroll and the MLAT

Carroll's work on aptitude began in the 1950's in a very specific context: the American government needed a cost-efficient way to predict success in foreign language learning for military and other personnel who were studying foreign languages in intensive courses. The development of the MLAT for diagnostic purposes provided an opportunity for an empirical approach to determining the nature of the factors underlying L2 aptitude. Carroll and his colleague Sapon embarked on a research program "guided when possible by theory, and where that was impossible, by hunches" (1962, p. 94). Initial validation

studies were carried out with the short week-long trial courses that the military had used for screening prospective candidates for language courses. The procedure followed for the validation phase involved running correlations between aptitude test scores and criterion measures such as final course grades, tests of foreign language skills, or instructor estimates of language learning ability. The MLAT (Carroll & Sapon, 1958) typically correlated with criterion measures in the .4 to .6 range, thus demonstrating a moderately strong relationship between aptitude and L2 learning outcomes (Carroll, 1981).

On the basis of a factor analysis, Carroll proposed that L2 aptitude consists of four distinct abilities: phonetic coding ability, grammatical sensitivity, rote learning, and inductive language learning ability. The definitions and the corresponding subtests on the MLAT are described in Table 2.1.

While the MLAT is perhaps the best known aptitude test, other tests have been developed over the years. In general, these alternative tests conceptualize aptitude along the lines of Carroll's model. For example, Pimsleur's (1966) Language Aptitude Battery (PLAB) which was intended for use with junior high school students, is composed of tests of English vocabulary, language analysis ability, sound discrimination ability, and sound-symbol association ability; in addition, the student's grade-point average and information concerning interest in language learning are included in the overall score. Despite high hopes, the predictive power of the more recent aptitude tests such as the PLAB, the Defense Language Aptitude Battery (Petersen and Al-Haik, 1976), and the innovative VORD[1] aptitude battery (Parry & Child, 1990) is on the whole no greater than that of the MLAT.

Table 2.1

Carroll's Model of Foreign Language Aptitude

Component	Definition	MLAT subtest
phonetic coding ability	"an ability to identify distinct sounds, to form associations between those sounds and symbols representing them, and to retain these associations"	Phonetic Script Spelling Clues
grammatical sensitivity	"the ability to recognize the grammatical functions of words (or other linguistic entities) in sentence structure"	Words in Sentences
rote learning ability for foreign language materials	"the ability to learn associations between sounds and meanings rapidly and efficiently, and to retain these associations"	Number Learning Paired Associates
inductive language learning ability	"the ability to infer or induce the rules governing a set of language materials, given samples of language materials that permit such inferences"	Not tested on the MLAT

Source: Carroll, 1981, p. 105

Grammatical Sensitivity and Inductive Language Learning Ability

Carroll's model of aptitude includes two components which deal with the ability to analyze language structure: grammatical sensitivity and inductive language learning ability. In its final commercial form, the MLAT contained no test of the latter. This was the case for purely practical reasons: in the test-development phase, the task that involved learning the rules of an artificial language (called Tem-Tem) was found to be too cumbersome to administer (Carroll, 1981). The subtest on the MLAT which measures grammatical sensitivity is Words in Sentences. This is a grammatical analogies test in which the test-taker has to parse a sentence so as to be able to match an underlined word in one sentence with a word or phrase having the same grammatical function in a second sentence (see example in Appendix A). Other aptitude tests, such as the PLAB. the York Language

28

Aptitude Test (Green, 1974/75) and the Vord do include an inductive learning subtest. Of these, the most widely used is the PLAB's Language Analysis. For this test, a series of sentences from an unknown language (Kabardian) are presented along with their English translations; the students must infer the rules in the language and apply them to novel sentences.

In his model of aptitude, Skehan (1989, 1998) has chosen to collapse grammatical sensitivity and inductive learning into a single component. This is based, apparently, on the strong correlation that he found between a test of grammatical sensitivity and the York Aptitude test which measures inductive language learning ability (Skehan, 1990). Indeed, the cognitive operations required by Words in Sentences are similar to those required by inductive learning tests. The first requires the examinee to analyze the grammatical patterns of sentences in a well-known language (the L1), while the second involves extracting grammatical patterns from a set of sentences in an unknown language. Until recently, little research has examined the relationship between these two tasks. Alderson, Clapham, and Steel (1997) administered Words in Sentences and an inductive language learning test developed by Davies (1971) to British university students. The test consisted of a text in Swahili of which the first few sentences had been translated into English; students were asked to translate the next few sentences. The analysis revealed that grammatical sensitivity did not correlate with inductive language learning as measured by this task. Since different measures of inductive language learning were used in the studies by Alderson et al., and by Skehan, it is not possible to draw firm conclusions. Certainly, the relationship between grammatical sensitivity and inductive language learning ability, as it has been operationalized, is not straightforward.

Carroll's Conception of Aptitude

Carroll's view of aptitude reflects both the practical purpose of his research program, and the behaviorist paradigm then prevalent in psychological testing. Three

characteristics of this view merit attention. First of all, Carroll operationalizes aptitude in terms of rate (i.e., speed) of learning a foreign language in the context of some sort of formal instruction, either a language course or program of self-study. According to this view, learners with the highest aptitude are those who learn the fastest. Second, aptitude is seen as related to, but independent of, intelligence. Third, the ability to learn a L2 quickly is seen as a stable, possibly innate, psychological trait. In the next sections, each of the three aspects of aptitude will be discussed: aptitude as speed of learning, aptitude as distinct from intelligence, and aptitude as an innate trait.

Aptitude as speed of learning.

It is important to understand that Carroll's conception of a specialized talent for learning a L2 reflects the specific time and place in which he was developing the MLAT. There was a need at that time in the US to train a large number of people in foreign languages, a process which was very costly to the public purse, so it was necessary to find a way of ensuring that those individuals streamed into language training would succeed. These courses tended to be intensive in nature, consisting of full-time study for a period of eight to twelve months (Carroll, 1962). In such a setting, it is fairly clear why speed of learning would be the operationalization of talent. This contrasts with the choice made by researchers working in second language rather than foreign language contexts who operationalize talent as native-like attainment (Ioup, Boustagui, El Tigi, & Moselle, 1994; Novoa, Fein, & Obler, 1988; Obler, 1989; Schneiderman & Desmarais 1988).

Aptitude and intelligence.

Carroll (1981) maintains that L2 aptitude is not the same thing as intelligence since measures of aptitude and measures of intelligence do not share the same pattern of correlations with L2 achievement. Skehan (1998) reviews the studies which have examined the relationship between intelligence and language aptitude. Using simple correlations,

30

Gardner and Lambert (1972) and Skehan (1982) found that IQ and aptitude tests generated low to moderate correlations with each other, but that aptitude tests were better predictors of L2 proficiency. On the other hand, more complex research designs using factor analysis (Sasaki, 1993a, 1993b; Wesche, Edwards & Wells, 1982) have found stronger relationships between aptitude and IQ. Skehan (1998) points out that in both cases a second-order factor analysis was needed to bring out this relationship; the first order factors suggested some degree of separation. He argues that the difference in results may be due to the fact that the earlier studies by Gardner and Lambert and by Skehan were conducted with fairly unselected groups while the Wesche et al. and the Sasaki studies made use of subjects who had been streamed in some way. He concludes, provisionally, "that aptitude is not completely distinct from general cognitive abilities, as represented by intelligence tests, but it is far from the same thing" (p. 209).

The relationship between aptitude and cognitive abilities appears to be strongest, according to Skehan, when dealing with language analytic ability component of aptitude rather than with memory or with phonetic coding ability. Here we see that treating L2 aptitude in a monolithic way obscures the nature of the relationship between general cognitive abilities and specific linguistic ones. That IQ should be a more strongly related to grammatical sensitivity than to phonetic coding ability is something that becomes understandable as we examine further the nature of this ability.

Aptitude as an innate trait.

In his earlier writings, Carroll (1973) argued for the innateness of L2 aptitude, linking it to the Critical Period in which the learner has a "heightened capacity to learn ANY language" (p.6) which means that individual differences in aptitude arise because of different rates of decline in this capacity. More recently, he has taken a more neutral position:

an aptitude might be at least in part innate, or it could have developed over a long period as a result of the individual's experience and activities. In any case, aptitudes are regarded as being relatively enduring. Although it may be possible to improve an individual's aptitudes by special instruction or training, this is not easily done, and there seem to be limits to how far aptitudes can be changed.

(Carroll, 1985, p. 84-85)

In recent years, support for the innateness or at least stability of aptitude has come from two sources: (1) a study examining the relationship between rate of L1 development and L2 aptitude later in life, and (2) studies investigating the effect of language learning experiences on L2 aptitude. We begin with the first strand of evidence provided by Skehan (1986b, 1989, 1990) who reports on the findings from the follow-up study to the Bristol Language Project (Wells, 1985).

The Bristol Project follow-up study.

The aim of this study was to investigate the relationships among L1 development in early childhood, performance on L2 aptitude tests, and success in L2 learning in secondary school. Skehan (1986b, 1989, 1990) was able to administer aptitude and L2 proficiency tests to 103 of the individuals who had participated in the Bristol Language Project (Wells, 1985) as young children. In the original longitudinal study, Wells and his colleagues collected data from two samples of children ($\underline{N}$ = 128), one group from the age of 15 months until the age of 42 months and the other from age 39 to 60 months. A stratified sample was obtained by selecting children from a range of family backgrounds. Each child was audio-recorded once every three months by a recording set-up that was preprogrammed to sample at frequent but irregular intervals over the course of a day without there being an outside observer present. In addition to the spontaneous production data, tests of different kinds were administered in conjunction with each observation.

Parental interviews also provided other types of information such as the quality of literacy-based activities the children were exposed to.

The follow-up phase to the Bristol Project involved administering (1) measures of different components of L2 aptitude (verbal intelligence, grammatical sensitivity, inductive language learning ability and phonetic coding ability); and (2) standardized tests of speaking, listening, writing, and reading in French or German. Skehan found that aptitude and L2 achievement were strongly related to each other. On the other hand, L1 development and L2 success were not found to be directly related to each other. What is interesting is that strong correlations were found between some measures of L1 development and many of the aptitude measures, particularly the more 'language analytic' tasks. For example, significant, moderate-strength correlation coefficients in the range of .4 to .5 were found between grammatical sensitivity and 'MLU at 42 months', between grammatical sensitivity and 'range of adjectives and determiners', and between inductive language learning ability and 'vocabulary at 39 months' (Skehan, 1990; note that only the results for the first cohort ($\underline{n}$ = 53) of the original sample are available in published reports). In other words, speed of learning L1 was correlated with tests that purportedly measure potential speed of learning in L2. However, a simple innatist explanation is not the only interpretation of the results since 'nurture' variables such as family background indices also correlated moderately with grammatical sensitivity. Based on regression analyses using the wide variety of measures available, Skehan (1986b, 1989) concluded that aptitude is a product of two separate groups of influences: one which reflects an innate capacity for learning which is a 'residue' of L1 development, and the other which reflects the development of the ability to handle language in a decontextualized way.

Aptitude and the Effect of Language Learning Experience

Many people believe that the more languages a person knows, the more successful that person will be in further language learning. As we shall see, research investigating this issue has not unequivocally confirmed this belief. The discussion will begin with the correlational studies and then move on to experimental studies.

Correlational studies.

Eisenstein (1980) investigated whether bilinguals, being more experienced in language learning, would have higher levels of aptitude than monolinguals, and whether multilinguals would have higher levels of aptitude than bilinguals. MLAT scores and self-ratings on language abilities were provided by a sample of 93 young adults studying modern languages at a college in New York City; 57 of the students did not have any L2 experience before the age of 10, and 36 had been exposed to at least one L2 before the age of 10. Subjects were matched for socio-economic status. Bilinguals significantly outperformed the monolinguals on the MLAT and also rated themselves more highly on their L2 skills. Eisenstein also found that the multilinguals had higher college grades than the bilinguals. On the basis of these findings, one might argue that the experience of learning other languages causes individuals to have higher levels of aptitude. However, another explanation is possible: people generally do the things that they are good at, so it is possible that any group of multilinguals will include a greater proportion of individuals with higher levels of innate aptitude. Is aptitude the cause or the effect? The research design does not allow us to decide between a conclusion that multilingualism leads to improved aptitude or aptitude leads to multilingualism. Unlike the children in the Bristol Project Study, we do not know where the individuals in Eisenstein's study started from.

In contrast to Eisenstein's study, two other studies did not find an effect for experience. Sawyer (1992) administered the short version of the MLAT and a language experience questionnaire to a group of 129 students beginning short intensive courses in

various Southeast Asian languages. Both correlational and principal components analyses were performed to examine the relationships among language learning experience, aptitude, and L2 proficiency. Aptitude correlated moderately with final course grades and with some of the L2 proficiency measures. It did not, however, correlate with biographical variables such as the amount of exposure to other languages. Language learning experience, aptitude and proficiency all emerged as independent factors in the principal components analysis.

Harley and Hart (1997) compared the aptitude and L2 proficiency of learners who had had either an early or a late start to L2 learning. Four grade 11 French immersion classes ($\underline{N}$ = 65) were involved in the study; the students in two of the classes were continuing in a partial immersion program that had begun in grade 1 (the early group), while the students in the other two classes were continuing in a partial immersion program that had begun in grade 7 (the late group). A variety of tests were administered, including two memory measures, the Language Analysis subtest from the PLAB, and various L2 tests (more will be said about this study in later sections). Results showed that the early immersion students did not have higher levels of aptitude overall, despite having 11 years of immersion in French compared to the 4 years of the late immersion students.

Experimental studies.

A recent teaching experiment that has shown a positive effect for language learning experience on aptitude was conducted by Sparks, Ganschow, Fluharty, and Little (1995). In earlier research, Sparks, Ganschow and colleagues documented (e.g., Sparks & Ganschow, 1993) that students who have strong L1 skills generally have strong L2 aptitude and achieve higher grades in foreign languages at school, and that those who experience difficulties in L2 courses at school exhibit overt or subtle L1 problems that are similar to those found in learning disabled students. Sparks et al. (1995) compared changes in aptitude scores of learning-disabled and of non-learning-disabled high-school students after they had studied Latin. Additionally, they examined whether using a "multisensory,

structured language approach" to L2 teaching would improve the aptitude scores of the learning-disabled students. This method is characterized by direct and explicit teaching of the phonology/orthography of the L2 in a step-by-step approach, providing students with opportunities to hear, see and write the sound-symbols simultaneously. The MLAT was administered at the beginning and at the end of the school year; there were significant improvements on the aptitude test for the non-learning disabled students following the regular Latin program, and for the learning-disabled students who had the multisensory structured treatment in Latin, but not for the learning-disabled students who had the regular Latin program. The learning-disabled students who improved still scored well below mean for the non-learning disabled. This improvement in aptitude test scores may be due to the positive effect that studying Latin grammar would have on performing the type of sentence analysis involved in the Words in Sentences task. Unfortunately, Sparks et al. (1995) do not report the subtest scores for the MLAT to confirm this.

Two lab experiments made use of the Reber experimental paradigm described in Chapter 1 to investigate the issue of whether the experience of being bilingual leads to language learning 'expertise' (see McLaughlin, 1990, 1995 for general discussion of this approach). Nation and McLaughlin (1986) compared the strategies used by adult monolinguals, bilinguals, and multilinguals ($\underline{n}$ = 14 for each group) while they learned an artificial language (derived from a finite-state Markov grammar) under either an Implicit or an Explicit condition. In this study, two grammars were developed since all subjects experienced both conditions. In the Implicit condition, subjects were instructed to pay close attention to the stimuli while in the Explicit condition they were told that the stimuli were derived from a complex set of rules and that they should try and figure them out. The treatment phase was followed by a test of the rules of the grammar and an introspective questionnaire. The multilinguals were found to be significantly better at learning the rules of the language in the Implicit condition as compared to the bilinguals and monolinguals, but there was no difference in performance between the three groups in the Explicit

condition. There were no differences between the groups in their ability to articulate their learning strategies. The authors interpret these results as demonstrating that multilinguals have optimal strategies that help them to allocate their attentional resources more efficiently than monolinguals and bilinguals when dealing with learning that takes place under implicit conditions.

Nayak, Hansen, Krueger, & McLaughlin (1990) further explored differences between adult monolinguals and multilinguals in a study which involved learning a different type of artificial language, this time a phrase-structure grammar. The words of this language were CVC trigrams which, unlike the grammar used in Nation and McLaughlin (1986), were associated with abstract referents (e.g., class 'A' words referred to rectangular things). Monolingual and multilingual subjects ($\underline{n}$ = 48) were randomly assigned to either the Memory condition or to the Rule-discovery condition. The Memory group were instructed to memorize each sentence; subjects in the Rule-discovery condition were told, as in the earlier study by Nation and McLaughlin, that the order of the words in the sentences was based on a complex set of rules, and that they should try and discover the rules. Subjects were interrupted during the treatment phase and asked to verbalize what they were doing. After the treatment phase which lasted 25 minutes, subjects were given a vocabulary test and a test of syntax.

The Memory condition led to better performance on the vocabulary test but there were no significant differences between the monolinguals and multilinguals. Multilinguals in the Rule-discovery condition, however, were better than monolinguals on the test of syntax. While this appears to contradict the findings from Nation and McLaughlin (1986), the fact that the two studies involved different learning stimuli and different experimental treatments means that direct comparisons are not appropriate. In the study by Nayak et al., the analysis of the verbalizations revealed that multilinguals have a wider variety of strategies at their disposal and use them more flexibly. The authors suggest that the

multilinguals were not able to surpass the monolinguals in all aspects of the learning task because of the limited exposure to the target language.

Nature vs. Nurture

The studies described above provide conflicting evidence concerning whether L2 aptitude is influenced by language learning experiences. A few studies have demonstrated a relationship but others have not. There are two possible explanations for these mixed results: 1) problems with the research design and instrumentation, and (2) the complex nature of aptitude. With respect to the first point, we can point to methodological problems when comparisons are made between groups of individuals who differ in more than the amount of language experience; for example, the early immersion group and the late immersion groups studied by Harley and Hart (1997) not only differed in the amount of experience with the L2, but also received a different kind of L2 instruction. Research design issues also constrain the outcomes of experimental studies such as that of Nayak et al. (1990) where exposure to the target language is generally very brief. Instrument reliability and discriminatory power are frequent problems in SLA research, and the studies reviewed above are no exception. For example, Sawyer (1992) suggests that the questionnaire he used to collect information concerning language experience data from the heterogenous sample of learners was not sufficiently sensitive to discriminate among individuals.

The second point is suggested by the findings from the Bristol Project follow-up study (Skehan, 1986b) which provide us with a picture of both innate mechanisms and early environmental influences leading to the development of L2 aptitude. We can link this nature-nurture interaction to a distinction that has been made in psychology with respect to intelligence testing, and which Chapelle and Green (1992) have brought to the attention of the field of SLA. It is the difference between 'crystallized' and 'fluid' abilities (Cattell, 1987; Horn, 1989). Fluid ability reflects physiologically-based aspects of mental

functioning, while crystallized abilities are associated with education or experience. Chapelle and Green (1992) write:

> Crystallized abilities refer to learned knowledge such as language, math, mechanical information and skills, and social judgements. These all have in common their reference to knowledge gained from either school or life experience in a content area. Second-language ability, in this framework, would be considered a crystallized ability. Fluid ability, in contrast, is independent of any content area. (p. 53)

Fluid abilities are those involved in, for example, such tasks as problem-solving based on classification or analogy.

It seems likely that the reason aptitude seems to have both innate and experience correlates is that tests of aptitude, and measures of L2 learning, typically involve both fluid and crystallized abilities. For example, consider what is involved in the MLAT subtest Words in Sentences. Test-takers are required to conduct a grammatical analysis of isolated sentences disconnected from any meaningful context. They then must compare this analysis with that of a second sentence and choose the appropriate response. Although Carroll stresses the fact that knowledge of metalinguistic terminology is not required for the grammatical sensitivity test, only awareness of grammatical function, it seems unlikely that a person could perform the task without having had some kind of training in doing sentence analysis. It is not surprising, then, that Carroll (1979; 1990) found that scores on Words in Sentences correlated with a grammar test in which subjects were required to identify sentence elements in terms of named grammatical functions (e.g., subject, verb, object, etc.), and with a self-rating of knowledge about grammar (but not with the amount of training in formal grammar). Recall also that some students learning Latin in the Sparks et al. (1995) study improved significantly in their performance on the MLAT. Performance on the grammatical sensitivity measure appears to require a certain amount of crystallized ability.

Carroll (1985) suggests that tests of inductive language learning ability (such as in PLAB) are measures of more general fluid ability. However, he adds that tests designed to measure the one ability also involve the other to some degree. Thus, in Words in Sentences, inductive reasoning is involved in solving the grammatical analogies, and on the inductive tasks, performance may be enhanced by awareness of grammatical categories. Along the same lines, Chapelle and Green (1992) argue that the strategic competence proposed to underlie test-taking performance (Bachman, 1990) is a fluid ability, and is thus found in all test situations.

From the preceding discussion of aptitude, we can summarize the main points concerning grammatical sensitivity and inductive language learning ability as follows:

- they are associated with speed of learning, and thus predictors of L2 success in intensive classroom learning settings;

- they are components of aptitude that are more likely to be related to general intelligence;

- they have both fluid and crystallized dimensions.

Part II: The Ability to Focus on Form as Learner Style

Aside from aptitude, the literature on individual differences in SLA contains a number of labels that deal with the learner's ability and predisposition towards analyzing language. These appear under several different rubrics including learner type, learner style, learning style, cognitive style, and learner orientation. I will use here the term 'learner style' to describe a person's characteristic approach to language learning tasks. We will begin the discussion with learner styles that are based on Carroll's L2 aptitude model, then move on to the cognitive style contrast between field independence and field dependence, and conclude with an examination of learner orientations which have emerged from process-oriented research based mainly on the analysis of oral language data.

40

Learner Style Based on Aptitude Profiles

The componential character of Carroll's model of aptitude makes it possible to characterize patterns of learner strengths and weaknesses. Some individuals may be strong or weak in all of the components of L2 aptitude, while others may be strong in a particular component and relatively weak in the others. Skehan (1989) proposed hypothetical profiles of learners depending on the pattern of strengths and weaknesses in the components of aptitude (see Appendix B). Three empirical studies (Skehan, 1986a; Wesche, 1981; Harley & Hart, 1997) provide some evidence that such profiles are a useful way of accounting for learner differences.

Empirical Studies

Skehan (1986a, 1989) gathered data from two cohorts ($\underline{n}$ = 60 and 71) of adults studying Arabic in a ten-week intensive course in the UK. Using cluster analysis, a technique for condensing multivariate data, he was able to group learners according to similarities in their pattern of scores on a variety of aptitude measures and L2 achievement tests. These included the MLAT subtest Words in Sentences, a number of memory tests (digits, paired associates using Rumanian, the grammar rules of Indonesian, the rules of a board game, words from Finnish), and three tests of oral Arabic (conversation, interpreting, and translating); scores on the L2 tests were collapsed into one composite score.

Eight clusters were chosen for the data from the first cohort and seven for the second (cluster analysis is discussed in more detail in Chapters 4 and 5). Only three cluster profiles were readily interpretable. The analysis-oriented learners were strong on the grammatical sensitivity test but average in memory abilities. The memory-oriented learners were strong on the memory tasks but only slightly above average on the grammatical sensitivity test. Both the analysis-oriented and the memory-oriented learners were above

average on the Arabic language measure. A third type displayed average performance on the criterion measure and a relatively flat profile on the aptitude variables.

The findings of Skehan's study are limited by the fact that the aptitude-based profiles include only one measure of L2 proficiency. We do not know how the particular aptitudinal strength of each cluster relates to performance on specific L2 tasks. As Skehan admits, the groupings derived from cluster analysis are valid only to the degree that they can be interpreted by theory and research external to the analysis. In this case, Skehan relates his cluster profiles to those reported in Wesche (1981).

Wesche (1981) describes the language program of the Public Service Commission of Canada in which learners were matched to a compatible teaching approach. Students entering language training in English or French were given several placement tests including the MLAT, parts of the PLAB, and an interview. On the basis of this assessment, students were classified as having an analytic, memory-oriented, or flat profile. They were then placed in one of three types of language courses, each following a different teaching approach. In the Audio-visual course, learning activities included memorizing dialogues, doing structural drills and translation. The Analytic approach involved grammatical explanations and meaningful oral and written material, while instruction in the Functional approach was organized around interpersonal situations.

Those individuals who had a high score on the grammatical sensitivity subtest of the MLAT, but a low score on the memory subtest were placed in the Analytic class, those with the reverse profile (low on Words in Sentences but high on the memory tests) were placed in the Functional class, and those who had no clear strengths or weaknesses ("a flat profile") were placed in the Audio-visual class which was the default choice at the time. Wesche reports that such streaming led to greater student and teacher satisfaction. Indeed, when comparisons were made between analytic students placed in an Audio-visual class (i.e., mismatched) and those in an Analytic class (i.e., matched), significant differences were found in L2 achievement and in attitudes. While this streaming procedure was based

on the practical experience of the Canadian government language training program rather than an empirical investigation of learner styles, the distinction between analytic, memory, and flat profiles parallels that of Skehan's empirically-based groupings.

Finally, Harley and Hart (1997) examined the link between learner styles and the influence of different L2 instructional experiences, this time with secondary school students. Recall that the study by Harley and Hart examined whether early French immersion students would have a greater overall L2 aptitude than late immersion students. The researchers also investigated whether the different L2 learning experiences might lead to the development of different learner orientations. Specifically, they hypothesized that learners who began their learning of L2 in grade 1 would be characterized as having a memory orientation to language learning, while students who began their L2 studies later in grade 7 would be more analytical. The investigators' hypotheses were supported by the correlational and multiple regression analyses: for the early group, L2 proficiency was more likely to be associated with a measure of memory while for those learners who began L2 learning in adolescence, success in L2 was more likely to be associated with analytic language ability. These correlational findings do not tell us why this difference was found. It could mean that the timing and type of L2 instruction was responsible, or that the early vs. late immersion learners were from two different populations with a greater proportion of analytic learners in the late group. Indeed, the late immersion group was significantly higher than the early immersion group on the Language Analysis subtest but not on the two measures of memory ability.

Discussion

The different learner profiles found in these studies appear to reflect the fact that different populations are being compared. In Skehan's cluster analysis, the analysis-oriented learners were on average considerably older than the memory-oriented students (i.e., 46 vs. 27 years). Wesche (1981) mentions that the Memory group (i.e., those who

were assigned to the Functional approach) consisted of mostly female secretarial staff who were less educated and had "less overall verbal facility in English" (p. 136) than the older, mostly male professionals in the Analytic group. And finally, the late immersion students in Harley and Hart's study are likely to be a self-selected group since they have to be willing and able to deal with the demands of studying more complex subject matter in the L2. Early immersion programs, on the other hand, are more likely to draw students with a wider range of abilities and motivational characteristics. On the basis of the available evidence, one might conclude that people become more analytical with age and/or education. At present, there is no research that has investigated to what degree learners of the same age and educational experience vary in their style of learning.

Recently, Skehan (1998) has elaborated the distinction between analysis-oriented and memory-oriented learner styles by relating it to a cognitive model of L2 acquisition. Aptitude differences are related to two independent dimensions underlying language learning: degree of analysis and amount of memory (see Figure 2.1). Analytic learners are oriented towards analyzing linguistic structures which allows them to develop representations of the L2 which are rule-based. In contrast, memory-oriented learners prioritize fluency in communication which they achieve through the deployment of highly accessible lexical chunks. This type of representation is referred to as exemplar-based learning. The key difference between the two types of learners is in how L2 forms are represented in long term memory (see Table 2.2).

Both high analysis and high memory learners can be successful in L2 learning, but they will necessarily perform differently on L2 tasks depending on task requirements. Where fluency is required, the memory-oriented learner is at an advantage; where the analysis of linguistic structure is needed, then the analytic learner is favoured. Since the two dimensions of analysis and memory are independent, it is not necessary for learners to be strong in one and weak in the other. The model predicts that learners may also be high,

or low, in both analysis and memory abilities. More will be said about this framework in the overview section at the end of Part II.

Skehan's learner style framework brings together research findings from both the SLA and cognitive psychology literature. The next learner style distinction to be discussed here has also been extensively studied by both psychologists and SLA researchers.

high analysis

low memory ———————————+——————————— high memory
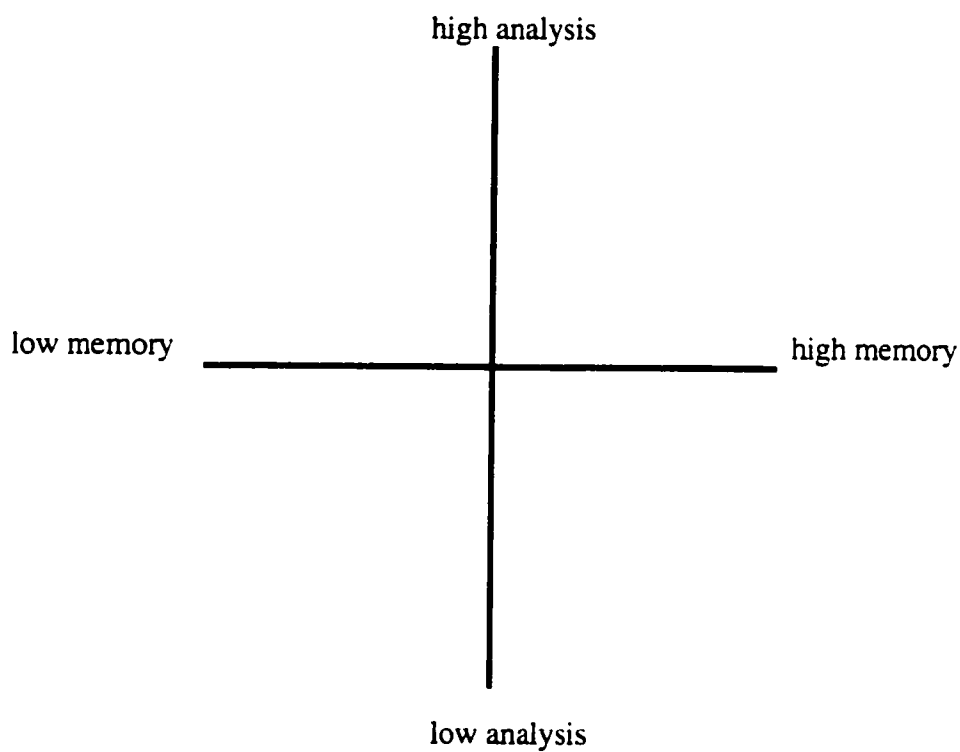
low analysis

Figure 2.1 Analytic and memory dimensions underlying language learning (Skehan, 1998)

Table 2.2

Learner Style Characteristics (based on Skehan, 1998)

	Analysis	Memory
High	• differentiated and organized rule-based representations • accuracy valued • orientation towards form	• wide range of lexicalized chunks • highly accessible representations for communication in real time • no need for complex analytic system • fluency prioritized over accuracy • predisposition to fossilize
Low	• smaller and less differentiated systems • less able to articulate about L2 system	• small repertoire of lexical elements • reliance on other resources to cope with communication in real time

(Based on Skehan (1998, p. 250)

46

Learner Style as Field Independence/Dependence

Cognitive style refers to individuals' preferences for certain modes of perceiving, processing and learning information (Oxford, 1990). The cognitive style distinction that has been investigated the most extensively with respect to L2 learning is 'field independence vs. field dependence' (FI/D). While the concept of FI/D has evolved over the years, most of the SLA research is based on the view prevalent in the 1970's:

> a bipolar, stable trait affecting how one thinks, feels and behaves. The FI person is analytic, confident, and self-reliant, whereas the FD person is holistic, uncertain, and dependent upon others.
>
> (Chapelle & Green, 1992, p. 49 based on Witkin, Moore, Goodenough, & Cox, 1977)

In a more recent formulation, FI/D includes three constructs: 'reliance on internal vs. external referents', 'cognitive restructuring skills', and 'interpersonal competencies' (Chapelle & Green, 1992; Witkin & Goodenough, 1981). The foundation of the FI/D distinction is the perceptual one between a reliance on internal vs. external referents; this is the basis for the value-neutral endpoints of the continuum. Cognitive restructuring skills and interpersonal competencies are the special abilities which are associated with each pole of this cognitive style: field independence is associated with the ability to disembed or restructure visual stimuli (i.e., cognitive restructuring ability), while field dependence is associated with skill in interpersonal relations. It is also proposed that individuals vary in the degree to which FI/D orientation is fixed; 'mobile' individuals show characteristics of both styles by being strong in analytical and interpersonal abilities (Chapelle & Green, 1992; Witkin & Goodenough, 1981).

FI/D and L2 Acquisition

Since language acquisition generally involves the development of both analytic and communicative skills, FI/D has been of particular interest to researchers who wished to

47

show how learners with diverse learning orientations can achieve success in different ways (Bialystok & Hakuta, 1994). Unfortunately, the results of such research are far from encouraging. In their review of ten empirical studies, Chapelle and Green (1992) found that generally FI/D correlates weakly, albeit significantly, with many types of language tests. However, when other cognitive abilities such as aptitude or IQ are partialed out, this relationship disappears. For example, in the study by Bialystok and Fröhlich (1978) the measure of FI correlated significantly with aptitude scores ($r = .43$) and with measures of L2 reading, writing and listening ($r = .32, .25, .21$, respectively), but when the effect of aptitude was removed, FI was no longer predictive of L2 achievement. Not surprisingly, many scholars have called for the abandonment of further research into the relevance of FI/D for L2 learning (R. Ellis, 1994; Skehan, 1989; for lively debate on this topic see Griffiths & Sheen, 1992; Chapelle, 1992; Sheen, 1993).

Part of the problem with evaluating the usefulness of the FI/D distinction for SLA is that almost all of the studies investigating the relationship between this cognitive style and L2 learning have made use of an Embedded Figures Task (EFT). This task requires the subject to find a simple geometric figure within a more complex cluster of figures. As Chapelle and Green (1992) point out, the EFT appears to measure only the cognitive restructuring ability associated with field independence. Cognitive restructuring ability means that the individual is able to disembed or restructure visual stimuli, to break up a pattern into its component parts, and to make judgments about them. Although a low score on the EFT is interpreted as reflecting the other end of the continuum, there is usually no independent measure of field dependence via the interpersonal skills ability associated with it. This is problematic since a cognitive style is defined as bipolar, while measures of abilities are not; a low score on a test of ability does not imply anything but a lack of that ability. Thus, the fact that one person did poorly on an EFT does not indicate what kind of abilities associated with field dependence that person might have.

48

In some studies a measure of interpersonal skills in the L2 is included in the research design as one of the criterion variables. For example, Johnson and Rosano (1993) studied the relationship between FI/D in foreign students at a Toronto university and their communicative ability in English. In this case, FI/D was measured using the Block Designs test from the Wechsler Adult Intelligence Scale-Revised in which the subject uses coloured blocks to reproduce pictured models; like the EFT, this task requires mental decomposition of the stimulus image into its component parts. This study was unusual in that the researchers actually found a substantial negative correlation (r = -.57) between performance on the Block Designs Test and a teacher rating of L2 communicative ability. Such a finding, however, is not very informative on its own -- why should individuals who score poorly on a visual restructuring task be rated by their teachers as having greater ability to communicate in specific situations (or vice versa)? Some measure of interpersonal skill would facilitate interpretation.

Cognitive Restructuring Ability

While the final word concerning the relevance of the FI/D construct for SLA awaits further research in which both ends of the cognitive style continuum are measured, it is precisely cognitive restructuring ability that appears to be most relevant to the present discussion. Nonetheless, we are still left with the fact that the EFT has only been shown to predict L2 test performance in some studies, and only on certain types of measures such as the cloze (Chapelle & Green, 1992). One might argue that this is to be expected from a visual-spatial task which does not involve the use of language. Other explanations, however, have been offered by Bialystok (1992) and Chapelle and Green (1992).

Bialystok (1992) explains the fact that field independence predicts some but not all kinds of L2 performance in terms of a specific model of language proficiency (Bialystok & Ryan, 1985a, 1985b). The model describes language performance in terms of two independent dimensions: analysis of knowledge and cognitive control. The first refers to

the learner's development of analyzed knowledge of the structure of language, while the second dimension has to do with attentional control over cognitive processing. (We will return to this framework in more detail in Part III of this chapter.). Bialystok (1992) investigated the relationship between field independence and L1 metalinguistic performance of seven to nine year old anglophone children. She found that two measures of field independence, one of which was an EFT for children, correlated with nonverbal intelligence but not with a test of verbal intelligence. Field independence also accounted for a significant portion of the variance on some, but not all, of the metalinguistic tasks when verbal IQ was partialed out. She argues that the measures of field independence, of nonverbal intelligence, and of certain types of metalinguistic tasks are similar in that they all make demands on the subjects' ability to attend to stimuli selectively. This reflects the dimension of 'cognitive control' in the Bialystok and Ryan model of language proficiency. No such relationship was found on the metalinguistic task conditions which required analyzed knowledge, the other dimension of the model. On the basis of these findings, Bialystok proposes that the performance of certain types of metalinguistic tasks depends on more general cognitive processes, the same that underlie performance on the EFT and on the test of nonverbal intelligence, rather than on language-specific processes. This study examined children's metalinguistic performance in their L1, but according to Bialystok, the same holds for the L2. Field independence will predict L2 performance on tasks which make similar demands on selective attention; L2 tasks that also require high levels of analyzed knowledge will display a less straightforward relationship with FI.

Along similar lines but using a different framework, Chapelle and Green (1992) conclude that cognitive restructuring ability is a fluid rather than a crystallized ability (Cattell, 1987; Horn, 1989). Recall that this distinction contrasts physiologically-based aspects of mental functioning with abilities gained from educational or life experience. In two studies, Chapelle (Chapelle, 1988; Chapelle and Green, 1992) has examined the influence of field independence on test performance by administering a cloze test in English

to native speakers. By doing so, it is argued, variance from crystallized language knowledge is minimized, thus allowing the variance from fluid ability to be more apparent. In both studies, performance on the EFT accounted for a significant amount of variance in performance on the cloze. In Chapelle and Green's view, fluid ability is involved in any task, and thus may be equated with the 'strategic competence' found in models of L2 proficiency (Canale & Swain, 1980; Bachman, 1990). Strategic competence refers to "non-subject-specific abilities needed to put language knowledge to use" (Chapelle & Green, 1992, p. 70). This linking of field independence to strategic competence is compelling since it brings together notions from SLA and from language testing. However, in Bachman's model strategic competence has a central role in orchestrating language performance; the degree to which this is true of field independence has yet to be established.

Field Independence and Aptitude

Chapelle and Green also argue that restructuring ability as a fluid ability should be part of a model of aptitude. Chapelle and Green propose that Skehan's 'language analytic ability' should be divided into two components: 'language analytic ability' and 'general analytic ability'. The first would be content-related, and thus reflect linguistic training and the experience of learning languages; the general analytic ability would be independent of any learned knowledge. In essence, they are arguing for a return to the four-part model of aptitude as proposed by Carroll, distinguishing between a relatively more crystallized grammatical sensitivity component and an inductive language learning ability component which is relatively more fluid (Carroll, 1985). According to Chapelle and Green, such a distinction within a model of aptitude clarifies the relationship between aptitude and intelligence. Skehan (1991) also has suggested that field independence might be another label for the language analytic ability component of his aptitude model. Given our present

understanding of the evidence, it appears to be appropriate, and parsimonious, to consider field independence as an aspect of L2 aptitude.

Similar style distinctions have been proposed in L1 and L2 acquisition research based on longitudinal case studies involving the qualitative analysis of production data. For example, Skehan (1989) links his aptitude-based learner types derived from quantitative methods to the learner styles based on qualitative analyses described by Peters (1981). In the following sections we will examine the relevant L1 and L2 literature based on this naturalistic sampling research methodology.

Learner Style Based on Linguistic Analyses

L1 Studies

Although much of the research into child language development has focussed on the universal aspects of acquisition, a number of studies have revealed the presence of qualitatively different approaches taken by children in acquiring their mother tongue (see Bates, Dale & Thal, 1995; Goldfield & Snow, 1985; Wells, 1986 for reviews of this literature). For example, in early lexical development, Nelson (1973) found that some children were more referential, apparently focussing on learning words for objects, while those children she called 'expressive' had fewer object labels and vocabularies containing items from a variety of form classes, including some frozen phrases like 'Stop it' or 'I love you'. In Nelson's longitudinal study, referential children reached the 50 word stage earlier than those who had a more heterogeneous vocabulary; thus referential style was correlated with rate of vocabulary acquisition. This referential/expressive style dimension has been confirmed by other studies (e.g., Ramer, 1976; Horgan, 1978). Snyder, Bates and Bretherton (1981) studied 32 children and found that the referential/expressive distinction was evident from the very beginning of language acquisition, and that it applied to both comprehension and production; they too found that referential children were more advanced

52

in their lexical development. Furthermore, they found that the precocious referential children were more likely to use words in a context flexible way, such that a word like 'doggie' was produced to refer to any dog, in a variety of contexts, rather than only when in the presence of the family dog.

This distinction between an early nominal style in lexical development and a style characterized by heterogeneity finds a parallel in the earliest stages of syntactic development. Bloom, Lightbown, and Hood (1975) found that children used one of two approaches to word-combining at the two-word stage in order to express the same set of semantic relations: a nominal approach that was characterized by telegraphic speech in which they combined content words (e.g., 'Kathryn sock'), or a pronominal approach whereby children combined a small class of function words with content words (e.g., 'I finish'; 'my sock'). Nelson (1975) showed that children who were considered to be referential at the one-word stage were more likely to use a nominal approach in their early grammar, while expressive children preferred the pronominal style.

The expressive/pronominal style has been linked with imitation, but this finding has been both supported and contradicted. For example, in the study by Bloom, Hood, and Lightbown (1974), one 'imitator' was pronominal while another was nominal. Bates, Bretherton, and Snyder (1988) argue that expressive children's imitative responses often contain forms that have not appeared in their spontaneous utterances, something which is not true for referential/nominal children. Other features associated with the expressive style are less consistent use of word order at the telegraphic stage and less frequent production of morphological overgeneralizations than referential/ nominal children (Bates et al., 1988, p. 48).

Several studies have reported the speech of expressive children as difficult to understand. Bates et al. (1988) point out, it is often the children whose speech is clear and easy to transcribe that are selected for study. This suggests that the nominal/referential child is over-represented in the literature. Peters (1981) argues that a child's speech may be

unintelligible to the researcher because the child is using a different unit of analysis than expected. This was true of the child she studied. Minh appeared to approach the task of acquiring language using what Peters called a 'holistic' or 'gestalt' strategy in contrast to the typical, 'analytic' approach that had been documented in other children. That is, she tried to approximate the sentence level features of the language in order to engage in communication rather than trying to extract smaller units from the input, working on them bit by bit. Minh appeared to focus on the prosody of whole phrases rather than on the segmentation of words.

Bates and her colleagues have attempted to synthesize the findings of style differences in different domains (semantics, grammar, and phonology) that have been reported in different studies into what they call 'a two strand model' (Bates et al, 1988, 1995):

Infants who are word oriented during the babbling period grow into children with referential vocabularies at the first word level, and into children who display nominal style in their first word combinations, followed by high rates of morphological overgeneralization during the acquisition of grammar. Conversely, children who are intonation oriented during the babbling period grow into children with expressive vocabularies in the one word stage, and then show a formulaic, pronominal style in first word combinations, followed by a pattern of grammatical learning characterized by undergeneralization and inconsistent application of rules. At the highest level of generality, these styles have been labelled 'analytic' v. 'holistic' or 'rote', respectively. (Bates et al., 1995, p.12-123)

A number of possible explanations have been offered to account for the style difference: these range from social factors such as maternal style and input to neurological factors such as hemisphere dominance, and cognitive factors such as intelligence (Bates et al., 1988). More recently, however, Bates et al. (1995) have "soberly reconsidered" their grand synthesis and suggest that the difference between the analytic and the holistic style may be a

matter of what size unit children prefer to process, a small one for the analytic child or a larger one for the holistic child. In both cases, the process of learning is the same.

L2 Studies

There are clear parallels between the kinds of stylistic differences reported in L1 acquisition and those that have been proposed in the context of L2 acquisition. The distinction proposed by Hatch (1974) between 'rule-formers' vs. 'data-gatherers' is very often quoted in the SLA literature. In this paper, Hatch discusses issues relating to interlanguage development based on the language data presented in 15 observational studies involving 40 child and adult L2 learners. She comments that some learners, the rule-formers, appear to move from developmental stage to developmental stage in an orderly fashion, producing utterances which are easy to classify within their systems. In contrast, other learners, referred to as data-gatherers, have language systems less easily characterized by a discrete set of rules; they use new forms and become more and more fluent, but without having necessarily sorted out the rules of the L2. There is an obvious similarity to the referential/nominal and expressive/pronominal children described in the L1 literature. One might speculate that young L2 learners carry over the same approaches to acquisition from L1 to the learning of L2. As we shall see, other distinctions, sometimes analogous, have been proposed by researchers as a way of accounting for different patterns of development in L2 acquisition.

Child learners of English in a naturalistic setting.

Wong Fillmore (1979) points out that the major dimension of individual differences in L2 acquisition is speed of learning: some learners are slow and laborious in acquiring the L2 while others pick it up quickly and easily. In contrast to Carroll, Wong Fillmore explains differences in rate of acquisition in terms of social and cognitive strategies that

support L2 learning. This conclusion is based on the analysis of data collected over the course of the school year from five Spanish-speaking children acquiring English at school in grade 1. Each child was paired with an English-speaking child and observed during weekly play sessions which were audio-recorded, transcribed and analyzed with respect to interlanguage development.

All of the children demonstrated progress in their acquisition of English during the school year, but their progress varied enormously. The most successful, that is, the one who learned most rapidly, was a girl named Nora who made use of a wide range of social and cognitive strategies. She was highly motivated to be a member of the L2 social group, spending more time than the other four children in the company of monolingual English speakers. She tended to be playful and experimental with language, and took risks in using newly acquired expressions. She was particularly adept at analyzing formulaic expressions. According to Wong Fillmore, this is one of the most important of the cognitive strategies for the development of L2 grammar.

All the children made use of unanalyzed formulas such as "Lookit"; "Wait a minute"; "Gimme", etc. Strategic use of such highly situational formulae made it possible for the learners to continue to participate in activities which would provide further input and output opportunities. Wong Fillmore concluded that the value of these formulae lies in the fact that they "constituted the linguistic material on which a large part of the analytical activities involved in language learning could be carried out" (p. 212). The process of analyzing these chunks into their recurring parts is the way these children acquired the grammar of English. Nora was particularly quick at figuring out which parts of the formulaic expressions could be varied and at decomposing them. Interestingly, Wong Fillmore notes that Nora also enjoyed figuring out how objects worked, pulling them apart and putting them back together. This behaviour contrasts with that of Jesus, one of the least successful learners in the study, who appeared disinclined to analyze his formulae and

tended to maintain them in the forms he learned them in, a tendency consistent with a general rigidity in his thinking.

Adult learners of German in a naturalistic setting.

There are parallels to the distinction between learner types described in the findings from the Zweitspracherwerb Italienischer und Spanischer Arbeiter (ZISA) project (Clahsen, Meisel, & Pienemann, 1983; see Meisel, 1983; Meisel, Clahsen & Pienemann, 1981 for English-language descriptions of the project). The project involved a cross-sectional study of 45 adults and a two-year longitudinal study of 12 learners. The participants were Spanish, Italian, and Portuguese 'guest workers', men and women with little formal education who had moved to Germany for economic reasons, and were acquiring German in a naturalistic rather than a classroom setting. Informal interviews provided the data for interlanguage analysis, as well as making it possible to gain information concerning psycho-social factors such as job activities, contact with German at work and after work, leisure activities.

The analysis of the syntactic development of these learners led to the elaboration of the Multidimensional Model in which the acquisition process is viewed as consisting of a developmental and a variational dimension (see Hudson, 1993; Larsen-Freeman & Long, 1990 for critiques of the model). According to Clahsen, Meisel and Pienemann the developmental dimension is common to all learners of a second language as it is determined by universal processing constraints. However, learners do follow different paths for the so-called 'variational features' of the L2 and these paths are determined by the learner's orientation. Learner types were first derived from the analysis of syntactic simplification strategies: some learners were norm-oriented, favouring accuracy, while others tended to simplify their utterances; of those who simplified, the ZISA Project researchers argued, some simplified but avoided deviation from the target norm, while others simplified in order to communicate more efficiently. The first type of simplification is referred to as

57

'elaborative' and the second as 'restrictive' simplification. The ZISA group looked for an explanation for the observed patterns of variation in the psycho-social data they had also collected. Using a variety of statistical analyses, they found that learners could be classified with respect to their degree of integration into German society. They concluded that elaborative simplification was more common among those with an integrative orientation while those who made frequent use of restrictive simplification tended to have a segregative orientation. (Note that Hudson, 1993 criticizes the statistical analysis of the psychosociological data, but see the response by Pienemann, Johnston, & Meisel, 1993).

Child L2 learners of German in a school setting.

Clyne (1985) investigated the L2 learning of children in two types of German as a second language program in Australia. One school offered L2 instruction in which German was used as a medium of instruction for part of the school day, while in the other school German was taught as an object of instruction. The evaluation of the learners' German included the administration of different types of tests (i.e., listening comprehension, cloze, reading, dictation, translation). Oral language samples were also collected. The analysis of the oral data led Clyne to distinguish between two types of learner orientations: a 'structural orientation' and a 'functional orientation'[2]. The structurally-oriented children were characterized by a genuine interest in how the German language works, and were more successful at keeping their languages apart. On the other hand, the functionally-oriented learners clearly assumed that all German speakers in Australia could speak English and thus felt it appropriate to code-switch into English. These students had strong comprehension skills in German and well-developed communication strategies. According to Clyne, structurally oriented learners will be distinguished from those with a more functional orientation by an avoidance of code-switching (although lexical L1 transfer will occur), and by a relatively greater amount of self-correction behaviour. He suggests that the

functionally-oriented learners are better suited to the meaning-focussed instruction of an immersion program rather than to form-focussed instruction.

Child and adult learners of Dutch in a naturalistic setting.

Finally, we turn to the study by Snow and Hoefnagel-Höhle (1979) which although departing from the research methodology used in the case studies above, provides empirical support for 'stylistic' differences in L2 acquisition. They studied 47 English speakers of different ages (from age 3 to adult) who were learning Dutch in largely naturalistic settings. As in the case studies, however, the study was longitudinal and focussed on the developing L2 abilities of the learners. Language abilities were assessed using measures of pronunciation, auditory discrimination, elicited imitation, grammaticality judgments, story retells, translation task, a Dutch language 'wug test' (based on Berko, 1958), story telling, and tests of active and passive vocabulary. Testing took place three times during the first year to year and a half of the subjects' residency in Holland. The data from each session were analyzed using factor analysis, but only at the final session did a distinct and interpretable factor structure emerge. It was found that pronunciation and auditory discrimination ability were clearly differentiated from grammar and vocabulary, with some learners very high on one factor but low on the other. Snow and Hoefnagel-Höhle suggest that these represent two separate components of L2 ability, and that some learners may be characterized primarily by their phonological abilities, while others are characterized by their ability to acquire syntactic and morphological rules. Interestingly, this distinction between grammatical/vocabulary vs. phonological ability was found to be true even for very young L2 learners in their sample.

Snow and Hoefnagel-Höhle recommend that research into individual differences in SLA should include both quantitative and qualitative research approaches. Indeed, confirmation for the components of L2 ability that they discovered in the study of English-speaking learners of Dutch is provided by Wong Fillmore's Hispanic children in

California.[3] Jesus, who was perhaps the least successful learner of the five, liked to play with sounds and was an excellent mimic; Wong Fillmore states that "he had a talent for remembering and precisely reproducing formulaic expressions that he heard" and "a talent for recognizing and picking up memorable and useful expressions" (1979, p. 226). On the other hand, he showed very little progress in analyzing formulaic expressions into their component parts. In contrast, Alej, another of the boys studied in this project, had great difficulty remembering and reproducing expressions, and his pronunciation was often unintelligible. Yet he learned more English in the year of the study than Jesus despite this 'deficit' because he was better at using other learning strategies, particularly the one relating to grammatical development via the analysis of formulae.

Overview of Learner Styles in SLA

An overview of the learner styles discussed in this chapter is presented in Table 2.3. At this point it is probably wise to consider the advice of Snow and her colleagues (Goldfield & Snow, 1988; Snow & Hoefnagel-Höhle, 1979) who warn against the temptation to equate learner types that have been identified in different types of research. Can we say the data-gatherer is field dependent while the rule-former is field independent? Is an analytic child like Nora doing the same thing as the analysis-oriented adults in the Arabic course described by Skehan? We must beware of making connections between different phenomena simply on the basis of a shared label. A case can, however, be made for making connections between some of the terms in the left-hand column because of the findings of the Bristol follow-up study (Skehan, 1986b, 1990). The longitudinal data provided by this study provides strong evidence for relating differences in mother tongue acquisition to later performance on aptitude tests, although the reasons for the association are not so clear. Further research is needed to clarify the connection between the implicit analysis of the child and the conscious analysis of the older learner.

Table 2.3

Learner Style Distinctions in L1 and L2 Acquisition Research

X	Y
analytic (referential/nominal)	gestalt/holistic/rote (expressive/pronominal)
analysis-driven	memory-driven
field independent	field dependent
rule-former	data-gatherer
standard-oriented (integrative)	restrictive simplification (segregative)
structure-oriented	function-oriented
grammar/vocabulary	phonology

Note. The visual display of the two learner style poles as X and Y is borrowed from Reid (1995, p. xiv).

Relatively little is known about the learner styles in the right-hand column of Table 2.3. Skehan (1998) has proposed that learners vary with respect to two independent dimensions, degree of analysis and amount of memory. Some learners prioritize accuracy while others emphasize fluency. This in itself is not new. The innovation is that this difference is linked to how these individuals learn and the type of knowledge they come to attain. That is, prioritizing communicative fluency over grammatical accuracy is connected to a greater reliance on memory abilities, and is also reflected in a tendency towards exemplar-based representations of knowledge. Skehan provides many different types of evidence in support of this framework, but the relevant SLA evidence is noticeably thin. Further investigation of the predictions made by this model needs to be carried out.

Using Skehan's framework, we can stop viewing learners from the perspective of a style dichotomy. The cognitive processes that underlie L2 acquisition are the same for all learners: they have to notice forms in the input, store them in long term memory, and retrieve them from memory for use in production. All learners have to use memory in order to acquire and use language; calling one learner style memory-oriented is somewhat misleading. Skehan's work suggests that the reason why one learner creates parsimonious rule-based representations of the L2 while another makes do with lexicalized exemplars has to do with some kind of predisposition. An alternative approach is to view the analytic learner as being at a more advanced stage of development than the non-analytic learner. This latter view of language acquisition is part of Bialystok's model and will be discussed in Part III of this chapter.

Part III: The Ability to Focus on Form as Metalinguistic Skill

Developments in cognitive psychology over the last few decades have led to a move away from viewing cognitive abilities as innate traits to seeing them as part of an information-processing model. From this perspective, the ability to perform a task like the MLAT subtest Words in Sentences, which requires the individual to focus on grammatical form, is a result of the development of a skill, specifically, metalinguistic skill. In general, this term is defined as the ability to look at language as an object. Ryan and Ledger (1984, p. 157) characterize metalinguistic development as:

> the gradual shift of attention from meaning to structure in tasks requiring deliberate control over language forms. The essential feature of the beginnings of linguistic awareness seems to be flexibility of strategy -- the ability to decenter, to shift one's focus from the most salient attribute of a message (its meaning and contextual

setting) to structure (the ordinarily transparent vehicle by which meaning is conveyed).

There is a vast body of research on this topic, and discussion here will deal with a few important points (see Birdsong, 1989 for a synthesis of both the L1 and L2 literature).

Metalinguistic Development

It is clear that metalinguistic ability emerges only after a child has attained some ability to use language in context; for example it is not until children have reached the age of 6-7 that they are able to find humour in lexical ambiguity. We often think of metalinguistic performance as relating to such tasks as grammaticality judgments or sentence parsing. However, it is best considered as involving a continuum of abilities ranging from "sporadic insights into aspects of language" at one end, to the ability to generate the elaborate metalinguistic explanations that linguists enjoy at the other (Sharwood Smith, 1991a, p. 20). In the L1 literature, a wide range of behaviour is classified as being metalinguistic. For example, in 1978, Clark proposed the taxonomy presented in Table 2.4.

The different types of awareness are presented in an order that roughly mirrors their sequence of emergence, from most to least basic. It is the sixth category pertaining to the ability to reflect on language which will be referred to here as metalinguistic skill. While this heading covers all levels of linguistic analysis, that is, the ability to reflect on phonological, morphological, syntactic, lexico-semantic and pragmatic aspects of language use, discussion will focus on the development of metalinguistic skill with respect to grammatical structure.

Table 2.4

Taxonomy of Metacognitive Skills and Awareness of Language (Clark, 1978)*

1. Monitoring one's ongoing utterances	• repairs own speech spontaneously • practices sounds, words, sentences • adjusts speech to the age, status and language of the listener
2. Checking the result of an utterance	• sees whether the listener has understood or not • comments on utterances of self and others
3. Testing for reality	• decides whether a word or description works or not
4. Deliberately trying to learn	• practices new sounds, words, sentences • does role playing
5. Predicting the consequence of using inflections, words, phrases, or sentences	• applies inflections to 'new' words out of context • judges out of context which utterance would be more appropriate for a specific speaker
6. Reflecting on the product of an utterance	• identifies linguistic units • provides definitions • constructs puns and riddles • explains why certain sentences are possible and how they should be interpreted

Note. This table is slightly adapted from Clark, 1978, p. 34.

Individual Differences in Metalinguistic Ability

As is the case with primary language ability, metalinguistic skill emerges as a function of age such that a seven year old is generally able to do what a five year old cannot. Gleitman, Gleitman, and Shipley (1972) found that five year olds in their study

could make judgments about sentences that were implausible or semantically odd, but did not notice syntactic violations that did not affect meaning. For example, "The colour green frightens George" was rejected, but "John and Bill is a brother" was accepted, even though the children themselves demonstrated control over subject-verb agreement in their spontaneous speech. As one might hope, the reverse pattern of responses was true for older children. Thus, there appears to be a developmental pattern such that children are able to reflect on the semantic aspects of utterances before they are aware of the syntactic aspects. It is also the case, however, that considerable variation is evident in metalinguistic performance among children of the same age. For example, in a study of children's ability to interpret ambiguity by Kessel (1970), some of the seven year olds outperformed some of the ten year olds. Van Kleeck (1982) accounts for such differences in terms of a model in which endogenous variables within the child interact with exogenous variables in the environment. She hypothesizes that a child's creativity and verbal intelligence will be predictive of metalinguistic skill as will such environmental factors as bilingual input, and parental valorization of literacy. Conversely, the model also predicts that some children who possess the cognitive constructs necessary for metalinguistic performance may still experience delays in metalinguistic development due to environmental deficits.

Differences in metalinguistic performance are not limited to the early years of life -- adults also differ in their metalinguistic performance. Thus, individual differences are found in both rate of development and in ultimate attainment. Investigation of the latter is reported in Gleitman and Gleitman (1970, 1979). They found that adults vary enormously in their ability to think about and comment on syntactic novelty as opposed to semantic novelty. Adult native speakers of English were asked to produce and recognize paraphrases of novel compound nouns. The compounds were presented with different stress contours ('bird-house black' vs 'bird house-black') which determine the appropriate interpretation of compounds when presented without a context. A group of clerical workers and a group of Ph.D. candidates were tested in their ability to paraphrase these compound nouns.

Surprisingly, the two groups differed "massively" in their performance; a closer examination of the results revealed that the group differences were largely due to subjects' responses to the compounds that presented syntactic rather than semantic oddities. The group of clerical workers focused on meaning and plausibility, while the highly educated group were able to focus on the syntax even when meaningfulness was sacrificed.

Gleitman and Gleitman (1979) conclude that not everyone can focus attention on a syntactic anomaly and perform an appropriate syntactic manipulation to repair it; this is so even when a speaker is in productive control of the construction for normal communicative use. Meaning is more readily brought into conscious attention. This parallels the developmental pattern of metalinguistic judgment abilities mentioned earlier where we find that children are able to make judgments concerning meaning before they can explicitly comment on the syntactic form of utterances.

The Relationship Between Primary Language Skills and Metalinguistic Abilities

Based on their studies of metalinguistic judgments in both children and adults, Gleitman and Gleitman (1979) have concluded that the ability to use language communicatively is independent of the ability to reflect on language usage. They argue that the ability to make metalinguistic judgments is "a single example of a more metacognitive organization in humans" (1979. p. 121) and thus has more in common with other metacognitive activities than with the lower level linguistic processes involved in language production and comprehension. Consequently, metalinguistic judgments are not necessarily direct windows on language competence.

Not everyone agrees with Gleitman and Gleitman that metalinguistic skill is independent of primary language skills. Some see primary language skills and metalinguistic ability as being interdependent in some way, either showing a unidirectional, causal relationship, or a reciprocal relationship. Others hesitate to ascribe causation to

correlated behaviour and propose a synchronous or continuous relationship. (For discussion of these views see Tunmer, Pratt and Herriman, 1984 and Van Kleeck, 1982, 1984). Van Kleeck (1982) argues that it is difficult to establish unequivocal evidence for either the independent or an interdependent position. She points out, for example, that Gleitman and Gleitman did not actually measure primary language abilities in their child (Gleitman, Gleitman, & Shipley, 1972) or adult subjects (Gleitman & Gleitman, 1970), while, on the other side, some researchers (Smith & Tager-Flusberg, 1980; Saywitz & Wilkinson, 1982) have measured language production or comprehension skills using tasks that could be considered decontextualized (e.g., Peabody Picture Vocabulary test) which could account for the correlations found between such measures and performance on metalinguistic tasks.

To better understand the relationship between primary language abilities and meta-level abilities, we need to examine the factors that promote the growth of metalinguistic skill. Two factors have been the focus of research: exposure to bilingual input and the acquisition of literacy skills.

The Influence of Early Bilingualism on Metalinguistic Abilities

The following anecdote from de Villiers and de Villiers (1978, p. 151) illustrates the oft-cited superiority of bilingual children over monolinguals with respect to language awareness.

Emir (4 yrs., bilingual):	I can speak Hebrew and English.
Danielle (5 yrs., monolingual):	What's English?

The relationship between early bilingualism and precocity in metalinguistic development is widely reported (e.g., Ben-Zeev 1977; Cummins & Mulcahey, 1978; Ianco-Worrall, 1972; Leopold, 1939-49). For example, in a classic study, Ben-Zeev (1977) compared the performance of children bilingual in Hebrew and English and groups of monolingual English and monolingual Hebrew children on (among others) a symbol substitution task. The task involved substituting one meaningful word for another, for example the word

'macaroni' for the pronoun 'I'. The child is required to produce sentences like 'Macaroni am warm' or 'The doll is going clean the house' (where the word 'clean' is substituted for the preposition 'into'). As predicted, the bilingual children were significantly better than the monolinguals at performing this task. It appears that since both languages of the bilingual provide words for the same objects "the magical idea that there is a determined relationship between a word and its referent becomes dissipated very early" (p. 1012). The bilinguals in this study were also superior in their ability to recognize verbal transformations and to make verbal descriptions of non-verbal transformations. Ben-Zeev concludes that bilingual children have a greater readiness to search for structure and to reorganize their perceptions in response to feedback.

It has also been found that the benefits of bilingualism are not necessarily limited to situations where both languages are spoken in the home -- even a very limited exposure to another language can have positive effects. This has been shown by Yelland, Pollard, and Mercuri (1993). They studied the acquisition of reading skills in Australian children in grade 1 who received one hour a week of instruction in Italian. After six months of limited contact with the L2, these 'marginal' bilinguals were found to develop word awareness more quickly than their monolingual peers. Although this gain disappeared over the course of the year, the marginal bilinguals outperformed monolinguals on measures of written word recognition at the end of the school year. The researchers conclude that the metalinguistic benefits of bilingualism are not necessarily dependent on some critical degree of competence in the L2. This position contradicts the hypothesis that individuals must attain a threshold of ability in the L2 in order for the cognitive benefits of bilingualism to manifest themselves (Cummins, 1987).

While many studies have shown advantages in metalinguistic performance for bilingual children when compared to monolingual children, the effects are not evident across all metalinguistic tasks. For example, Bialystok (1987) has shown that the bilingual child's advantage in metalinguistic development relates to enhanced processing control

rather than to linguistic knowledge. (We return to discussion of her model of proficiency later.) Furthermore, it is not clear whether the advantage is merely in rate of acquisition rather than in long-term effects.

The Role of Literacy in the Development of Metalinguistic Awareness

Although there is general agreement that literacy and metalinguistic ability are related, the nature of this relationship is a matter of debate. Some see metalinguistic ability as a factor in literacy development, others see literacy as a necessary condition for metalinguistic ability to develop, while still others argue for a reciprocal role for literacy. Recently, Olson (1994, 1996) has presented a convincing case for viewing our metalinguistic awareness not as a precursor but as a consequence of learning to read an alphabetic script. Traditional wisdom considers writing to be a transcription of speech, and the historical development of Western writing systems has been depicted as a linear progression from picture writing to word-based writing systems, to sound-based syllabic writing systems, culminating in the Greek invention of the alphabet. According to Olson, this view assumes that the inventors of alphabetic writing already knew that language could be analyzed into the structural categories of sentence, word and phoneme. Such an explanation assumes the very thing that needs to be explained -- where did this knowledge of the structure of language come from? The answer offered by Olson is that it was the historical accident[4] which led to the creation of the alphabet that made it possible for the Greeks to "hear" their language in a new way. Early writing systems were not attempts at transcribing speech but rather attempts at unambiguously representing meaning:

> writing far from transcribing speech tends to provide a model for that speech. To invent a writing system is, in part, to discover something about speech; to learn to read is, similarly, to discover something about one's speech, and ultimately, about 'what is said'. The script provides the model, however distorted, of one's speech. (Olson, 1994, p. 78)

69

Children in learning to read retrace this historical development to some extent. While preliterate children bring to the task of learning to read certain kinds of metalinguistic knowledge, it is the alphabetic script that provides the model which allows phonological form to be represented and thus brought into consciousness; "the models of language provided by our scripts are both what is acquired in the process of learning to read and write and what is employed in thinking about language: writing is in principle metalinguistics" (Olson, 1994, p. 89). Olson argues that people who have learned an alphabetic script come to hear words as composed of the sounds represented by the letters of the alphabet while those unfamiliar with an alphabet do not. Thus, one would predict differences in metalinguistic abilities between literates and illiterates, and also among literates who acquire different kinds of writing scripts. Evidence in support of the first prediction is provided by, for example, Karanth and Suchitra (1993) who showed that the ability to perform a metalinguistic task such as making grammaticality judgments is directly tied to the acquisition of literacy skills. The second prediction is confirmed by the findings of Read, Zhang, Nie, and Ding (1986) that Chinese readers of traditional character scripts cannot detect phonemic segments, whereas those who can read the alphabetic script called 'Pinyin' are able to do so.

If literacy in an alphabetic script has a direct influence on the development of metalinguistic skill, one would predict that literacy in two languages would have a powerful influence on metalinguistic abilities. Indeed, Bialystok (1987) found that biliterate children outperformed monoliterate bilinguals on metalinguistic tasks. Further support for the effect of biliteracy is offered by Thomas (1988) who compared English monolinguals to Spanish-English bilinguals in learning French at the university level in the US. The bilinguals significantly outperformed the monolinguals on tests of French grammar and vocabulary, and on measures of comprehensibility of written compositions. However, the bilinguals who had received formal instruction in Spanish and English were significantly better than the monoliterate bilinguals on the two discrete-point French tests; they also attempted more

structures and made fewer errors on the compositions. The author concludes that the bilinguals, particularly the biliterate bilinguals had enhanced metalinguistic skills which were helpful in the form-oriented teaching approach used in the French class.

While the existence of individual differences in metalinguistic skill among adult native speakers is beyond dispute, little research or theory has been mounted to account for this variation. It would seem likely that here too we are dealing with an interaction between endogenous and exogenous variables. The main focus of research in the metalinguistic skill development paradigm has been on accounting for variation in performance on language tasks, particularly with school-aged children. We will focus the discussion here on two models which have been very influential in SLA, those of Bialystok (Bialystok, 1991, 1994a; Bialystok and Ryan, 1985a, 1985b) and of Cummins (1987, 1991).

Bialystok's Model of Metalinguistic Ability

The Analysis/Control model associated with Bialystok was briefly described in the section on cognitive style in Part II of this chapter. The model itself has been modified over the years; we begin with the version presented in Bialystok and Ryan (1985a, 1985b). The theoretical framework proposed by Bialystok and Ryan accounts for differences in performance on conversational vs literacy vs metalinguistic tasks. They argued that metalinguistic skill is not unitary but rather composed of two constituents: the dimension of analyzed knowledge and the dimension of cognitive control. Variability in performance on all types of language tasks, not just metalinguistic ones, can be understood with reference to these two independent dimensions. With respect to the first, knowledge of the language system (grammar, semantics, phonology, and discourse) varies in terms of the degree to which it is analyzed. Unanalyzed language is conceptualized as a routine or pattern that has limited application to new contexts. As language forms become more analyzed, the learner is able to deliberately use the systems for specific purposes. The second dimension, cognitive control, involves the selection and coordination of information, and the

automaticity with which the selection and co-ordination operations can be executed. This dimension is thus described in terms of attention allocation and of automaticity in processing.

More recently, Bialystok has moved away from modelling language ability in terms of knowledge vs process typical of her earlier work (Bialystok & Ryan, 1985a, 1985b; Bialystok & Sharwood Smith, 1985; see Hulstijn, 1990 and Bialystok, 1990 for discussion of the evolution of this model.) Language performance is now seen as reflecting two processing components, the process of analysis and the process of control. The first component refers to "the progressive analysis or restructuring of mental representations of language" (Bialystok, 1991, p. 117) while the second concerns attention to these representations. Control is defined in terms of the process of selecting the information that will be attended to in real-time solutions to problems. Automaticity is now considered to be an outcome of selective attention during on-line processing. The discussion that follows refers to the later version of the theoretical framework.

The Analysis/Control model explains language performance in terms of task demands and the abilities required to perform tasks. Different tasks make different demands on the learner's level of analyzed knowledge and different degrees of cognitive control. The model accounts for development as increments in analysis and in control (Bialystok, 1994a). Over time the learner's mental representations of language structures become increasingly explicit. In their initial unanalyzed state, language representations are organized around function and meaning; as the process of analysis takes place, they become organized around formal structures. This increasing analysis of mental representations leads to greater accessibility to knowledge and thus greater flexibility in its use. The main factors influencing this process are cognitive growth and the acquisition of literacy. Development in control is associated with the learner's increasing ability to execute intentions and direct performance in real time. When problems can be solved with less attention, they appear to be performed with fluency or automaticity. Bilingualism has been

shown to lead to enhanced control in children (Bialystok, 1987). This view of language proficiency is supported by considerable experimental research by Bialystok and others (e.g., Bialystok, 1986, 1988; Ricciardelli, 1993).

Analysis/Control and Metalinguistic Performance

Metalinguistic tasks typically make the greatest demands on both analysis and control (Bialystok, 1991). But even metalinguistic tasks vary along the two dimensions: correcting errors requires more analyzed knowledge than simply identifying them. Figure 2.2 illustrates how different metalinguistic tasks can be classified by the Analysis/Control framework. The most demanding metalinguistic tasks are those which require the individual to direct attention away from meaning and towards the formal features of language. Consider, for example, the study by Gleitman and Gleitman (1979) presented earlier. When confronted with the syntactically odd compound word 'eat house-bird', a clerical worker paraphrased it as 'everyone is eating up their pet birds' while a Ph.D. candidate responded with 'a house bird who is very eat'. The highly educated speaker was able to deal with the syntactic oddness of the stimuli, revealing a willingness to recategorize words and the ability to avoid being distracted by the non-meaningfulness of the utterance. The role of selective attention on such tasks is very clear.

The processes of analysis and of control proposed by Bialystok are general cognitive constructs. The framework explains both first and second language acquisition as involving the same processes, although the starting points are different for the child acquiring a L1 and the older L2 learner (Bialystok, 1994a). The L2 learner begins with the analyzed representations of language gained from the L1, and with developed procedures for directing attention. Given this framework, we would predict that individuals who can perform metalinguistic tasks in the L1 will be at an advantage in L2 learning when compared to learners who cannot perform such tasks.

We now turn to a similar model proposed by Cummins (1983, 1987).

73

High control

judge anomaly

symbol substitution

counting words in sentences

sun/moon problem

segment text

Low analysis   **High analysis**

produce rhyme, synonymy,
paraphrase

detecting errors   correcting sentences

judge correct sentences

Low control
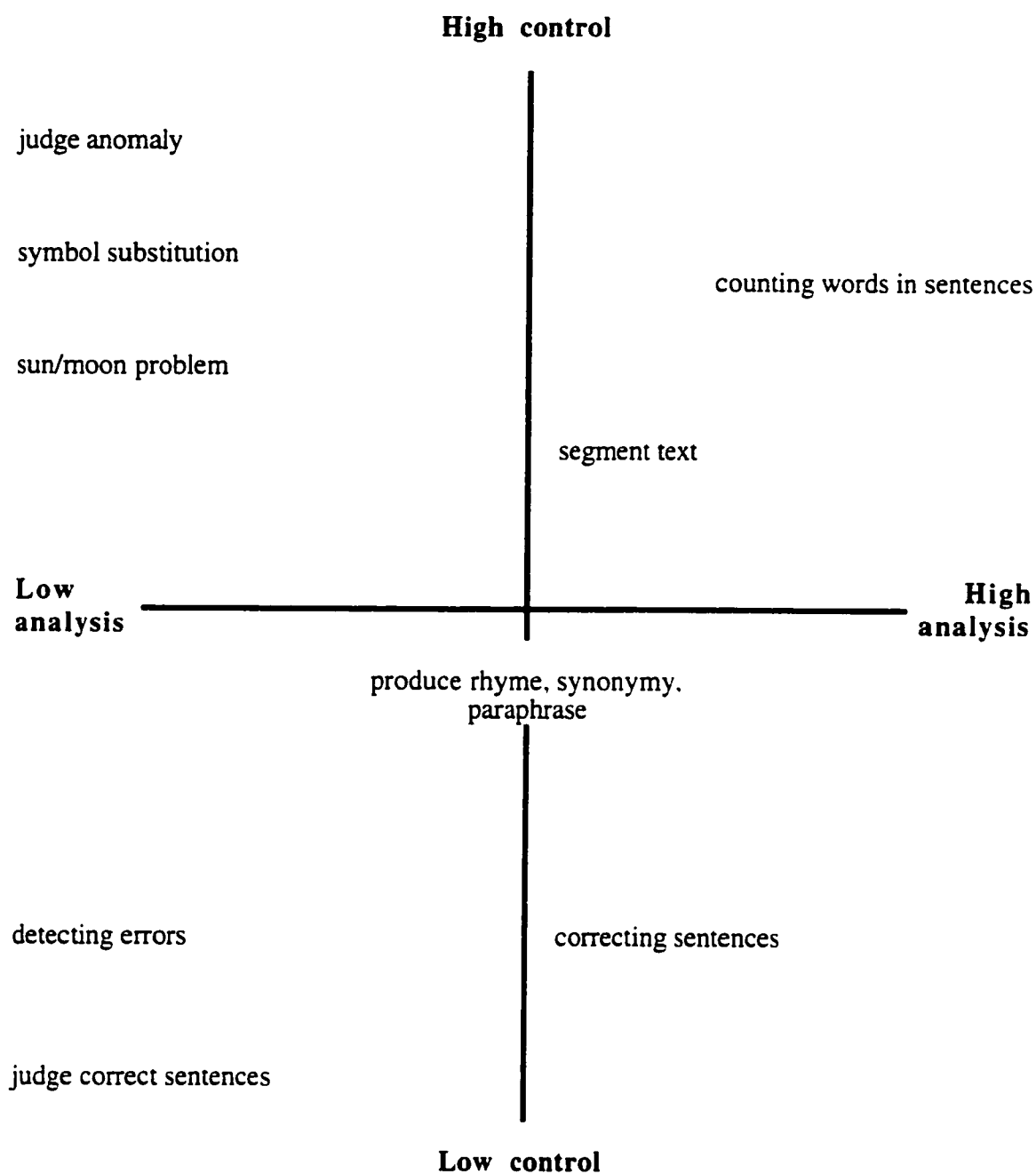
Figure 2.2 Metalinguistic uses of language according to Bialystok's Analysis/Control framework (Bialystok, 1991, p. 131)

Cummins' Model of Bilingual Proficiency

Cummins (1983, 1987) proposes that language proficiency can be conceptualized along two continua, one relating to the range of contextual support available for expressing or receiving meaning, and the other in terms of the degree of cognitive involvement required by the task. Context-embedded communication is characterized by feedback and negotiation of meaning, and a range of paralinguistic and extralinguistic cues. The production and comprehension of context-reduced communication, on the other hand, is primarily based on linguistic cues. Basically, the distinction captures the difference between tasks which permit top-down processing and those where bottom-up processing is required. Cognitive difficulty is conceptualized in terms of the amount of information that the individual must process simultaneously or in close succession in order to perform the task.

In earlier formulations, Cummins (1980) distinguished between 'basic interpersonal communication skills' (BICS) and 'cognitive-academic language proficiency' (CALP). Within the new framework, BICS covers tasks which are context-embedded and cognitively undemanding, while CALP characterizes tasks which are context-reduced and cognitively demanding. Academic language tasks tend to be context-reduced in nature, but their placement on the vertical dimension of cognitive involvement depends on the learner's level of cognitive development and amount of practice or exposure to similar tasks or situations. Metalinguistic tasks are categorized as being context-reduced but vary in the degree of cognitive difficulty depending on the age of the individual and the nature of the task (e.g., explaining language rules is more cognitively demanding than judging the grammaticality of a sentence). O'Malley and Chamot (1987) illustrate how typical classroom activities can be classified with respect to the two dimensions of Cummins' framework (see Figure 2.3).

Cummins (1987) argues that the abilities involved in context-reduced, cognitively demanding tasks (i.e., quadrant D in Figure 2.3) transfer across languages, given

sufficient input and motivation. There is considerable empirical support for such cross-linguistic transfer of skills. Countless evaluations of early French immersion programs have documented that students who learn to read and write in their L2 first do not need to be taught to read in English (see Genesee, 1987 for discussion). The literacy skills transfer from the L2 to the L1 and students catch up to their peers in regular English programs within a relatively short period of time. For minority-language children it has been shown that literacy skills developed in the L1 transfer to the L2 (see Cummins' 1991 for a review). We can summarize the main points from the preceding discussion on metalinguistic skill as follows:

- the development of metalinguistic ability is characterized by individual differences in both rate of development and in ultimate levels of skill;

- metalinguistic tasks and the abilities required to perform them can be analyzed in terms of their linguistic and their cognitive demands.

Both Bialystok's and Cummins' models make the prediction that metalinguistic skills will transfer across languages. It is, however, uncertain in what way metalinguistic ability in the L1 (or other languages) may be related to the development of more communicative skills in the L2. This issue will be examined in greater detail in Chapter 3.

Cognitively Undemanding

A

Developing survival vocabulary
Following demonstrated directions

Playing simple games

Participating in art, music, physical education

Engaging in face-to-face interactions
Practicing oral language exercises and communicative language functions

Answering lower level questions

B

Developing academic vocabulary
Understanding academic presentations accompanied by visuals
Participating in hands-on science activities
Making models, maps, charts, and graphs in social studies
Solving math computational problems

Solving math word problems assisted by manipulative and/or illustrations
Participating in academic discussions
Making brief oral presentations
Using higher level comprehension skills in listening to oral texts
Understanding written texts through discussions, illustrations, and visuals
Writing simple science and social studies reports with formal provided
Answering higher level questions

Context Embedded | **Context Reduced**

C

Engaging in predictable telephone conversations
Developing initial reading skills: decoding and literal comprehension
Reading and writing for personal purposes: notes, lists, recipes, etc.
Reading and writing for operational purposes: directions, forms, licenses, etc.

Writing answers to lower level questions

D

Understanding academic presentations without visuals or demonstrations
Making formal oral presentations
Using higher level reading comprehension skills: inferential and critical reading
Reading for information in content subjects
Writing compositions, essays, and research reports in content subjects
Solving math word problems without illustrations
Writing answers to higher level questions
Taking standardized tests in content subjects

Cognitively Demanding

Figure 2.3 Classification of language and content activities within Cummins' model of proficiency (from Chamot & O'Malley, 1987, p. 238)

77

Trait vs. Skill Approaches to Focus on Form Ability

In the three parts of Chapter 2, we have examined how the ability to focus on form has been treated under three different headings: as a component of aptitude, as learner style, and as metalinguistic ability. Each is a representative of a particular approach to studying individual differences. The first approach can be characterized as being nomothetic. This means that the learner attribute is viewed as a consistent trait in the population; every person is described as having some amount of this trait (Meier, 1994, p.36). A researcher working within this paradigm will typically administer a battery of aptitude and language measures to a large, heterogeneous sample and seek to account for differences among individuals (Loman, 1989).

The nomothetic perspective contrasts with the ideographic approach which considers individuals to possess unique characteristics which may be shared with all, some or no other individual (Meier, 1994). The typical ideographic research design is the case study of one or a few individuals over a long period of time. Many of the studies presented in the discussion of learner style (Part II) fall clearly in this tradition.

Finally, the third approach to studying language analysis ability differs from the trait models by viewing metalinguistic ability as an acquired set of cognitive competencies rather than as an innate capacity (Loman, 1989). Here the focus is on test item difficulty: the researcher contrasts learners at different stages of development (i.e., grade or proficiency level) on tasks that vary in their cognitive-linguistic demands and seeks to explain differences between the group means. Despite these real theoretical and methodological differences, however, there are a number of commonalities to be found across the different treatments of this particular learner attribute.

Common Threads

A case can be made that aptitude, learner style and metalinguistic skill are different manifestations of some common underlying capacity. The basis for this claim rests on the

78

pattern of correlations between aptitude, and some measures related to learner style, and those between aptitude and metalinguistic ability. As we saw in Part I of this chapter, Skehan's (1986b, 1989) follow-up study to the Bristol Project established a relationship between measures of L1 development and later L2 aptitude. The L1 measures included qualitative, linguistic indices which form the basis of what Peters (1981) referred to as the 'analytic' strategy for language acquisition. Measures of aptitude, on the other hand, appear to be related to metalinguistic skill. Recall that Carroll (1979) found that performance on Words in Sentences was highly correlated with scores on a test of knowledge of grammatical functions and of subjects' self-rating on knowledge about grammar. In a study conducted with college-aged francophones, Masny (1987; Masny & d'Anglejan, 1985) found that the French version of the MLAT, and in particular Words in Sentences, was a significant predictor of performance on an English L2 grammaticality judgment task. Furthermore, Alderson, Clapham, and Steel (1997) found, in a population of first year British university students studying French, that the MLAT correlated moderately with some (but not all) measures of metalinguistic knowledge in English and in French, and with L2 proficiency measures of a metalinguistic nature.

While these studies provide evidence of a link between performance on the MLAT and metalinguistic skills, the relationship is not straightforward. The grammatical sensitivity measure does not correlate highly with every metalinguistic task. In the Alderson et al. study, this might be due to the problem of restricted range since the subjects were all students who had chosen to study French at university. In other cases, the lack of relationship may be explained in terms of task demands. In Bialystok's terms, we should not expect correlations between tasks that make different linguistic and cognitive demands. For example, in Masny's study, Words in Sentences was a significant predictor of the ability to detect errors on the grammaticality judgment task, but was not a direct predictor of the ability to correct errors (the detection score was, however, a significant predictor for the correction measure). The correction task requires very explicit knowledge of English (L2)

morphology and syntax and thus a strong correlation with performance on the analysis of French (L1) sentence functions would not be predicted. This illustrates the need to adopt a finer grained approach to classifying metalinguistic tasks in order to clarify the relationship between aptitude and metalinguistic test performance.

There is a common thread that underlies this learner attribute which has been referred to variously as grammatical sensitivity, language analytic ability, or metalinguistic ability, and which we have discussed as the ability to focus on form. It is an ability that allows the learner to extract linguistic forms from their pragmatic context and manipulate them in some way. The person who has this skill is not only able to decontextualize language, but is able to work with language as form in addition to language as meaning. From this perspective, the role of selective attention in performing metalinguistic tasks is crucial.

In Chapter 3, the classroom and laboratory research which has investigated the influence of the ability to focus on form on L2 learning will be the subject of discussion.

NOTES

[1] This test is named after the word for 'word' in the artificial language.

[2] Clyne, while claiming to be following the work of Nicholas (1987), uses the term 'function' in the macro sense (i.e., communicative use) rather than in terms of grammatical function which is Nicholas' usage (Nicholas, personal communication).

[3] Note that Snow and Hoefnagel-Höhle do not themselves make the connection to Wong Fillmore's study.

[4] Olson (1994) points out that alphabetic writing arose in a straightforward fashion when a syllabic script developed for a Semitic language where vowel differences are not significant was adapted to Greek where vowels are highly significant. Semitic characters were used for vowels needed in words that were all vowels, for example, and this led them to hear vowels in the syllables used in the Semitic script and then dissolve them into CVCV pairings.

CHAPTER 3

REVIEW OF EMPIRICAL STUDIES AND RESEARCH QUESTIONS

In the previous chapter, three different theoretical approaches to describing a learner's ability to focus on language form were presented, and the similarities and differences were examined. The first part of Chapter 3 consists of a survey of the classroom and lab studies that have dealt with the relationship between the ability to focus on form, whether defined as aptitude, learner style, or metalinguistic skill, and L2 learning. These studies are organized according to five instructional contexts: form-oriented classrooms, content-based instruction, university L2 courses, communicative classrooms, and laboratory experiments. In the second part of the chapter, the salient features of these studies are discussed leading to the presentation of the research questions and hypotheses for the present study.

Review of Classroom and Lab Studies

As noted in Chapter 2, the development of Carroll's aptitude model was based on a very extensive research program followed by validation studies involving all types of learners. During the time period in which this was done, the predominant teaching approach was that of audiolingualism. The subsequent 'fall from grace' of the audiolingual method and the shift towards the communicative approach to L2 teaching in the late 1970's meant that aptitude research was tarred with the same brush. Consider the views expressed by Cook (1991, p.76)[1]:

> Such tests are not neutral about what happens in the classroom nor about the goals of language teaching. They assume that learning words by heart is an important part of L2 learning ability, that the spoken language is crucial, and that grammar consists of structural patterns. In short MLAT mostly predicts how well a student

will do in a course that is predominantly audiolingual in methodology rather than in a course taught by other methods.

In his review of the aptitude literature, Krashen (1981) argues that

a. aptitude is related to success in 'learning' contexts and on tests of 'learned' knowledge under conditions that favour monitoring;

b. aptitude is not correlated with success in 'acquisition' settings, or on tasks that make use of 'acquired' knowledge;

c. therefore, aptitude is not relevant to L2 acquisition.

Note that in his discussion, Krashen draws conclusions about the influence of aptitude on L2 'acquisition', although, in fact, his review deals almost exclusively with grammatical sensitivity. Many would agree with Krashen's conclusion based on common sense: analyzing grammar and learning to communicate effectively in conversation are two different things, so practicing one will not help you with the other. In general, aptitude research has fallen by the wayside, and Krashen's view about the (non)role of cognitive learner variables for the acquisition of L2 in communicative contexts has been left largely untested until quite recently. As we shall see, recent studies do provide some counter-evidence to Krashen's claims.

Whether one agrees with Krashen or not, it is clear that the interpretation of the research findings concerning the contribution of learners' language analysis abilities requires consideration of both the type of instructional environment in which learning takes place and the type of L2 performance that is used as the criterion. In the following sections, we will examine studies in which students' analytic abilities, either as a component of L2 aptitude, as field independence, or as metalinguistic skill has been assessed. The review is organized around the following types of instructional setting:

- form-oriented classrooms
- content-based instruction

- university L2 courses

- communicative classrooms

- laboratory experiments

Under each of the headings, the studies are presented in chronological order with a summary at the end of each section.

Research in Form-oriented Classrooms

The studies that are grouped under this heading are generally those which antedate the widespread adoption of a communicative approach to teaching. The validation testing of the MLAT in the 1950's and the PLAB in the 1960's (Carroll, 1962; Pimsleur, 1966) provides evidence of the relationship between L2 aptitude and success in L2 classes where the teaching has followed a structural syllabus. Carroll (1981) states that the MLAT was found to be a predictor of success in many unpublished studies conducted in different types of L2 classrooms.

Similar findings are found in Gardner and Lambert (1972), a volume which contains reports of the findings from a number of studies conducted in high school French classes in the United States (Maine, Louisiana and Connecticut), as well as reprints of earlier studies conducted in Montreal (e.g., Gardner, 1960). The research design of these studies involved administering a large array of measures including the MLAT, IQ tests, attitude-motivation questionnaires as well as tests of many different aspects of L2 achievement such as listening and reading comprehension, knowledge of vocabulary and grammar, oral production, and grades in French. These multivariate data sets were then subjected to a factor analysis designed to simplify the set of interrelationships. Across the various settings, the overall pattern of the results was the same. In almost all of the studies the first factor, the one that accounted for the greatest amount of the variance, was an Aptitude-IQ factor which was a strong predictor of achievement in reading, vocabulary, grammar, grades in French, and overall academic average. This factor consistently loaded

83

highly on the Words in Sentences subtest. The motivational factor emerged as an independent source of influence on L2 learning.

In a later study conducted in different communities in English Canada, Gardner, Smythe, Clément, and Gliksman (1976) found that aptitude consistently correlated with French grades. A total sample of approximately 1,000 learners in regular French as L2 classes in grades 7-11 were studied, and measures of attitudes, motivation, and anxiety were administered along with the aptitude test.[2] While the authors focussed on the role of motivation, aptitude was the predictor that correlated most consistently with the linguistic criterion measures. Not only did aptitude correlate with performance on listening comprehension and grammar tests, it also correlated (at around $r = .3$) at all grade levels with a rating of fluency and pronunciation accuracy that was derived from samples of free speech and oral reading. Overall, aptitude entered into stronger relationships with French at the more advanced levels.

As its name indicates, the Good Language Learner Study (Naiman, Fröhlich, Stern, & Todesco, 1978/1996) sought to determine which learner characteristics are associated with success in L2 learning. The result was a "groundbreaking, seminal study that for the first time yielded empirical findings on the styles and strategies of successful learners" (Brown, 1997, p. 511). The research project consisted of an interview study with adults and a classroom study with adolescents. The learners interviewed in the first stage of the project were highly educated, mainly anglophone adults who had been identified as being highly proficient in one or more L2s. The researchers concluded from this self-report data that aptitude, early first exposure and length of exposure were not as significant as such factors as motivation-attitudes, optimal learning conditions, personality traits such as sociability and persistence, and learning techniques. Yet, the quantification of the data shows that nearly all of these successful learners approached language learning in an analytic way; for example, they considered their language learning to be either "a highly

conscious and systematic process" (68%), or "one comprising both conscious and unconscious elements (26%)" (Naiman et al, 1996, p. 25).

The second phase of the project made use of insights gained from the interview study and attempted to isolate some of the critical variables associated with the good language learner. The classroom study involved administering French tests and measures of learner traits, as well as observing student classroom behaviour, and interviewing students concerning their learning strategies. The participating students were attending regular French L2 classes at the grade 8, 10 and 12 levels in six different schools in Ontario. The participants of the study ($\underline{n}$ = 36) were randomly selected from a pool of potential subjects identified on the basis of teacher ratings and their performance on a L2 listening test. Testing included measures of field independence, cognitive style, personality, motivation and attitudes. The researchers chose not to include aptitude and intelligence measures because of time constraints, and the desire to focus on the (then) neglected aspects of individual differences. As a criterion variable, the selected students were given an elicited imitation task in French.

The multiple regression analysis revealed that field independence was a predictor of performance on both the listening and imitation task (accounting for 18% and 11% of the apportioned variance, respectively) at the grade 12 level, but not at the grade 10 level, and to a limited extent at the grade 8 level (a total of 2% of the variance on the imitation task only). The authors conclude that attitudinal variables are more important at the early stages of L2 learning, but that field independence and other cognitive variables are more important at more advanced stages of learning, thus confirming the findings of Gardner et al. (1976). The significance of field independence was also revealed in the analysis of the elicited imitation data. For each grade, students were classified as field independent and field dependent according to their scores on the EFT. It was found that the two groups differed significantly in their ability to reproduce French: the field dependent learners were more

likely to omit a greater number of segments of the stimulus sentences than were the field independent learners.

The L2 measures used in these early studies would now be characterized as discrete-point rather than integrative and as form-focussed rather than meaning-focussed. Even the oral production tasks such as elicited imitation used in the Good Language Learner Study involve the ability to handle decontextualized language rather than requiring learners to demonstrate their capacity to communicate meaning. Not surprisingly, scores on the elicited imitation task correlated with performance on the decontextualized listening test. None of these studies included a test of the ability to use the L2 for communication. Thus, we are left with the question: do learners with greater language analysis ability who are learning in form-oriented classrooms perform better on communicative tasks even if such activities have not been a focus of classroom instruction? Later studies such as those of Bialystok and Fröhlich (1978) and Spada (1979) shed some light on the matter.

Bialystok and Fröhlich (1978) investigated the role of various individual factors (aptitude, cognitive style, attitude and strategy use) in L2 learning. The measures of the independent variables were the MLAT, an EFT, and a questionnaire measuring attitudes and strategy use. The L2 tasks contrasted in modality (oral vs written) and the degree to which they were code-oriented (i.e., form-focussed) or function-oriented (i.e., communicative). Four French achievement measures were developed: an aural grammar test (error identification and correction task), a written discrete-point grammar test, a listening test, and a reading test. These tests were administered to three grade 10 and three grade 12 regular French as a second language classes in Ontario. Aptitude consistently accounted for the largest portion of the variance on the L2 achievement tests. The multiple regression analysis revealed that Words in Sentences accounted for most of the variance on the aural grammar, written grammar, and reading tests, but not on the listening test. The latter was the measure that required the least explicit grammatical knowledge of French, since the students listened to dialogues and dramatic scenes and answered comprehension

86

questions; the emphasis was on contextual meaning rather than formal aspects of the L2. This would suggest that grammatical sensitivity is not as relevant for contextually-dependent receptive skills.

While Bialystok and Fröhlich attempted to contrast formal and communicative abilities, the measures they used did not involve spontaneous production of any kind. Spada (1979), however, included the analysis of oral production data in her study.

Spada (1979) compared the L2 performance of three groups of grade 7 students in Quebec who varied in their degree of contact with English. These French L1 learners had all had the same kind of audiolingual ESL classroom instruction. On the basis of their responses to a questionnaire, students were divided into low-, medium- and high-exposure groups (n = 11, 13 and 9 respectively). Three types of measures were administered: an aptitude test measuring inductive language learning ability (a French version of Green, 1974/5), a grammaticality judgment test in English, and an oral production task. By chance, it turned out that the medium-exposure group had higher levels of aptitude as measured by the inductive language learning test than the two other groups whose group means were almost identical. This makes it possible to compare the results of a group of high aptitude students to a group of lower aptitude students who have the advantage of a greater amount of exposure to meaningful input. As one would predict based on the aptitude research, the high-aptitude group on average had significantly higher scores than the low-aptitude, high-exposure group on the grammaticality judgment task. Spada also found that on many of the accuracy and fluency measures of the oral language elicited from the students, the high aptitude students were as good as or superior to the group who had much more contact with English outside of class.

One could argue that Spada's results suggest that learners with greater language analysis ability are better able to deploy knowledge gained from structural instruction on communication tasks even when compared to those having the advantage of more practice in using the L2 for communication. While other interpretations are possible (for example,

Krashen might argue that the medium-exposure group had a lower affective filter and thus made better use of the comprehensible input they received), this early, small-scale study suggests that the ability to analyze grammatical structure is involved in more than just performance on discrete-point grammar tests.

Spada's study finds support in the results from the Bristol Language Project follow-up study reported in Skehan (1986b, 1989, 1990) which investigated the relationship between L2 aptitude, L1 development and later L2 achievement (see Chapter 2, Part I). The students had been studying French or German for 1-2 years at the secondary school level in the UK[3]. Testing included measures of verbal intelligence, aptitude (grammatical sensitivity, inductive language learning ability and phonetic coding ability), and achievement in French or German as measured by standardized tests of speaking, listening, writing, and reading. Aptitude and all L2 achievement measures were found to be strongly related to each other. The measure of grammatical sensitivity used (from the EMLAT battery) correlated significantly and strongly with the tests of all four L2 skills (ranging from r = .66 to .69); similar results (r = .65 to .74) were found for the York inductive language learning test. The fact that these correlations are so much stronger than has been generally found in the literature is probably due to the reliability of the L2 tests and the heterogeneity of the sample (Skehan, 1990). It is important for the present discussion to note that the tests of grammatical sensitivity and of inductive language learning correlated strongly and positively with the tests of speaking ability.

Summary of Findings from Form-oriented Classroom Settings

The findings of the studies conducted in contexts where L2 teaching is form-oriented can be summarized as follows: learners who have higher levels of grammatical sensitivity are at an advantage on tests that require explicit, grammatical knowledge, or when L2 skills such as listening and reading comprehension are tested in a decontextualized

way. However, there is some evidence to suggest that learners with higher levels of analytic ability also have an advantage in making use of their knowledge gained from explicit instruction when performing oral communicative tasks.

In the next two sections, we turn our attention to the research that has been conducted in instructional settings that can be characterized as being meaning-oriented rather than form-oriented. We begin with an examination of studies conducted in immersion or bilingual education classrooms where the L2 is used as the medium of instruction.

Research In Content-Based Instructional Settings

A study by Tucker, Hamayan, and Genesee (1976) examined the role of affective, cognitive and social factors in L2 acquisition. They compared three groups of grade 7 students: an early immersion group ($n = 17$), a late one-year immersion group ($n = 28$), and a group of students who were in the regular English program and consequently getting French instruction for a few hours a week ($n = 19$). Tests administered included a standardized test of verbal and non-verbal intelligence, the MLAT, an EFT, an attitude-motivation instrument, questionnaires assessing personality and language background, and the following measures of French proficiency:

- multiple-choice achievement test of spelling, listening comprehension, vocabulary and grammar;
- a multiple-choice reading comprehension test;
- a listening comprehension test consisting of multiple-choice questions concerning a news broadcast from French radio;
- ratings of oral production in terms of comprehensibility, pronunciation, vocabulary, grammar, and amount of communication.

The predictor variables were first factor-analyzed, allowing the number of variables to be reduced to four factors ('IQ-Aptitude-Use', 'Personality', 'Attitudes-Motivation' and 'Cognitive Style'); the factor loadings were then used for multiple regression analyses. The results of the multiple regression showed that the 'Attitudes-Motivation' factor was

associated with performance on all four of the L2 measures; the 'Intelligence-Aptitude-Use' factor was associated only with the listening comprehension test, and the 'Cognitive Style' factor was a predictor on the achievement test.

It was discussed earlier that performance on the MLAT, particularly the grammatical sensitivity subtest overlaps to some degree with measures of intelligence. Genesee (1976) examined the relationship between intelligence and performance on academic vs interpersonal communication skills. Students from three different grade levels (4, 7 and 11) were tested. The primary school children (N = 124) had begun their schooling in an early total immersion program; the grade 7 subjects (N = 166) were near the end of a one-year late immersion program; and the grade 11 students (N = 155) had also been in a late immersion program in grade 7 and had taken post-immersion French courses during their secondary schooling. The students were classified as being below-average, average and above-average intelligence on the basis of IQ test scores. Comparison groups of students at each of the three grade levels were also selected from the regular English program who received regular French L2 instruction. Students' French language abilities were tested using multiple-choice tests of listening and reading comprehension, a standardized French language achievement test, and a math achievement test; interpersonal communication skills were assessed by an individual interview with a native speaker. These interviews were audio-recorded and rated along five dimensions (listening comprehension, pronunciation, grammar, vocabulary, and communicativeness) by native speakers. The results of the statistical analyses for each grade confirmed the widely documented finding that IQ is related to academic achievement: students with higher scores on the IQ test scored higher on the three tests of French academic skills; the above-average IQ students outscored the average IQ students, who outscored the below-average students. However, Genesee found no consistent association between IQ level and performance on the measures of interpersonal communication for any of the grade-levels.

Genesee and Hamayan (1980) studied individual differences among 52 anglophone children in early total French immersion in grade 1 in Montreal. Since they were dealing with 6 year old children, the usual battery of IQ and aptitude tests was not deemed appropriate. On the other hand, field independence with its focus on non-verbal analytic ability seemed like a potential predictor of individual differences in the immersion classroom context where the learner "must learn to differentiate and organize linguistic rules from a background of oral language" (p. 96). The predictor variables included a version of the EFT for children, a non-verbal intelligence test, an attitude questionnaire, personality, and teacher ratings of classroom behaviour. L2 learning was measured using tests of French achievement (spelling, synonyms/antonyms, oral comprehension of a text, picture/word matching, vocabulary, logical sequencing), and listening comprehension (non-verbal responses to items). Oral production skills were measured using native speaker ratings of elicited speech; each child was recorded while describing a cartoon and while conversing informally with a native speaker interviewer.

Using the same research methodology as Tucker et al. (1976), Genesee and Hamayan first reduced the predictor variables through factor analysis and then entered them into a multiple regression analysis with the L2 criterion variables. They found that the factor that included FI/D and non-verbal reasoning correlated with the French achievement test and the listening comprehension test, but not with oral production skills. The authors conclude that individual differences in FI/D do influence some aspects of L2 learning in a meaning-oriented classroom environment, even at a young age.

A study by Geva and Ryan (1993) provides evidence of how performance on a metalinguistic task in L1 is related to L2 learning in a different content-based instructional setting. The subjects in the study were 73 grade 5-7 anglophone children studying in a bilingual English-Hebrew school where the L2 was used for about 40% of the school day. Tests administered included measures of intelligence, the Clause Completion task, a teacher rating of oral proficiency in Hebrew and a forced-choice cloze test in Hebrew. The Clause

Completion task is a multiple-choice test which measures analyzed knowledge of English conjunctions. In the correlational analyses, the Clause Completion task correlated significantly with both the L2 cloze (r = .37) and the teacher rating in Hebrew (r = .41). The correlations remained significant even when intelligence and grade level were partialed out (r = .26 for both the L2 cloze and the oral rating). In contrast to the immersion studies, Geva and Ryan found that performance on an analytic task in L1 predicted performance on both paper-and-pencil tasks and assessments of oral proficiency in L2.

Finally, we turn to the study by Harley and Hart (1997) which was discussed in Chapter 2 related to the issue of whether aptitude is influenced by experience, and concerning aptitude-based learner profiles. Here we will limit discussion to the findings concerning the language analysis aspect of aptitude. The study involved a comparison of grade 11 students who had begun French immersion in grade 1 (n = 36) vs. those who began in grade 7 (n =29). Harley and Hart found that the two groups of learners differed considerably in the pattern of relationships between analytic ability and the L2 measures. The researchers administered the Language Analysis subtest of the PLAB and a variety of L2 paper-and-pencil measures: a yes/no vocabulary recognition test, a listening comprehension test, a cloze, and a written response to an open-ended question. For the vocabulary test, learners had to identify words that they knew from a list of real French words and plausible sounding pseudo-words. The listening test required students to listen to five tape recorded excerpts from oral texts in French and respond to written multiple-choice comprehension questions. The written production data were analyzed in two ways: the number of error-free T-units provided a measure of accuracy, and a rating of how well task requirements were fulfilled provided a measure of communicative effectiveness. In addition, oral production data were collected from a subsample of 18 students from the early start group and 20 from the late start group. The oral data included elicited imitation and cartoon descriptions; the analysis of the description data focussed on the range of lexical verbs produced.

For the learners who began their L2 exposure in late immersion, performance on the Language Analysis subtest correlated moderately and significantly (around $r = .4$) with the cloze test and the decontextualized yes/no vocabulary recognition test. In addition, the aptitude subtest correlated at a similar level with measures of both accuracy and of communicative effectiveness on the written task. In contrast, analytic ability did not correlate significantly with the listening comprehension test nor with the oral production measures. The correlations with elicited imitation and oral verb vocabulary were around .30 but were not statistically significant given the small size of the subsample that took part in the oral tasks.

The results for the high school students who had an early entry into French immersion revealed a very different pattern: the Language Analysis subtest correlated with the test of listening comprehension but with no other French proficiency measure. Unlike the late-entry group, there was no relationship between performance on the Language Analysis test and grammatical accuracy in written production. The findings suggest that older learners will tend to rely on different cognitive abilities than learners who begin intensive L2 instruction at an earlier age. Harley and Hart also admit that such differences may arise in response to differences in the instructional focus of early vs late immersion programs.

Summary of Findings from Content-Based L2 Settings

The picture of the role of analytic abilities that emerges from the French immersion research literature is as follows: measures of language analysis ability tend to be associated with performance on tasks that require a focus on form or that test L2 skills in a decontextualized way. In particular, several studies found correlations between the language analysis measure and performance on tests of listening comprehension in which learners responded to multiple-choice questions (Tucker, Hamayan & Genesee, 1976; Genesee & Hamayan, 1980; Harley & Hart, 1997 (i.e., for the early immersion learners)).

On the other hand, analytic ability did not correlate with performance on integrative measures used to assess oral production. It should be noted, however, that almost all of these studies used rating scales to assess oral production.

The next teaching/learning context we will examine is that of the university language class.

Research in University Settings

A number of studies have been conducted in the university setting examining the relationship between field independence and L2 learning. For example, Stansfield and Hansen (1983) administered an EFT to almost 300 beginning level students of Spanish. They found that performance on the EFT correlated significantly but modestly with students' final course grades, ratings of oral proficiency by the teacher, and scores on a cloze test. However, when academic ability was partialed out, the correlations were only significant for the cloze.

The results from Stansfield and Hansen may reflect the fact that they were dealing with beginning level students of a foreign language in the US. A different pattern of results was found by Chapelle and Roberts (1986) in their study of international students $(N = 61)$ enrolled in an intensive ESL program. A number of tests were administered at different times in the year. The predictor measures included an EFT and measures of several affective variables; the criterion variables were the TOEFL, a grammar test, a cloze, a dictation and a test of oral communicative competence. The oral test consisted of an individual interview in which the subject was engaged in role-plays which were rated for grammatical, pragmatic, and sociolinguistic competence by the researchers. The multiple regression analyses indicated that field independence was a significant predictor for all the L2 measures except the cloze.

Chapelle and Robert's finding was not replicated by Johnson and Rosano (1993) who studied international students ($\underline{N}$ = 45) in an intensive ESL program in Toronto. They measured the learners' performance on two decontextualized measures of proficiency (i.e., vocabulary and verbal analogies) and two measures of communicative ability (i.e., rating of students' pragmatic competence in L2 by teacher and a metaphor interpretation task). In this study, FI/D was measured using the Block Designs test from the Wechsler Adult Intelligence Scale-Revised in which the subject uses coloured blocks to reproduce pictured models; like the EFT, this task requires mental decomposition of the stimulus image into its component parts. Field independence correlated positively with the analogies test ($\underline{r}$ = .40), but negatively with the number of metaphor interpretations generated by the subjects ($\underline{r}$ = -.49) and with the teacher ratings of L2 communicative ability ($\underline{r}$ = -.57). This is the only study that has shown field dependence (i.e., a low score on the Block Design task) to be related to communication skills.

Summary of Findings from University Setting

The inconsistent relationships found between field independence and different aspects of L2 proficiency revealed in these three studies reflects the overall pattern in the FI/D research literature (see Chapelle & Green, 1992; R. Ellis, 1994 for a review). Even with respect to one particular measure of L2 performance, the cloze, the findings are contradictory. Certainly, it does not appear that the cognitive restructuring ability purportedly measured by the EFT or the Block Designs test is a straightforward predictor of L2 learning success.

In the next section, we will consider studies that were conducted in L2 teaching/learning settings in which instruction is described as consisting of a mixture of both form-oriented and communicative activities.

95

Research in Communicative Classrooms

Even some fifteen years after the publication of Krashen's 1981 article in which he contended that no studies showed a link between aptitude and 'acquisition', few empirical studies conducted in communicative language teaching contexts exist to support or refute this claim. Reves (1983) is often cited as relevant to this issue. She studied Arabic native speakers in Israel who were acquiring Hebrew "informally" in a bilingual community and studying English "formally" at school. The learners were grade 11 and 12 students who had been exposed to Hebrew since early childhood and had studied English since grade 5. An important point is that in grade 5 these learners had also started to receive instruction in Hebrew which focussed on the written language. It is thus not the case that the exposure to Hebrew as a L2 was purely informal.

The predictor variables were personality, L2 aptitude, cognitive style, learning strategies and language attitudes. The aptitude measures consisted of Arabic adaptations of Number Learning and Words in Sentences from the MLAT, and tests in Arabic developed by the researcher to measure word learning, word recognition, and imitation abilities. The six measures of L2 performance were ratings of oral fluency and of grammatical accuracy, and course grades for each language. The multiple regression analysis revealed that aptitude accounted for significant percentages of the variance on all criterion measures for both Hebrew and English. However, with respect to the components of aptitude, Reves found that the Imitation subtest accounted for the largest share of the explained variance while Pattern Recognition (the Arabic adaptation of Words in Sentences) only accounted for a very small portion of the variance on the ratings of oral accuracy in Hebrew and in English, and on the English final grade. While this study shows the relevance of aptitude for later stages of L2 learning, the results indicate that grammatical sensitivity is not as important as other aspects of aptitude with respect to oral fluency.

Two studies which have examined learning in communicative classrooms are Horwitz (1987) and Ehrman and Oxford (1995). Horwitz (1987) sought to distinguish the abilities that facilitate the development of grammatical competence from those that facilitate the development of communicative competence. She hypothesized that the social cognitive abilities associated with a learner factor called 'conceptual level' (Hunt, Butler, Noy, & Rosser, 1978) would be associated with communicative competence, while aptitude would be associated with grammatical competence. The developmental stages of conceptual level provide an index of a person's "cognitive complexity and interpersonal maturity": Stage 0 is characterized by self-centeredness; Stage 1 describes the individual who is concerned with social acceptability; in Stage 2 the individual is more open to new ideas and more tolerant of ambiguity and differences of opinion, but is still mainly concerned with being independent; and finally at Stage 3 individuals explore different perspectives when seeking solutions to a problem, taking into account their own ideas and feelings and those of others (Horwitz, 1987, p. 148).

American high school students in their second year of French $(N = 61)$ completed the MLAT and a written measure of conceptual level, as well as a French grammar test and three oral production tasks; these last were recorded and later rated for communicative effectiveness. A valuable aspect of the research design is that Horwitz selected the five classes that took part in the study on the basis of classroom observations, thus ensuring that the learners had comparable experiences of meaning-oriented activities in their French class; at least a third of observed class time involved communicative activities.

In accordance with the researcher's hypotheses, the MLAT correlated moderately and significantly with the grammar test $(r = .41$ for the MLAT as a whole and .49 for Words in Sentences), and conceptual level correlated with communicative competence $(r = .54)$ as measured by the oral tasks. In contrast to what was expected, there were also similar strength relationships (in the .4 range) between aptitude and communicative competence on the one hand, and between conceptual level and grammatical competence,

on the other. The two predictor variables also correlated with each other; it was thus necessary to determine the shared and unique variance accounted for by the two measures. For both the French grammar test and the ratings of oral production abilities, most of the apportioned variance was accounted for by the shared variance of conceptual level and grammatical sensitivity. For the test of communicative competence, conceptual level on its own accounted for a sizeable share of the variance. Horwitz suggests that the cognitive aspects of conceptual level such as inductive reasoning are responsible for the relationship with grammatical competence, while the interpersonal aspects contribute to the development of communicative competence. For the purposes of the present discussion, it is significant that there was a clear relationship between grammatical sensitivity and ratings of L2 oral language performance.

Ehrman and Oxford (1995) found a positive relationship between aptitude measures and L2 success in a more recent large-scale study. They examined relationships between a variety of learner variables (i.e., aptitude, learning strategies, learning styles, personality traits, motivation and anxiety) and proficiency ratings in speaking and reading a foreign language. The subjects were US government employees who had participated in intensive language instruction that was largely communicative but also included some features of audiolingual teaching such as drills and dialogues. A total of 855 individuals participated in the study, although data were not collected from everyone on all measures; for example, only 282 people completed the MLAT. Ehrman and Oxford found that the MLAT and an 'observed aptitude' rating by the instructor were the predictor variables that were most strongly correlated with the Foreign Service Institute ratings of speaking and reading. These proficiency measures correlated with the grammatical sensitivity subtest at around .4, and with the MLAT as a whole at .50, which is in keeping with the ranges found by Carroll during the audiolingual era. The authors conclude that aptitude is still a good predictor of L2 learning success even when teaching is communicative in orientation.

Summary of the Findings from Communicative Classrooms

The studies by Horwitz and by Ehrman and Oxford are suggestive that L2 aptitude, particularly grammatical sensitivity, is a good predictor of success even when the instruction follows a more communicative approach. These were, however, classes in which the teaching included both form-focussed and meaning-focussed activities and the actual percentage of each is not specified. Thus, we do not have evidence concerning the role of aptitude in what Krashen would consider an 'acquisition' classroom environment.

We now turn to studies conducting in the laboratory where the type of instruction is carefully controlled and documented.

Research in the Laboratory

Two lab studies that examined the interaction of grammatical sensitivity and type of instruction (implicit vs explicit) are reported in de Graaff (1995) and Robinson (1997). While some researchers have used the MLAT to ensure that the treatment groups in their lab experiments were equivalent with respect to aptitude characteristics (e.g., N. Ellis, 1993 and DeKeyser, 1995), de Graaff and Robinson used aptitude as an independent variable in the research design.

De Graaff (1997) studied the learning of a modified version of Esperanto called 'eXperanto'. The participants (N = 54) were randomly assigned to either an implicit or an explicit learning condition. Aptitude was measured as a pretest variable using Dutch versions of the MLAT Words in Sentences and Paired Associates subtests, and a lexical inferencing test. All three aptitude measures were collapsed into one variable for the correlations with the proficiency measures. The learning materials which were presented on a computer consisted of 10 lessons of approximately 1.5 hours each. The rules were embedded in dialogues which were contextualized in a setting. Learning activities included translation and production exercises. In addition, the explicit group received grammatical

explanations of the target structures. Proficiency measures were administered three times, once during the instructional phase, and twice after the treatment. These measures included sentence judgments, a gap-filling task, a translation task, and a sentence correction task.

De Graaff found that participants in the explicit instruction condition outperformed those in the implicit learning condition, and that aptitude correlated with performance on the proficiency measures. There was no interaction between the influence of aptitude and treatment condition: aptitude affected test performance in both conditions to the same extent. In other words, his findings support the idea that aptitude is relevant in both implicit and explicit instructional settings. We do not know, however, what the role of grammatical sensitivity alone is since de Graaff used a composite aptitude score.

In contrast to de Graaff's use of an artificial language, Robinson (1997) used English in a learning experiment conducted at the University of Hawai'i. Two syntactic rules were chosen for the study: one was judged by 15 ESL teachers to be an "easy" rule based on its structural complexity and the complexity of the pedagogical explanation it would require; the other rule was judged to be "hard" by the same criteria. The participants in the study were 104 learners of English as a L2, most of whom had Japanese as their L1. Only those who showed by their performance on a pre-test that they were unfamiliar with the target structures were selected for the study. Aptitude was measured using the English version of Words in Sentences and Paired Associates subtests of the MLAT.[4]

Subjects were randomly assigned to one of four training groups: implicit, incidental, rule-search, and instructed conditions. All subjects were presented with sentences on a computer for the same length of time. What they did after viewing the sentences varied according to the condition. In the implicit condition, participants were told that they were to perform a memory task, and were asked questions concerning the location of words in sentences. The incidental group answered yes/no comprehension questions and received feedback concerning the content of their responses. Subjects in the rule-search condition were told to try and find the rules in the sentences, and then were asked after each

sentence whether they had found the rules or were still looking; no feedback was provided. Finally, in the instructed group, participants began by viewing explanations of the easy and hard rules and were asked to apply the rule to each sentence as they saw it. Following each sentence, there was a metalinguistic question to which they responded with a 'yes' or 'no' and were given feedback. The target rules were categorized as either hard or easy based on the ratings of experienced ESL teachers. Learning was evaluated by a grammaticality judgment task administered after the training. In addition, subjects filled out a questionnaire in their L1 aimed at assessing their level of awareness of the target structures.

On the post-test, the instructed group outperformed the other groups with respect to both the easy and hard rules. With respect to the role of aptitude, the results on the grammaticality judgment task revealed that grammatical sensitivity was significantly correlated with learning in the implicit, rule-search, and instructed conditions, but not in the incidental condition. Indeed the strongest correlations were with accuracy on the easy and hard rules in the implicit condition (r = .69 and .75, respectively). In this condition, those who responded that they had looked for rules, and those who were able to verbalize rules were more accurate on the judgment task. Furthermore, those who were able to look for rules or could verbalize the rules were significantly better on Words in Sentences.

On the surface, it looks as if learners who had grammatical sensitivity were able to derive explicit knowledge out of a stimulus which was not structured to promote learning. In fact, the task demands in the implicit condition, which involved responding to questions such as "Were the words X and Y next to each other?" required careful processing of the form of the stimulus sentences. Robinson concludes that task demands determine the role individual differences in grammatical sensitivity and memory played in each learning condition. When an individual's abilities matched the demands of the task "aptitude often led to awareness, which often was associated with superior levels of awareness" (Robinson, 1997, p. 81).

The findings for the incidental condition are intriguing. This was the experimental treatment that most closely resembles normal language use where the focus is on comprehension. The performance of the participants in this condition on the judgment task was the poorest of all four groups; there was no correlation with either grammatical sensitivity or memory ability. This suggests that when time for learning is very limited, as was the case here (about 25 minutes total), and the input is not enhanced in some way, even learners with grammatical sensitivity are not able to notice and abstract the rules of the target language.

Summary of the Findings from Lab Experiments

Both de Graaff and Robinson found that aptitude correlated with learning outcomes under controlled laboratory conditions. In both studies, learning was assessed using form-focussed measures. The studies thus leave unanswered the question of how different types of instruction affect the development of functional skill in the L2. Robinson's finding that grammatical sensitivity was not correlated with learning in the incidental condition is of particular interest here. It is consistent with the proposal by VanPatten (1990) that beginning level students are not able to focus on both form and meaning at the same time. However, other studies (e.g., Horwitz, 1987) have shown a relationship between grammatical sensitivity and performance on both form-focussed and communicative measures of L2 proficiency. Clearly, the characteristics of the instruction and of the test tasks must be taken into account in order to make sense of these different results.

As always with lab studies, we are uncertain about the generalizability of the findings unless they are twinned in some way with findings from ecologically valid research. The last part of this chapter presents a synthesis of the overall pattern of research findings.

Research Questions and Hypotheses

Statement of the Problem

This review of classroom and laboratory research relating to the role of aptitude and other related learner variables on L2 learning should make it very clear that Cook (1991) and Krashen (1981) have overstated their case. Many studies conducted in a variety of teaching/learning contexts have revealed that learners' analytic abilities, whether measured as grammatical sensitivity, field independence, or metalinguistic awareness, are associated with success in L2 learning. The nature of the relationship between the predictor variables and learning outcome variables does vary, however, depending on the particular criterion measure. The available research findings suggest the following:

Generalization 1:

Learners who have higher levels of language analysis abilities have an advantage on tests of metalinguistic knowledge and decontextualized tests of language skills.

Generalization 2:

Learners who have higher levels of language analysis abilities are not advantaged on measures of oral communication.

The first generalization is fairly consistent in all the studies reviewed. Analytic learners surpass their non-analytic classmates when tested on their explicit knowledge of L2 grammar, or when L2 proficiency is evaluated using tasks that require decontextualized language skills. For example, measures of listening comprehension call upon decontextualized language skills when a test consists of many short listening texts that have no connection to each other, or to any larger meaningful purpose or context, and learners have to respond to multiple-choice questions. Krashen (1981) has argued that this relationship exists because of the similarity between aptitude tests and the form-focussed tasks prevalent in the teaching contexts (e.g., grammar-translation or audiolingual classes) where the research was carried out. In his view, the type of L2 instruction experienced by

the learners is responsible. By implication, a communicative classroom would not lead to such a relationship. However, no study has investigated this issue in L2 classes where the primary focus of instruction is on the development of interpersonal communication skills. If the relationship still holds in such an environment then the explanation must focus on the influence of learner traits on the acquisition process and on test performance.

Before moving on to the second generalization, we should consider the fact that even on decontextualized measures of L2 proficiency, only a portion of the variance is accounted for by the measure of analytic ability. Other factors necessarily contribute to L2 performance. As was discussed in Chapter 2, Part II, it has been suggested by Skehan (1989; 1998) that learners who have strong memory abilities can excel in L2 learning despite lower analytic abilities. An illustration of this is offered by CJ, a talented language learner studied by Obler (1989). CJ scored only at the 50th percentile on Words and Sentences but scored exceptionally high on many tests of verbal memory. A memory oriented learner is likely to be highly successful in settings where communication is emphasized. However, Skehan (1998) has recently hypothesized that analysis-oriented learners develop rule-based knowledge representations while memory-oriented learners develop exemplar-based representations. It follows from this that learners like CJ who have a high memory, low analysis profile (see Table 2.2) will be at a disadvantage when performing tasks which require more explicit knowledge of the L2. As yet, little empirical research has addressed this question.

Table 3.1

The Relationship Between Cognitive Variables And Oral Production

	Research context	Analytic ability measure	Oral production measure	Findings
Tucker et al. (1976)	French immersion; grade 7	MLAT; GEFT; and IQ	ratings of quantitity, comprehen-sibility, pronunciation, vocabulary, and grammar	no relationship
Genesee (1976)	French immersion; gr. 4, 7, 11	IQ	ratings of grammar, vocabulary, pronunciation, and communi-cativeness	no relationship
Genesee & Hamayan (1980)	French immersion; grade 1	GEFT	global rating	no relationship
Reves (1983)	Arabic medium schools in Israel; grade 11, 12	Arabic version of MLAT	ratings of accuracy and fluency	significant predictor
Horwitz (1987)	FSL classes high school	MLAT	rating of communicative effectiveness	moderate relationship
Ehrman & Oxford (1995)	US Foreign Service Institute	MLAT	ratings of speaking and reading proficiency	moderate relationship
Harley & Hart (1997)	French immersion	PLAB	accuracy of repetition and range of lexical verbs	no relationship

105

With respect to Generalization 2 above, we find that the relationship between aptitude and performance in oral communication is unclear. Table 3.1 offers an overview of the studies reviewed in this chapter which have included measures of oral production. In most of these studies, oral production was assessed using rating scales. Procedures vary across studies: speech samples were rated by native speakers in the Montreal studies (Genesee, 1976; Genesee & Hamayan, 1980; Tucker et al, 1976), by trained Foreign Service Institute (FSI) raters in Ehrman and Oxford (1995), and by the principal researcher in Horwitz (1987). Except for the FSI scale used by Ehrman and Oxford, the studies used undefined scales. This suggests the possibility that the contradictory results may be due to the assessment instrument. Small but significant differences in language ability among learners cannot be captured on a scale that covers a wide range of ability (e.g., absolute beginner to native speaker). Furthermore, problems with reliability and validity plague even some of the widely used rating scales (see Upshur & Turner, 1995 for discussion). Further research into the relationship between aptitude and oral proficiency would benefit from better rating scale construction; for example, Upshur & Turner, 1995 suggest using "empirically derived, binary-choice, boundary definition scales" as one way of addressing reliability and validity problems. Another approach is to assess oral production by examining the aspects of the learner's interlanguage (i.e., their developing L2 grammatical system). The studies by Spada (1979) and Reves (1983) suggest that aptitude may be related to differences in accuracy in oral production.

The Present Study

The present study addresses some of the issues raised in the preceding discussion by investigating the importance of analytic abilities for learners in a communicative language learning environment. The participants in this study are attending an intensive ESL program offered at the grade 6 level. In their ESL classrooms, English is used for all

'transactions' between teachers and students. Unlike immersion and bilingual education programs, the emphasis is almost exclusively on the development of oral rather than written skills through participation in a variety of communicative activities. The school in which the study was conducted is situated in an area outside of Montreal where, for the vast majority of the students, the classroom is the only place where they have the opportunity to interact in English.

Since learners are in an intensive language program, one would expect aptitude to play a role in determining levels of success at the end of the program (Carroll, 1981). But as we have seen, no research has taken place in a learning environment that promotes interpersonal rather than academic skills. In addition, most of the studies have studied high school, university or adult learner populations, thus leaving the question open as to whether aptitude variables are predictive of performance by pre-adolescents.

The focus of this study will be on 'grammatical sensitivity' since it is the label that most transparently refers to the learner's predisposition towards focussing on form. This term is used in Carroll's model of aptitude. He defined it as "the ability to recognize the grammatical functions of words (or other linguistic entities) in sentence structure" (Carroll, 1981, p. 105), a definition which is closely linked to the requirements of Words in Sentences. Other researchers have defined and operationalized this concept differently. For example, Willows and Ryan (1986) define grammatical sensitivity as performance reflecting the control and analysis components underlying metalinguistic ability while Emmorey, Bellugi, Friederici, and Horn (1995) define it as the ability to detect ungrammaticality. In this study, grammatical sensitivity will be defined as "the ability to focus on grammatical form".

Research Questions

The research questions and the research hypotheses for each question, posed in relation to grade 6 students in an intensive ESL program, are as follows:

Research Question #1:

Is grammatical sensitivity associated with performance on measures of L2 proficiency when learning takes place in a communication-oriented instructional environment?

Hypothesis #1:

There is an association between grammatical sensitivity and performance on L2 proficiency measures that are form-focussed in nature.

Research Question #2:

What is the relationship between degree of grammatical sensitivity and the profiles of successful vs. less successful L2 learners?

Hypothesis #2:

Higher levels of grammatical sensitivity distinguish successful learners from less successful learners when success is defined in terms of performance on form-focussed measures of L2 proficiency.

Research Question #3:

Is degree of grammatical sensitivity associated with stages of interlanguage development in oral production?

Hypothesis #3:

There is a positive relationship between grammatical sensitivity and stages of interlanguage development in oral production such that higher levels of grammatical sensitivity are associated with higher stages of development.

Chapter 4 presents the procedures used in the study that was designed to address these research questions.

NOTES

[1] This statement appears unaltered in the second edition of the book (Cook, 1996. p. 101).

[2] Gardner, Smythe, Clément and Gliksman (1976) do not provide information about which measure of aptitude was used, nor do they give a breakdown of the subtest scores.

[3] Skehan (1986b, 1989a, 1989b) does not give any information about the nature of foreign language teaching that the learners had experienced. Thus, I have assumed that they received form-focussed teaching, based on the time period in which these students were in secondary school.

[4] It is important to note that the English versions were administered because no appropriate measures existed in the subjects' L1. Robinson cites a study by Sawyer that demonstrates the reliability of an English version of the MLAT for Japanese students. However, the real problem is that differences in performance on the English MLAT may reflect proficiency differences in addition to differences in grammatical sensitivity, thus making the groups non-equivalent.

CHAPTER 4

METHOD

This chapter describes the method of data collection and data analysis used to investigate the research questions described in Chapter 3. The first sections describe the research setting and the characteristics of the learners and the teaching-learning environment. This is followed by a description of the instruments used in the study. The final sections present the three data analysis procedures used to analyze the data collected.

Research Setting

The participants in this study of individual differences were all learning English in an intensive ESL program in Quebec during the 1994-95 school year. Intensive ESL is a form of instruction that has become increasingly popular in French language schools in Quebec over the last ten years. According to Watts and Snow (1993), almost 22,000 students had participated in intensive ESL programs offered by 31 school boards over the approximately 15 years covered by their survey. The intensive model, which is considered to be "experimental", is not an official program of the Ministry of Education of Quebec but rather an expansion of the regular ESL curriculum which is designed for a recommended 120 minutes a week in each of grades four, five and six (Gouvernement du Québec, 1981). Schools vary in the amount and distribution of time allotted to intensive ESL, as well as the grade level at which it is offered (see Watts & Snow, 1993 for details). Students in the present study received five months of ESL instruction in grade 6.

The Intensive ESL School

While there are a large number of school boards in Quebec offering intensive ESL instruction, there is only one school in Quebec in which all students are involved in an intensive learning experience. In this school, five classes begin the year with an all-day program in ESL while another five classes study the grade 6 French language curriculum. At the end of January, in the middle of the school year, the classes switch. The school provides an English-language environment in that all members of the staff are bilingual to various degrees, and the use of English outside of the ESL classroom is encouraged through a number of different ways (e.g., English tables in the cafeteria, announcements given in English over the public address system). The salient characteristics of this school are summarized as follows.

- French is the L1 for almost all students
- students have little contact with English outside of school
- prior to the year of intensive study, all students have had a maximum of two years of regular ESL for 2 hours/week
- every student in the school is in grade 6 (10 classes of 30 students each)
- every student in the school spends one half of the school year in intensive ESL (= 4.5 hours/day)
- English is never used as the medium of instruction for subject matter teaching
- the ESL program emphasizes oral communication skills over reading and writing skills
- students are encouraged to use English outside the ESL class
- students from all levels of academic ability are eligible for the school's special program

There are many characteristics of this school that make it an ideal place for classroom research. Students are fairly homogenous with respect to out-of-class exposure to English; and the teachers work closely together in planning instructional activities. The data for this study were collected in conjunction with a larger study of L2 learning in this intensive ESL school during the 1994-95 academic year (Lightbown & Spada, 1997).

111

A Communicative Learning Environment

The purpose of this study is to examine L2 learning in a setting in which the focus of instruction is on the development of interpersonal communication skills rather than on academic language ability. It is thus crucial to ensure that learners were indeed in such an environment. All of the ESL teachers involved in this study were known to the author through earlier research activities. For three of the five ESL teachers, this included both formal and informal observations in intensive ESL classes over a period of ten years (see Spada, Ranta, & Lightbown, 1996 for discussion of this collaboration). However, these observations took place before these teachers began teaching at the Intensive ESL School. Thus additional information concerning the nature of L2 teaching in the school was gathered through a teacher questionnaire and structured observations.

As mentioned earlier, the content of most intensive ESL programs is based on the regular primary level MEQ program which is organized around a set of functions and notions, and emphasizes the development of oral rather than written skills, and fluency rather than accuracy (Gouvernement du Québec, 1981). The ESL program at the Intensive School is consistent with these principles. Learners participate in a variety of oral activities such as games, puzzles, surveys, interviews and discussions which are organized around themes relevant to the interests of students of this age (White, 1996).

In an earlier phase of the research, a questionnaire concerning teaching practices was sent out to all the school boards offering intensive ESL programs. Four of the five ESL teachers at the Intensive ESL School responded to the questionnaire. Their responses reflect a commitment to communicative language teaching. For instance, when their opinion about grammar was elicited, they commented:

> " Learn to speak first, as you learned your mother tongue, then proceed with learning grammar to improve the quality of the language"

> "...Teaching grammar or concentrating on it to learn a language is a false concept"

> "I believe students should learn to communicate before learning grammar structures"

"... They can communicate with kids their age and understand the messages without the grammar"

Finally, all teachers in the school were observed and the observations coded using Part A of the Communicative Orientation of Language Teaching (COLT) Observation Scheme (Spada & Fröhlich, 1996) for one full day in May-June 1995. This instrument permits the systematic description in real-time of features characteristic of communicative teaching such as the extent to which classes are teacher- or learner-centred, the amount of instructional time devoted to focus on form or focus on meaning, the amount of time spent on different skills, and the type of interactional organization. These observational data provide further confirmation of the teachers' commitment to a communicative approach to L2 teaching. The salient features of ESL instruction among the five classes at the Intensive ESL School were:

- English was consistently used for all functions in the classroom (i.e., pedagogical, disciplinary, transactional, etc.), and frequently used in the school environment, outside of the classroom;

- pair- and group-work activities predominated over teacher-centered activities;

- student-teacher interaction was characterized by negotiation of meaning;

- grammar was never presented in isolation;

- when focus on form occurred, it usually took place when homework was corrected.

On the basis of this information, we may conclude that learners in the school were indeed experiencing a communication-oriented approach to L2 teaching.

Learner Population

The participants in this study are the cohort of grade 6 students (N=150) who began their ESL instruction in September 1994 and who were in the French part of their intensive year from February to June 1995. Using this cohort makes it possible to examine

113

maintenance and further growth (or loss) in English during the period when the instructional focus is on subject matter taught in French. Students at the Intensive School represent a range of academic abilities. The only requirement for admission to this special program is that they not be fluent in English to begin with, and that they have the motivation and parental support to carry them through the demands of the intensive nature of instruction -- students must be prepared to cover the usual ten-month French language arts, math, and social studies content in a five-month period.

This population of learners is relatively homogenous: the vast majority of students in the school are francophones from unilingual families, their exposure to the L2 prior to participation in the intensive program has been limited to the classroom and exposure to English-language media, and the type of L2 instruction they have experienced has been meaning-oriented rather than form-oriented. Although the learners are older children (aged 11-13 years), they are cognitively mature enough to perform tasks that demand metalinguistic skills, such as those typically used in measuring grammatical sensitivity.

Instruments

Measuring Grammatical Sensitivity

A measure of grammatical sensitivity was needed in order to answer the research questions of this study. The vast majority of studies reviewed in chapter 3 made use of either the MLAT or the PLAB. There is not, however, a readily available version of either for use with French-speaking pre-adolescents. Carroll developed a form of the MLAT for younger learners (the EMLAT) but it is not available in French. A number of tests suitable for younger learners test inductive language learning ability rather than grammatical sensitivity. This is true of the Language Analysis subtest of the PLAB and the Aptitude test developed by Green (1974/5). The latter was translated into French as part of a study of L2 learning in audiolingual classrooms in Quebec (Spada, 1979). Green's test is basically a

lesson in Swedish grammar: various grammatical forms (e.g., articles, present tense, plurals) are presented and the examinee has to write new forms by analogy, construct sentences in Swedish, and translate them into English. This test was rejected on the basis of ecological validity. In this particular learning context, the focus is on authentic communication in English, both inside and outside the classroom; even translation between French and English is frowned upon. It was felt that it would be difficult to justify to the teachers and students the use of a test in which the learners would be required to figure out the grammatical rules for Swedish. Thus, new measures of grammatical sensitivity had to be developed.

Development of the Grammatical Sensitivity Measure

In the piloting phase, two different tasks were developed. The first task, called 'Trouve l'imposteur' ('Find the Imposter') was based on a variant of Words in Sentences used in a computerized aptitude test developed by Meara (1994). In Meara's test, subjects are presented with pairs of sentences that form part of a coherent text recounting the story of Cinderella. As with the MLAT, a word in the key sentence is underlined, and the test-taker has to choose an underlined word in the second sentence that fulfills the same grammatical function. It was felt that using a coherent text of a well-known story would decrease artificiality of the task. However, Meara's test was designed for adults and needed to be modified for the younger learners in this study. In the version that was eventually used in the test piloting, there were 15 pairs of sentences; in the key sentence one word was circled, and in the second sentence three words were underlined. The test taker's task was to identify the part of speech (i.e., "la catégorie") of the key word, and then find which of the three underlined words was not of the same part of speech.

The second grammatical sensitivity task involved detecting and correcting errors in written French. It is a characteristic of French that to spell it correctly one must perform a grammatical analysis of the sentence. Many words may sound the same but are spelled

differently, and the appropriate form can be based on the form class (e.g., the infinitive of some verbs is spelled with '-er' while a past participle is spelled with an '-é'), or require an explicit parse of the sentence (e.g., a past participle in a passé composé verb phrase with the auxiliary 'avoir' does not agree in gender and number with its subject but is marked for agreement with a pronominalized direct object that precedes it). Francophone students in grade 6 do not generally have complete control over the rules of written French since this requires years of practice and feedback. Thus, the rationale for using an error detection task in French as a measure of grammatical sensitivity was the assumption that the ability to detect grammatical spelling errors will reflect the learner's sensitivity to grammatical structure. The use of sensitivity to error has been used as a measure of grammatical sensitivity in other research (Emmorey et al., 1995).

The error detection task called 'Trouve les erreurs' was adapted from one which was developed by Lyster (1987) for use in a French immersion class. It consisted of 20 isolated sentences, including intralingual and interlingual (English-based) errors. Subjects were asked to indicate where the error was and to correct it.

The two tests were piloted on two classes of grade 6 students in another intensive ESL program in Quebec. This ESL population differs from that of the Intensive School in that students receive their English instruction all year round for half of the school day. However, they are in grade 6, and live in an area in which contact with English outside of school is relatively limited.

Results from these 60 students indicated that the Trouve l'Imposteur task was extremely easy for all students -- almost all students got 12 or 13 out of 15. Thus, all students had sufficient metalinguistic knowledge to identify words sharing the same form-class. On the other hand, the error detection task Trouve les Erreurs provided a greater range of scores. It was found that items that involved English-based errors discriminated the least, because virtually all students saw that they were incorrect. Changes were made to increase the number of errors which involved 'grammatical spelling', that is, where the

116

correct form is based on a grammatical analysis and not on the sound of the word. For example, the verb form that sounds like [parl] (i.e., 'speak') can be spelled as 'parle', 'parles', or 'parlent' depending on the grammatical distinctions of person and number. In its final form, 'Trouve les Erreurs' included twenty French sentences of which three had no error, and the remaining sentences consisted of the following: eight sentences with grammatical errors that did not involve homophony, six grammatical errors which were homophonous with the correct form, two non-standard grammatical forms, and one sentence in which a verb phrase was missing a tense marker and had an incorrect, homophonous grammatical ending. The complete test is presented in Appendix C.

Learning Preferences Questionnaire

In addition to the direct measure of learners' ability to perform metalinguistic analysis, a self-report questionnaire concerning learner preferences was included in the test battery. It was predicted that analysis-oriented learners would report a greater preference for form-focussed activities based on the common sense notion that 'you prefer to do the things that you are good at'. The 'Ma façon d'apprendre l'anglais' ('My way of learning English') questionnaire was based on one created by Willing (1989). The new measure was in French in order to assess student orientation independently of their knowledge of English. Students were presented with the description of 15 different types of classroom activities accompanied by an illustration. Some of these activity types, such as learning tongue-twisters or playing games, would be quite familiar to these students. Other activities, such as consulting grammar books, would not have been a part of their ESL classroom experience (although, they would have used them in their French academic courses). For the analysis, these activities were classified as being either more communication-oriented or more form-oriented. Activities considered to be form-oriented were those which required the learner to attend to linguistic form, either exclusively as in

117

the case of tongue-twisters, or in addition to a focus on meaning as in the case of writing texts. These activities are listed in Table 4.1.

Table 4.1

Types of Activities Used in the Learning Preferences Questionnaire

Communication-oriented activities	Form-oriented activities
• play games	• do grammar exercises
• listen to cassettes	• get corrected by the teacher
• do oral activities	• recite tongue-twisters
• watch videos	• consult a grammar book
• perform skits	• study vocabulary lists
• watch television	• look words up in the dictionary
• listen to the teacher's explanations	• write texts
	• read stories

Below the illustration of each activity is a four-point Likert scale and the student is asked to circle the phrase that corresponds to how well that activity helps them to learn English, either 'beaucoup', 'un peu', 'très peu', or 'pas du tout' ('a lot', 'a bit', 'very little' or 'not at all', respectively). Student responses for each item were scored by assigning a numerical value for each point on the scale (i.e., 4 = 'beaucoup', 3 = 'un peu', 2 = 'très peu', and 1 = 'pas du tout'). The Learning Preferences Questionnaire is provided in Appendix D.

In the next section, the tests administered to measure English development over the course of the school year are described.

Measures of English Proficiency

L2 proficiency was evaluated using paper-and-pencil and oral measures that had all been used in earlier research with students in intensive ESL programs, either in the Intensive ESL school or elsewhere in the province of Quebec (Lightbown & Spada, 1997; Martens, 1987; White, 1996; White et al., 1991). Pre-tests were administered in September before students began their five months of intensive ESL instruction; immediate post-tests were administered in January at the end of the intensive ESL portion of the school year; and delayed post-tests were administered in June after the students had had five months of intensive subject matter instruction in French (i.e., language arts, math, social studies, etc.).

Paper-and-pencil Tasks

The paper-and-pencil tasks included group-administered tests of listening and reading comprehension, vocabulary recognition, and grammaticality judgments. The tests chosen for this study are described below; more details about each test are found in Appendix E.

Aural Vocabulary Recognition test.

The Aural Vocabulary Recognition (AVR) test consists of pages showing eight pictures to which a letter of the alphabet is attached in random order. From an audio tape, learners hear a word pronounced once (e.g., 'carry') and have to indicate on an answer sheet the letter corresponding to the word heard. There were 64 items on this test when given as a pre-test; as an immediate post-test, 16 additional items were included.

Listening Comprehension test ('Test de compréhension auditive').

For this test students demonstrate their comprehension of short sentences by choosing an appropriate picture. From a tape, students hear ten sentences twice, and have

to find which of twelve pictures shown on a single page best matches the content of the utterance. For example, they hear "He's looking for his key " and have to match it to a picture of a man standing by a car, searching for something in his pocket. The student writes the letter corresponding to the picture on an answer sheet. The stimulus sentences and pictures do not involve any grammatical contrasts such as a single item vs. plural items. This test was administered as a pre-test only.

Cloze test.

This is a rational deletion cloze consisting of a short passage of 54 words with ten blanks. The scoring procedure allowed for any plausible word choice and great lee-way was given for spelling. Thus, for the target word 'school', responses such as 'scool' or 'skool' were accepted. The Cloze was administered as a pre-test and an immediate post-test.

MEQ Listening/Reading Comprehension test.

This is a test that was produced in the early 1980's by the Ministère de l'Education du Québec (MEQ) as part of the needs analysis preceding the adoption of a province-wide ESL program based on a communicative teaching approach. It was administered to Secondary III (grade 9) students in 1982, and has been used by Lightbown and Spada in studies involving intensive ESL learners at the grade 5 and grade 6 level for many years (see Lightbown & Spada, 1997 for a report containing comparisons among some 40 groups on this measure.) Two forms of this test were used; the 32-item version was used as an immediate post-test in January, and the 53-item version was used as a delayed post-test in June. The format of the MEQ test is multiple choice; there are three distracters for each of the items. While the tasks vary in each of the six sections, all require the students to listen to recorded English statements, questions, or descriptions and to select the best picture, statement, or response shown in their test booklets.

120

Grammaticality Judgment task.

The Grammaticality Judgment task was developed by Martens (1988) in her study of intensive ESL learners and their knowledge of third person possessive determiners. Unlike the MEQ test which samples from a wide range of English structures and usage, this task focusses on a specific aspect of the L2 that is known to be developmental in nature and difficult for francophones (Lightbown & Spada, 1990; White, 1996; Zobl 1985). On this grammaticality judgment task, students are given a little booklet which presents the story of a boy who is celebrating his birthday. The text contains 32 different errors of which 21 involve third person singular possessive determiners. The other 11 errors were distractors, involving different aspects of English. Students were told that there were some errors in the text and that they were to put a cross through the error; they were not required to correct the error. An overall score was determined by counting all the identified errors. The test was administered as an immediate post-test.

Yes/No Vocabulary test.

The Yes/No Vocabulary Test is based on the tests developed by Meara to measure the size of the vocabulary of ESL learners (Meara & Buxton, 1987; Meara & Jones, 1988; Meara, Lightbown, & Halter, 1994). A similar test was used by White (1996) in her study conducted in intensive ESL classes. Students are presented with a list of frequent English words (e.g., foot, west, thin), and of plausible-looking non-words such as 'jarvis', 'churchlow', 'savourite'; one third of the 180 items are non-words. The test-takers are asked to put a check mark beside the words whose meaning they "know". They are told that they will be penalized for guessing. A score is derived based on signal detection theory in which the number of 'hits' (i.e., when a real English word is checked off) are adjusted in relation to the number of 'false alarms' (i.e., when a non-word is guessed to be a real word).

Oral Production Tasks

In addition to the pencil-and-paper tests, oral production data were collected from a sample of ten randomly selected students from each of the five ESL classes. The selected students took part in an individual interview in January 1995 at the end of the five month ESL component of the academic year. Each selected student participated in two activities designed to elicit particular language forms. These student interviews were audio-recorded and later transcribed for analysis. The oral production tasks are described below.

'Five Questions' picture card game.

This task was first developed by L. White et al. (1991) for use in a study involving intensive ESL students. To play the game, the student is given four similar picture cards and the interviewer chooses one card from a second set of the same pictures. Neither shows their picture to the other. The learner is required to ask five questions in order to determine which of the four pictures the interviewer has. There were three different picture sets used in all, not including the first practice set in which the interviewer took the role of the questioner and modeled the type of questions that could be asked. The learner was allowed to ask both yes/no and wh-questions; if the student did not use a wh-question during the first two picture sets, the interviewer prompted for wh-questions (e.g., saying something like, 'Ask me a question with 'where'.). See Appendix F for examples of the pictures and a sample of student production on this task .

Possessive determiner elicitation.

A task to elicit third person possessive determiners was developed by White (1996) for use with similar ESL students. The learner is asked to describe cartoon pictures which are from the For Better or for Worse series (Johnston, 1977, 1985). Each cartoon shows a child with one or two parents in the midst of a familiar problem or predicament: a mother

reacts unhappily to her daughter's self-administered haircut; a father shows sympathy to his young child covered in unnecessary band-aids while the mother smiles behind her hand; a sequence of six pictures in which a child falls off his tricycle and waits to cry until he sees his mother; a girl explains to her parents why she and the bedroom are covered with make-up. For each picture, the interviewer asked the student: "What is the problem?" or "Why is this cartoon funny?" while both student and interviewer looked at the picture. Answering these questions usually led the student to create several obligatory contexts for his and her (e.g., She cut her hair and her mother is upset). An example of the pictures, and a sample from a student transcript are provided in Appendix G.

Classification of Paper-and-pencil Tasks

The research hypotheses for the first two research questions propose that there will be differences in learners' performance depending on the nature of the L2 tasks, specifically whether they require the learner to pay attention to the grammatical form of language in addition to the meaning. Thus, the Aural Vocabulary Recognition test, the September Listening Comprehension test, and the Yes/No Vocabulary test are not considered form-focussed because they test vocabulary in ways that do not require grammatical analysis on the part of the learner. The other paper-and-pencil tests administered in this study vary in the degree to which they are form-focussed. The Grammaticality Judgment task, the MEQ Listening/Reading Comprehension test, and the Cloze all involve reading and thus require the learner to focus on the written form of English. In addition, on each of these tasks, the learner must make a judgment concerning formal aspects of the L2. This is clearest in the Grammaticality Judgment task where sentences are judged to be correct or incorrect, and on the multiple-choice MEQ test where the learner must decide which of the given responses is the appropriate one. On the Cloze, we can also consider the learner to be making a judgment relating to language form: if a

learner can think of more than one word to fill the blank, he or she has to decide which word offers the best syntactic and semantic 'fit'.

Data Collection Schedule

Not all of the measures were administered at all testing sessions. Testing of the students took place at the beginning of the school year in September 1994 (pre-tests), at the end of their ESL program in January 1995 (immediate post-tests), and then at the end of the school year in June 1995 (the delayed post-tests). The oral production task was administered to a sample of students in January only (see Table 4.2 for an overview of the data collection). It is particularly important to emphasize that the L1 grammatical sensitivity task 'Trouve les erreurs' was administered in June, that is, at the end of the intensive program, rather than at the beginning. This decision was made for practical reasons: there is a general prohibition on the use of French in the ESL classrooms, so it was decided that the grammatical sensitivity test would be administered during the French phase of the year.

Table 4.2

Testing Schedule for September-January ESL Cohort, 1994-95 (N=150)

September 1994	January 1995	June 1995
• Aural Vocabulary Recognition-pre	• Aural Vocabulary Recognition-post	• Yes/No Vocabulary
• Listening Comprehension	• Grammaticality Judgment	• MEQ-53-items
• Cloze-pre	• MEQ-32-items	• Trouve les Erreurs
	• Cloze-post	• Learning Preferences Questionnaire
	• Oral elicitation tasks for Questions and Possessive Determiners (n=50)	

124

The following sections deal with the procedures used to analyze the data in order to address the three research questions.

Data Analysis Procedures

As discussed in Chapter 2, individual differences in L2 acquisition have been studied using different research methodologies, depending on whether the researcher adopts a nomothetic or ideographic approach. The distinction often boils down to the use of quantitative vs. qualitative data analysis techniques. Both quantitative and qualitative analyses were used to address the research questions in the present study. These were principal components analysis, cluster analysis, and interlanguage analysis.

Principal Components Analysis

Principal components analysis was used to address the first research question concerning the relationship between grammatical sensitivity and L2 proficiency. Many of the studies reviewed in Chapter 3 used correlation which is the prototypical analytic procedure for investigating the relationship between two variables. However, when many variables are involved, it is preferable to use some type of multivariate analytic technique such as principal components analysis in order to reduce a complex set of correlations into fewer dimensions. This makes it easier to determine how many distinct abilities underlie an array of language test scores, what these abilities are, and what contribution each test makes to the measurement of each ability (Woods, Fletcher and Hughes, 1986). Geometrically, the procedure involves placing axes through clusters of data points plotted in a multidimensional space (Gould, 1996). The first axis, or the first principal component, represents the largest amount of variance in the data; the second axis lies at right angles to the first and accounts for the second largest amount of the remaining variance, the next axis

is at right angles to the first two and so on with each axis orthogonal to the preceding ones, and each encompassing steadily decreasing amounts of information.

Many introductions to principal components analysis contrast it with factor analysis or common factor analysis as it is sometimes called (Stevens, 1986; Woods, 1983):

> In both approaches linear combinations of the original variables (factors) are derived, and often a small number of these account for most of the variation or the pattern of correlations. In factor analysis a mathematical model is set up, and the factors can only be estimated, whereas in components analysis we are simply transforming the original variables into the new set of linear combinations (the principal components). (Stevens, 1986, p.337-338)

While both methods yield similar results, there are many who argue that principal components analysis is the better choice (see Stevens, 1986, p. 338; Wilkinson et. al.. 1992, p. 76-79) because of the indeterminacy problem associated with the common factor model[1].

After the components are extracted, the next step is to examine how the components are related to the original variables. Component loadings are the correlations between each component and each variable from the data set. By examining these correlations the researcher can name the construct underlying the component.

> if the first component loaded high and positive on variables 1, 3, 5, and 6, then we would interpret that component by attempting to determine what those four variables have in common. The component procedure has empirically clustered the four variables, and the job of the psychologist is to give a name to the construct that underlies variability and thus identify the component substantively.
> (Stevens, 1986, p. 339)

Thus, for example, we might interpret a component as 'auditory ability' if it loaded highly (by convention, above r = .30) on different listening tests and not on measures of writing skills. Interpretation of the components is facilitated by rotating the axes; for example, using the orthogonal rotation known as Varimax, each factor tends to load high on a smaller number of variables and low or very low on the other variables (Stevens, 1986). In

126

the present study, the principal components analysis was conducted using the Factor program of SPSS-X 6.1.

Cluster Analysis

The cluster analysis technique was used in order to establish the learners' profiles for the second research question. Cluster analysis is a procedure (or more accurately, a group of procedures) that assembles entities into piles according to their similarities or dissimilarities. Like the principal components analysis, it offers the researcher a way of dealing with a multivariate dataset. The cluster analysis technique has been championed by Skehan (1986a, 1989) for use in the study of individual differences since it allows the researcher to identify subgroups which are maximally similar to each other, and different from other subgroups. Unlike factor analysis and multiple regression, cluster analysis is not limited by the assumption that variables are related to each other in a linear or an additive fashion.

In general, a cluster analysis consists of the following steps (Lorr, 1983; Romesburg, 1984):

1. assemble a set of objects (i.e., individuals), measure their attributes and enter the data into a data matrix;

2. standardize the data, if necessary;

3. choose a measure of similarity/dissimilarity and a clustering method;

4. decide on the number of subgroups present in the data.

A cluster analysis involves grouping objects that are similar in terms of their geometric characteristics when represented as points in some "attribute space". Dissimilarities between objects correspond to the metric distance between the respective points (Lorr, 1983). In this study, the Ward's method (Ward, 1963) which uses Euclidean distance between cases as an index of similarity was chosen as the clustering method. This was the method used by Skehan (1986a) in his study of learner profiles. The data matrix was

cluster analyzed using the Cluster program available on SPSS-6.1. After an appropriate cluster solution was chosen, the means for each cluster were calculated and submitted to a one-way Analysis of Variance for each of the eight dependent variables.

Interlanguage Analysis

The third research question was addressed by the qualitative analysis of the oral production data. Rather than looking at learner production as being correct or incorrect, interlanguage analysis involves looking for evidence of development in learners' non-target-like language system. Thus, an overgeneralization error such as 'He speaked' indicates that the learner has internalized the rule for marking the past in regular verbs, but has not yet acquired the full system of tense marking that covers irregular verbs. This approach is particularly useful for measuring progress in the early stages of L2 acquisition. In order to facilitate comparisons between learners, an index of development is needed (Larsen-Freeman & Long, 1991). In this study, two grammatical features of English were chosen as indices of development -- question forms and third person possessive determiners. For both of these target features, it is possible to classify a learner's production in terms of a developmental stage within a framework of stages. Both analytic frameworks have been used in studies conducted with learners in intensive ESL classrooms in Quebec (for question forms: Spada & Lightbown, 1993; for possessive determiners: White, 1996, 1998).

Question Form Stages

It is widely reported in the L1 and L2 literature that the rules of question forms in English are acquired in predictable stages (e.g., Brown, 1973; Ravem,1973; Wode, 1978). The descriptions of the developmental stages used in this study (see Table 4.3) are from Spada and Lightbown (1993), based on the work of Pienemann, Johnston and Brindley (1988). Progression from Stage 1 to Stage 5 is characterized by increasing mastery over the

rules for subject-auxiliary inversion in yes/no and wh-questions[2]. Each question was coded, and then two scores were obtained: highest stage and most frequent stage. A learner's 'highest stage' score was determined by identifying the highest stage for which there were at least two different examples.

Table 4.3

Developmental Stages for Question Formation in English (from Spada & Lightbown, 1993)

Stage 1	Single words or fragments A spot on the dog? A ball or a shoe?
Stage 2	SVO with rising intonation A boy throw the ball? Two children ride a bicycle?
Stage 3	Do-fronting Do the boy is beside the bus? Do you have three astronaut? Wh-fronting What the boy is throwing? Where the children are standing?
Stage 4	Wh- with copula BE Where is the ball? Where is the space ship? Yes/No questions with aux inversion Is the boy beside the garbage can? Is there a dog on the bus?
Stage 5	Wh- with auxiliary second What is the boy throwing? How do you say "lancer"?

Possessive Determiner Stages

The speech elicited by the second oral production task was analyzed according to the framework of acquisition stages for third person singular possessive determiners developed by White (1996; 1998). The stages are based on the theoretical and empirical work of several researchers (Felix, 1981, Felix & Hahn, 1985; Lightbown & Spada, 1990; Zobl, 1984, 1985). The stages in White's framework describe the learner's evolving ability to produce his/her during a communicative task. The choice between his and her is difficult for French-speaking learners of English because of the way the two languages differ in their assignment of gender. In French the choice between masculine and feminine third person determiners is based on the grammatical gender of the possessed noun (e.g., 'Robert voit sa mère'), while in English, the choice depends on the natural gender of the possessor ('Robert sees his mother'). In White's study, this framework adequately captured developmental changes in the use of his/her among learners similar to those of the present study.

Each student was assigned to a stage according to the descriptions presented in Table 4.4. In general, the procedure used by White was followed. However, since the data samples relevant to the possessive determiner analysis was smaller than those used in her study, the criterion used for Stages 4-7 was reduced from four instances to three.

For both the question form and the possessive determiner analysis, the transcripts were coded by one researcher and then checked by another to confirm reliability in the coding. All differences in stage assignment were discussed until agreement was reached.

130

Table 4.4

Developmental Stages in the Acquisition of English Possessive Determiners by Francophone Learners*

Stage 1	pre-emergence: avoidance of his and her and/or use of definite article
Stage 2	pre-emergence: use of your for all persons, genders and numbers
Stage 3	emergence of either or both his and her
Stage 4	preference for his or her (accompanied by over generalization to contexts for the other form)
Stage 5	differentiated use of both his and her (not with kin-different gender)
Stage 6	agreement rule applied to either his or her (kin-different gender)
Stage 7	agreement rule applied to both his and her (kin-different gender)
Stage 8	error-free application of agreement rule to his and her (all domains, including body parts)

*Adapted from White (1996, p. 179)

This chapter has described the data collection and data analysis procedures. The next chapter presents the results of the analyses with respect to each of the three research questions.

NOTES

[1] There is also considerable confusion in terminology. See Gould (1996, p. 276) for a useful discussion of this.

[2] The stage framework of Pienemann et al. (1988) includes a sixth stage, which involes embedded questions in which the subject and auxiliary are uninverted. The learners in this study were beginners and did not produce such questions.

CHAPTER 5

RESULTS

This chapter begins with a brief overview of the test and questionnaire results. This is followed by presentation of the results from the principal components analysis, the cluster analysis, and then the interlanguage analysis. In each case, the results are examined with respect to the relevant research question and hypothesis.

Test and Questionnaire Results

The means and standard deviations for all the tests administered to the participants are presented in Table 5.1 All were included in later analyses except for the Aural Vocabulary Recognition test administered in January. As the mean of 89% (54/64) indicates, this test proved to be very easy for all of the students who had just completed five months of intensive ESL instruction. Therefore, it was not included in the correlational analysis, the principal components analysis, or the cluster analysis. The reliability estimates[1] for the various measures are given in Appendix H.

The Learning Preferences Questionnaire was designed to help identify learners whose learner style was analytic or metalinguistic in nature. The distribution of student responses is given in Table 5.2 below. In order to interpret the results of the questionnaire, items were grouped according to whether they were more communication-oriented or more form-oriented in nature. Note that activities are classified as being relatively more form-oriented when they required the learner to attend to linguistic form either exclusively or in addition to meaning. Thus, the items 'reading stories' and 'writing texts' were classified as being form-oriented because learners have to deal with the written form of words in order to extract or express meaning.

Table 5.1

Descriptive Statistics for September, January and June Tests

Measure	N	Maximum possible score	Mean	Standard Deviation
Aural Vocabulary Recognition-pre (Sept)	149	64	35.13	8.06
Listening Test (Sept)	149	10	6.28	2.04
Cloze-pre (Sept)	149	10	2.12	1.76
Aural Vocabulary Recognition-post (Jan)	146	64	57.15	8.15
Cloze-post (Jan)	146	10	7.13	1.52
MEQ-32 items (Jan)	146	32	23.98	3.91
Grammaticality Judgment task (June)	145	32	12.14	7.39
MEQ-53 items (June)	146	53	44.28	5.87
Yes/No Vocabulary Test (June)	144	100	72.7	15.0
'Trouve les Erreurs' (June)	146	20	13.27	2.87

Note. The 'N' varies due to students being absent for one of the tests.

The distribution of responses to the questionnaire shows that these students prefer communicative over form-focussed activities. On average, two-thirds of the students reported that communicative activities were very helpful for learning English. Only 8% of the students used the low end of the scale for these activities. This indicates their overall satisfaction with the program since students reported learning best from the kind of

activities they had experienced in their ESL class (note that they completed the questionnaire in June, five months after their ESL class had ended).

Table 5.2

Distribution (in Percentages) of Responses on Learning Preferences Questionnaire (N= 147)

Communication-oriented activities:	Students responses (in %)			
	'beaucoup'	'un peu'	'très peu'	'pas du tout'
play games	51	39	6	5
listen to cassettes	42	35	20	3
do oral activities	72	23	4	2
watch videos	84	14	1	1
perform skits	68	25	5	1
watch television	85	13	1	1
listen to the teacher's explanations	70	20	3	3
Mean for communication-oriented activities	**66**	**5**	**6**	**2**
Form-oriented activities:				
read stories	59	33	5	3
write texts	38	49	10	4
recite tongue-twisters	19	36	27	18
get corrected by the teacher	46	36	12	5
study vocabulary lists	38	45	13	5
look up words in dictionary	46	33	15	6
consult grammar book	16	26	33	26
do grammar exercises	16	44	24	15
Mean for form-oriented activities	**31**	**38**	**19**	**11**

Note. As all percentages are rounded up to the nearest whole number, some row and column totals do not add up to 100.

The responses to the form-focussed activities reveal a much greater spread. On average, about a third of the students felt that these activities were of little or no help to them in learning English. This includes both activities which they had experienced in their ESL class (e.g., reciting tongue twisters and using the dictionary), and others which were not part of the ESL program (e.g., consulting a grammar book). This relatively greater dissatisfaction with form-focussed tasks could be interpreted as reflecting their preference for 'fun' activities over 'serious' ones, or easier tasks over more effortful ones.

In Chapter 4, it was hypothesized that analytic learners would rate metalinguistic tasks more favourably than would non-analytic learners. Is self-report a reliable method for classifying learners? From Table 5.2, we see that 16% of students reported that consulting a grammar book was very useful and a similar percentage responded in this way to the item 'doing grammar exercises'. To evaluate response consistency, learners' responses to the two metalinguistic items were compared. It was found that almost every student who responded very favourably on one also responded favourably on the other. However, almost half of the students who reported that they liked to consult grammar books scored below the mean on Trouve les Erreurs. The reverse was also true, students who displayed analytic ability through their performance on Trouve les Erreurs did not necessarily prefer consulting grammar books or doing grammar exercises. It thus appears that the self-report questionnaire was not a reliable way to obtain information about learners' analytic ability since the reported preference did not necessarily reflect learners' performance on an analytic task. On the basis of these observations, it was decided not to use the questionnaire results in the statistical analyses.

We now turn to the results of the principal components analysis, the cluster analysis and the interlanguage analysis with respect to the three research questions and hypotheses.

Research Question #1

The first research question relates to the overall relationship between grammatical sensitivity and L2 test performance. Specifically, the question is as follows:

Is grammatical sensitivity associated with performance on measures of L2 proficiency when learning takes place in a communication-oriented instructional environment?

This question is addressed by the results of the principal components analysis.

Preliminary Analyses

A principal components analysis uses a variance-covariance matrix as its input. As with simple correlations, it is sensitive to the influence of outliers. Preliminary analysis of the data included Pearson correlations and scatterplots in order to identify outliers. This led to the scores of three students being excluded from the analysis. In addition, individuals who had negative scores on the Yes/No Vocabulary Test were excluded; a negative score on this measure indicates that the test taker has accepted more non-words than real words, making the score uninterpretable.

The Pearson correlation matrix is displayed in Table 5.3. Most of the correlations are significant which is not surprising given the sample size (N = 133 with listwise deletion). The French Trouve les Erreurs measure correlates moderately with the pre-test and post-test Cloze, the June administration of the MEQ Listening/Reading Comprehension Test, the Grammaticality Judgment task, and the Yes/No Vocabulary Test. It correlates weakly with the AVR pre-test, and the January MEQ test. These correlations are consistent with the hypothesis that grammatical sensitivity will be associated with performance on form-focussed L2 proficiency measures. The next phase of the analysis made it possible to get a clearer view of the relationships among the variables.

Table 5.3

Pearson Correlations Among Test Variables (N =133)

	AVR	Listen. Sept	Cloze Sept	Cloze Jan	MEQ Jan	MEQ June	Trouve Erreurs	Gram. Judg.	Yes/ No
AVR									
Listen. Sept	.58								
Cloze- Sept	.55	.33							
Cloze- Jan	.46	.43	.46						
MEQ-Jan	.41	.36	.43	.48					
MEQ- June	.61	.45	.45	.56	.59				
Trouve Erreurs	.28	.08#	.39	.40	.26	.32			
Gram. Judg.	.40	.32	.38	.60	.57	.58	.42		
Yes/ No	.46	.26	.36	.51	.47	.60	.34	.53	

Note: All correlations are significant at p< .01 except for the one indicated by a #.

Principal Components Analysis

After the initial factor^2 structure is extracted from the variance-covariance matrix, the researcher must decide how many components are to be retained. Initially the number of components in the output of the analysis is equal to the number of variables (Grimm &Yarnold, 1995; Stevens, 1986). A number of conventions have been established to guide this decision. One may choose to (1) retain only those components with eigenvalues (i.e.,

the amount of standardized variance accounted for by the extracted component) greater than 1, or (2) examine the visual plot of the eigenvalues (the scree test) and retain those eigenvalues in the sharp descent before the values level off, or (3) retain as many as will account for a specified amount of variance (generally, at least 70% of the total variance). Stevens (1986) cautions that blind observance of 'stopping rules' could lead to the elimination of components which account for less total variance but may have practical significance (p. 341).

Using the first two procedures with the intensive ESL test data, the first two components would be retained, accounting for 63.3% of total variance. It was then decided to include the third component in order to bring the total amount of variance accounted for up to 72.1%. (The initial output of the principal components analysis and scree plot are displayed in Appendix I.) The three components were rotated using the Varimax rotation in order to arrive at loadings which would be more easily interpretable. The matrix of rotated factors is given in Table 5.4. The factor loadings (i.e., the correlation between each variable and each component) which are at least .40 and stronger are in boldface. Note that once the factors are rotated, it is no longer the case that the first component necessarily accounts for the largest amount of variance in scores.

Table 5.4

Principal Components of Test Variables with Varimax Rotation

Variable	Factor 1	Factor 2	Factor 3
AVR-1 (Sept)	.32	**.80**	.18
Listening (Sept)	.27	**.81**	-.15
Cloze-1 (Sept)	.17	**.67**	**.50**
Cloze-2 (Jan)	**.57**	**.40**	.34
MEQ-32 (Jan)	**.75**	.31	.06
MEQ-53 (June)	**.73**	**.45**	.10
Trouve les Erreurs (June)	.25	.02	**.89**
Grammaticality Judgment (Jan)	**.75**	.18	.31
Yes/No Vocabulary (June)	**.80**	.13	.17

As described in Chapter 4, the interpretation of the components is arrived at by making inferences based on the significant component loadings. The highest loading variables are the ones that contribute most to defining the component. According to Hypothesis #1, we would expect that Trouve les Erreurs would load on the same component as tests that require a focus on form.

Factor 1.

This component loads significantly on all the post-test L2 variables but not on the pretest ones. The highest loadings are on the Yes/No Vocabulary Test, both administrations of the MEQ test, and the Grammaticality Judgment task. It also loads to a lesser degree on the January administration of the Cloze. The grammatical sensitivity measure did not load together with the form-focussed proficiency measures. This factor will be referred to as 'Post-test English Proficiency'.

Factor 2.

The second component loads significantly on the two pre-test listening tests (the AVR and Listening), the September and January administrations of the Cloze, and the June (53-item) but not the January (32-item) version of the MEQ Listening/Reading Comprehension Test. The highest loadings are on the two listening tests given in September, and the lowest are on the January Cloze and the June MEQ test. The high loadings on the pre-test variables motivate calling this factor 'Pre-test English Proficiency'. The post-tests which are also associated with this factor must involve some of the same underlying ability. This will be discussed further in Chapter 6.

Factor 3.

This component appears to be a 'Grammatical Sensitivity' factor since it loads very highly on Trouve les Erreurs; it does not load on most of the L2 variables, which means that it reflects mainly variable-specific variance. However, it does load significantly on the pre-test Cloze, accounting for 25% of the variance explained by the component. This is the only form-focussed task that is associated with performance on the grammatical sensitivity measure; the loading for the Grammaticality Judgment task misses the .40 cutoff for practical significance.

Evaluation of Hypothesis 1

Taking into consideration the results of the principal components analysis, we can evaluate the first research hypothesis, according to which:

There is an association between grammatical sensitivity and L2 proficiency

measures that are form-focussed in nature.

The factors extracted by the principal components analysis are considered to be independent of each other; 'Pre-test English Proficiency', 'Post-test English Proficiency', and

'Grammatical Sensitivity' are separate factors in accounting for the performance of these learners on the tests administered. Consequently, it appears that Hypothesis 1 is not supported by the results of this analysis. It was hypothesized that grammatical sensitivity would be associated with performance on form-focussed L2 measures. In fact, Trouve les Erreurs did not share variance with any of the post-test variables. Indeed, the 'Grammatical Sensitivity' factor is largely variable-specific, loading very highly on the French error correction task and not on much else. While it did load on the pre-test Cloze, which is considered a form-focussed task, this correlation does not provide sufficient evidence to support the hypothesized relationship.

Research Question #2

The second research question examines the significance of grammatical sensitivity when learners are grouped according to their learning profile using the technique of cluster analysis. Specifically, the question is:

What is the relationship between grammatical sensitivity and the profiles of successful vs. less successful L2 learners?

Cluster Analysis

Before examining the output from the cluster analysis, a few points about this type of analysis are in order. Cluster analysis is a mathematical way of sorting a collection of objects (or individuals) into piles. As with any kind of sorting procedure, it is possible to group a set of objects/individuals according to any number of justifiable criteria. The cluster statistical program produces an output which provides solutions for any given number of clusters from 1 to n. In this case, this means that the solutions range from having each of the 135 individuals considered as a unique cluster of one, to having one cluster which includes all 135 individuals. The task of the analyst is to find the best solution somewhere in between those extremes. One method is to use the coefficient of the squared Euclidean distance provided on the printout (Norusis, 1985; Skehan, 1986a). These values indicate

141

how homogenous the clusters are at each merging of the data; a sharp increase can be used as an indicator that further agglomeration is 'forcing' the data. The coefficient values for the last stages of the clustering process are presented in Table 5.5.

As Skehan has pointed out, the choice of the solution is not determinant (hence its bad reputation among some statisticians). Inspection of the increase in coefficient values reveals a number of possible breaking points as the agglomeration process moves from many clusters to fewer clusters (i.e., moving down the table). Three breaking points present themselves as being possible candidates: between 7 and 6 clusters, between 4 and 3, and between 3 and 2. Skehan (1986a) argues that a small number of clusters is generally easier to interpret. While some of the learner style literature has identified three learner types (i.e., analysis-oriented, memory-oriented, and flat profile), the large jump in the coefficient indicates that the move from four clusters to three is somehow 'forcing' the data.

Consequently, on the basis of the change in the coefficient and a visual inspection of the clusters on a dendrogram, the four-cluster solution was chosen. This solution provided four groups of sufficient size to make statistical comparisons possible ($\underline{n}$ = 53, 41, 28, 13, respectively). The means of the four clusters on the nine variables are presented in Table 5.6. One-way analyses of variance were performed on the cluster means for each of the nine variables; to avoid the increased probability of Type I errors when doing multiple one-way ANOVA's, the alpha level was set at .001. The Tukey HSD post-hoc test of significance was used to determine which clusters were different from each other. The nine ANOVA tables are found in Appendix I. It should be noted that that the numbering of the clusters is not meaningful in itself but simply reflects the point at which one cluster breaks off from another.

Table 5.5

Increase in the Error Sum of Squares

Number of clusters	Last increase in coefficient (squared Euclidean distance)
12	17.414
11	17.798
10	20.516
9	22.528
8	24.278
7	**28.164**
6	**45.663**
5	47.829
4	**52.252**
3	**113.903**
2	409.326
1	1221.607

Table 5.6

Means and Standard Deviations for Test Variables by Cluster

Measure	Max. score	Cluster 1 (n=53)	Cluster 2 (n=41)	Cluster 3 (n=28)	Cluster 4 (n=13)
AVR-1 (Sept.)	64	40.9 (6.4)	30.0 (4.9)	36.8 (6.7)	23.8 (2.7)
Listening (Sept.)	10	7.5 (1.8)	**5.3** (1.5)	6.5 (1.5)	**4.1** (1.1)
MEQ-32 (Jan.)	100	83.4 (7.4)	**69.2** (9.9)	76.9 (8.5)	**63.2** (9.0)
MEQ-53 (June)	100	91.9 (4.7)	78.8 (7.3)	85.7 (6.3)	63.1 (11.9)
Trouve les Erreurs	20	15.1 (2.3)	13.4 (2.2)	**10.4** (2.0)	**11.3** (3.0)
Gram. Judgment (Jan.)	32	18.8 (4.5)	**9.4** (6.6)	**9.3** (5.4)	4.5 (3.3)
Yes/No Vocab. (June)	100	82.5 (7.0)	**69.4** (11.0)	**72.9** (12.0)	42.0 (11.0)
Cloze-1 (Sept.)	10	3.4 (1.8)	**1.3** (1.1)	**1.6** (1.4)	**0.9** (1.3)
Cloze-2 (Jan.)	10	8.4 (1.1)	**6.5** (1.1)	**6.5** (.9)	5.4 (1.0)

Note. Means in any row which are **NOT** significantly different from each other are underlined and in boldface.

144

Interpreting the Clusters

The next step is to characterize the clusters with respect to the pattern in test performance. Figure 5.1 presents the means for each cluster in z-scores. This means that it is possible to make comparisons across the different tasks. The two extremes -- high and low -- will be presented first and then the middle clusters. The representation of the clusters in Figure 5.1 is especially helpful for characterizing the profiles.

Cluster 1.

This cluster's mean scores rank highest on all the measures. As a group, they are significantly better on all the pre-tests, the immediate and delayed post-tests, and on Trouve les Erreurs in comparison to all the other clusters. Their performance profile is uniformly high whether the tasks are form-focussed or meaning-focussed.

Cluster 4.

This cluster has the lowest mean score on all the L2 pre- and post-tests. They also performed poorly on the grammatical sensitivity measure (Cluster 3 has a lower mean score but it is not significantly different from that of Cluster 4). A striking feature of their profile is their performance on the June MEQ test. Recall that the MEQ Listening/Reading Comprehension Test given in June was a longer version of the test that had been administered in January. Whereas the other three clusters improved their average performance on the second administration by approximately 10%, Cluster 4 learners as a group failed to improve their scores. This suggests that the learners in this cluster had difficulty working on their English while they were engaged in intensive academic work in French. This would also account for their poor performance on the Yes/No Vocabulary Test which was also given in June.
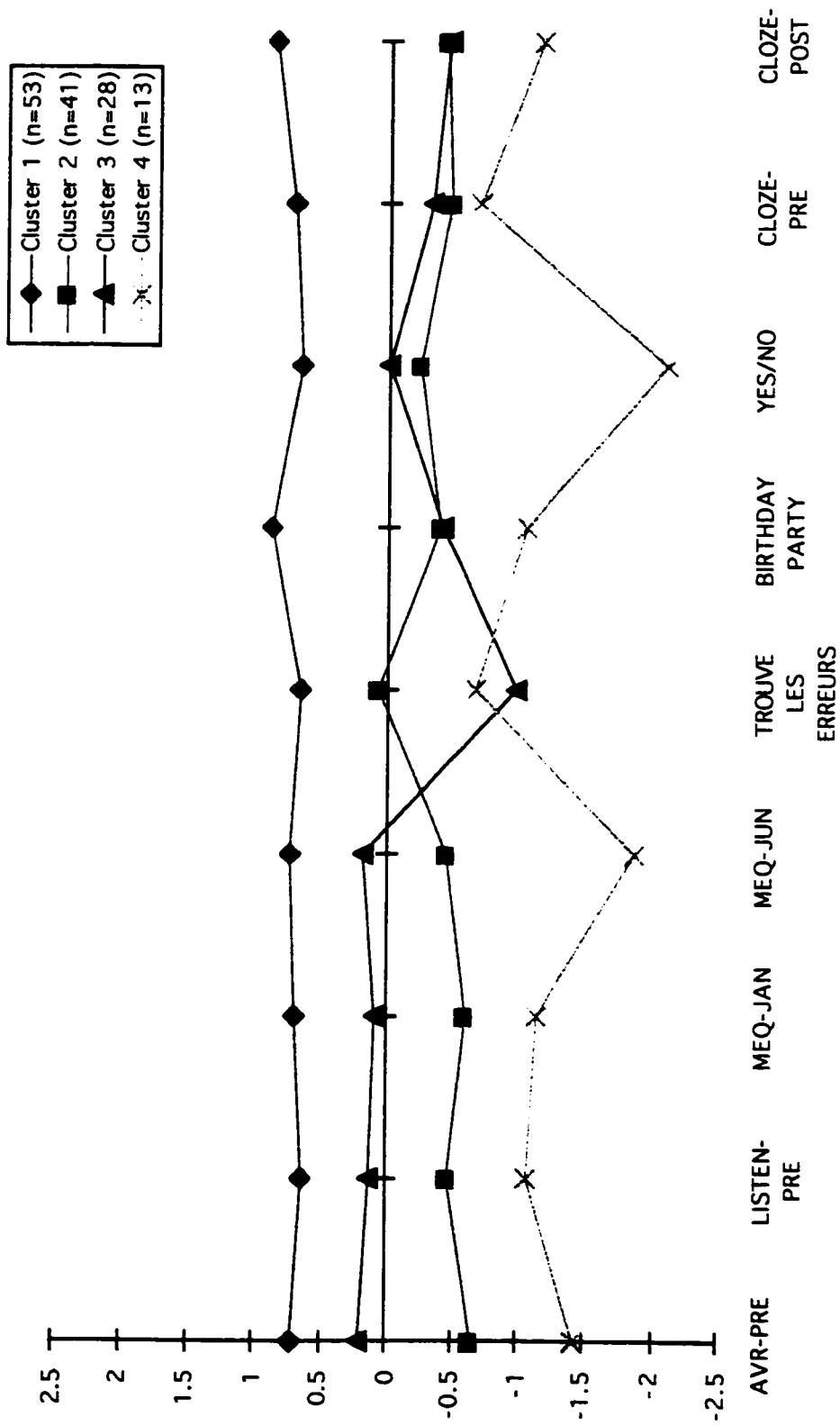
Figure 5.1 Profiles of four clusters based on standardized scores

146

Cluster 2.

Cluster 2 can also be characterized as a group of weak learners in this population. They performed at a level comparable to Cluster 4 on two listening measures (the Listening Comprehension pre-test and the MEQ test given in January), and on the pre-test Cloze. Unlike the learners in Cluster 4, they improved their MEQ test performance in June. Their scores on all the L2 measures were below the overall mean. On Trouve les Erreurs, however, Cluster 2 ranked second among the four clusters. They demonstrated a relatively greater ability to analyze French than the other group of weak learners. They were also better than Cluster 3 learners who otherwise had higher scores on the L2 measures.

Cluster 3.

In many ways, this cluster appears to consist of the 'average' learners. However, they do not have a 'flat profile'. They scored around the overall mean (cf. Table 5.1) on all the listening comprehension tasks (Aural Vocabulary Recognition-pre, Listening Comprehension-pre, and the January and June administrations of the MEQ test) and on the Yes/No Vocabulary Test. On the four listening tasks (but not the Yes/No Test), Cluster 3 learners are significantly better than Cluster 2 and Cluster 4 learners. This pattern contrasts with their performance on the September Cloze where they are not significantly different from either Cluster 2 or Cluster 4. Furthermore, on the January Cloze and on the Grammaticality Judgment task, Cluster 3 learners were not significantly better than the generally lower ranking Cluster 2 learners. This relatively poorer performance on form-focussed L2 tasks is paralleled by their performance on the grammatical sensitivity measure in French where they did less well than Cluster 2 and were not statistically different from Cluster 4. The strengths of learners in this cluster appear to lie in their listening and vocabulary skills rather than in grammatical analysis.

147

Evaluation of Hypothesis #2

The results of the cluster analysis allow us to evaluate the hypothesis for the second research question, which is as follows:

Higher levels of grammatical sensitivity distinguish successful learners from less successful learners when success is defined in terms of performance on form-focussed measures of L2 proficiency.

The performance profile of the highly successful Cluster 1 learners is consistent with the hypothesis: they are high in grammatical sensitivity and high on all the form-focussed proficiency measures. The profile of Cluster 4 learners is also consistent with the hypothesis because they ranked lowest on the form-focussed tasks and on the grammatical sensitivity measure. Similarly, Cluster 3 learners' performance on Trouve les Erreurs predicts their poor performance on the form-focussed tasks. However, Cluster 2 poses a problem for the hypothesis. These learners are weak on form-focussed L2 tasks, but average in grammatical sensitivity according to the L1 measure. Thus, grammatical sensitivity does not predict their performance on such tasks.

On the basis of these findings, we may conclude that Hypothesis #2 is only partially supported.

Research Question #3

The third research question examines the relationship between grammatical sensitivity and oral production. Specifically, the question is:

Is degree of grammatical sensitivity associated with stages of interlanguage development in oral production?

148

Interlanguage Analysis

In order to test the hypothesis that greater levels of grammatical sensitivity are associated with higher stages of interlanguage development in oral production, two tasks were administered to a sample of 50 students. One task was designed to elicit question forms, while the other elicited possessive determiners. Ten students from each of the five ESL classes were selected at random to be interviewed individually at the end of January. Due to missing test scores, oral analyses were performed on the transcripts from 47 students. The subjects were assigned into three groups according to whether they scored below the mean (7-11), around the mean (12-15), or above the mean (16-19) on the Trouve les Erreurs test. This led to a distribution of 11 students in what are referred to as the High Grammatical Sensitivity (GS) group, 25 in the Mid GS group and 11 in the Low GS group. The disproportionately large number of learners in the Mid GS group reflects the fact that, for the whole sample, scores clustered around the mean on this measure. Since the analysis involves looking at the number of students from each group for each stage of the two target grammatical features, the unbalanced group sizes pose a problem. The usual solution is to use percentages; however, this leads to distortions when the cell sizes are small, as in this case. Therefore it was decided to randomly choose 11 individuals from the pool of 25 Mid-GS learners using a table of random numbers. Thus the interlanguage analysis is based on the data from 33 participants, 11 in each group.

Question Forms

All of the subjects from the sample provided at least 15 questions for analysis. The results of the question analysis for the three GS groups are presented in Table 5.7 It can be seen from the table that High and the Low GS groups do not differ in any obvious way

from each other for either the highest question form stage or the most frequent question form. The data is presented visually in Figure 5.2

It is possible that the stage framework is not sufficiently sensitive to differences in learners' production of questions. The Pienemann, Johnston and Brindley (1988) stage framework is based on word order and not on morphological aspects of question forms. Thus a student who says "Do there is a dog on the bus?" and "Did she had a bus in your picture?" would both be considered stage 3 which is characterized as a do-fronting stage. This led White, Spada, Lightbown and Ranta (1991) to calculate syntactic accuracy scores for the learners in their study in order to measure the effect of form-focussed instruction.

In order to determine whether a qualitative difference in question forms is masked by the stage analysis, the yes/no questions produced by the high and low group GS groups were examined. The first step was to see whether they differed in accuracy by calculating the percentage of well-formed questions for each individual using the same criteria as in White et al. (1991). Since these learners are beginners, accuracy was measured in terms of fronting the appropriate auxiliary, but not considering tense, or person marking. Thus, 'is the dog have' is wrong, but 'did he has' is acceptable. This analysis revealed no differences between the groups: for both the High and the Low GS groups the accuracy scores ranged from 14 to 100% with a mean of 52.4% for the high and 52.9% for the low. Further examination of the transcripts for evidence of greater syntactic complexity did not reveal a difference between the groups.

Table 5.7

Distribution of Learners in Developmental Stages for Question Forms

GS score	Highest Stage			Most Frequent Stage		
	2	3	8	2	3	4
High* (n=11)	0	3	8	2	4	4
Mid* (n=11)	2	3	6	3	6	1
Low (n=11)	0	2	9	4	4	3

* A learner's 'most frequent question form stage' score is excluded because it was coded as 'other'.
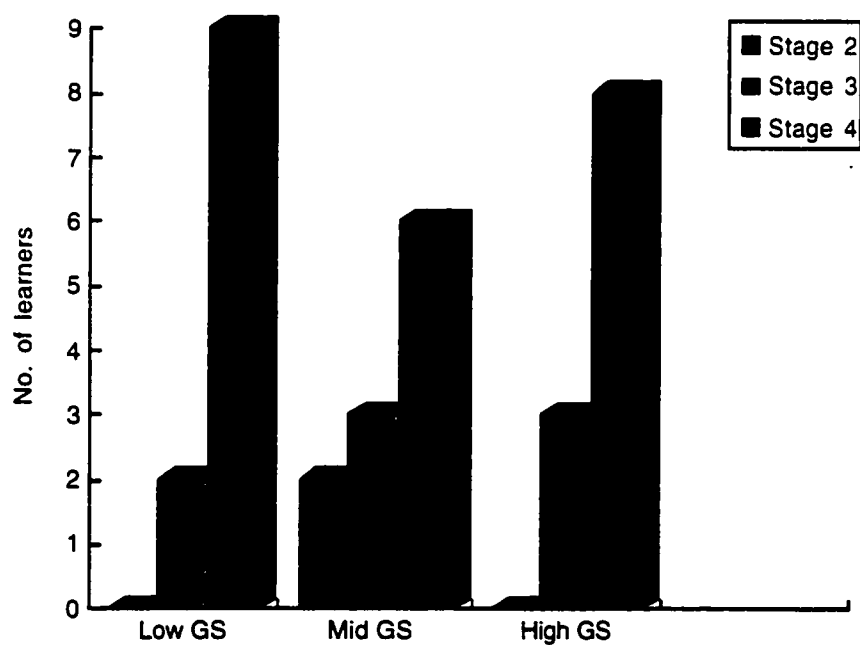
Figure 5.2 Distribution of highest question stage

Possessive Determiner Stages

The data used in the possessive determiner analysis come from the description of four cartoons; two of the pictures were biased towards contexts for <u>his</u> and two were biased towards contexts for <u>her</u>. The results are presented in Table 5.8. Here, in contrast to the results of the question analysis, the ability to use these possessive determiners appropriately tends to be associated with differences in grammatical sensitivity. There are more learners in the higher stages in the High GS group than in the Low GS group. This difference becomes clearer when the categories are collapsed along the lines recommended by J. White (1998, p. 105):

1. Pre-emergence: learners in Stages 1 and 2 do not use gender-marked possessive determiner forms

2. Emergence: learners in Stages 3 and 4 use his and/or her but show no evidence of applying the English agreement rule

3. Post-emergence: learners in Stages 5 to 7 gradually develop the ability to use the English agreement rule

4. Target-like performance: Stage 8 assumes targetlike use of the English possessive determiner rules in all semantic domains

Figure 5.3 illustrates the distribution of learners in the Pre-emergence, Emergence, and Post-emergence categories; no student in the sample reached Stage 8.

152

Table 5.8

Distribution of Learners in Developmental Stages for His/Her

GS Group	Developmental Stage						
	Pre-emergence		Emergence		Post-emergence		
	1	2	3	4	5	6	7
Low (n=11)	0	3	0	6	1	1	0
Mid (n=11)	1	1	2	6	1	0	0
High (n=11)	0	0	1	5	2	2	1
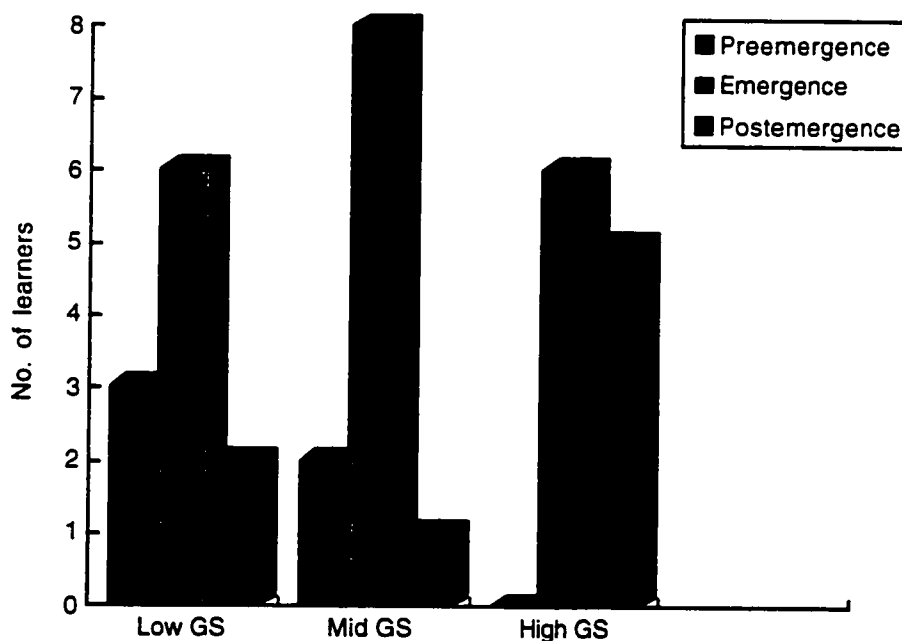
Figure 5.3 Possessive determiner stage

The High GS group contrasts with the Low GS group by (1) having no learners in the Pre-emergence stages, and (2) more learners in the Post-emergence stages. The low and mid

GS groups do not clearly differ from each other; both have some learners in the Pre-emergence stage and most of the others in the Emergence stages.

Cluster Membership as a Predictor

It appears, on the basis of the comparison between the High and Low GS groups, that the development of question forms in oral production does not reflect differences in learners' level of grammatical sensitivity but that the development of possessive determiners does. However, the small sample size available for the interlanguage analysis together with the modest reliability of the grammatical sensitivity measure (α = .62) may have led to some individuals being misclassified. An alternative procedure is to group the learners according to their cluster membership.

Cluster analysis, as we have seen, is a procedure that groups learners on the basis of similarity in patterns of scores. Appendix K displays the learners' classification according to their GS score and according to their cluster membership. Since students were chosen for the oral interview at random, the clusters are not equally represented in the sample. There are 15 oral samples from Cluster 1, 18 from Cluster 2, but only 8 from Cluster 3, and 5 from Cluster 4. The z value of each subsample mean was calculated (Welkowitz, Ewan, & Cohen, 1991) in order to confirm that the subsamples were not different from those of the 'parent' cluster. None of the four subsample means proved to be significantly different at the .05 level from their respective parent clusters. The distribution of learners into question form and possessive determiner stages by their cluster membership is presented in Table 5.9.

The most useful comparison to be made is between Cluster 1 and Cluster 2 which have similar cell sizes. Here it is possible to address Hypothesis #3 because Cluster 1 learners outperformed Cluster 2 learners on Trouve les Erreurs (see Table 5.6). Figures 5.4 and 5.5 display the results for the subsample from Clusters 1 and 2^3. The comparison of these learners reveals that the trend found for the possessive determiners in the initial

154

analysis is confirmed here: 40% of Cluster 1 learners are in the Post-emergence stages, while only 17% of the Cluster 2 learners are in the higher stages; no learners in Cluster 1 are in the Pre-emergence stages but 17% of Cluster 2 learners are. A similar pattern is found for the question stage analysis. A greater percentage of the learners in Cluster 1 produced Stage 4 questions, while no learners in this cluster had Stage 2 as their highest question stage. The learners in Cluster 2 had a more even distribution across the stage categories.

Table 5.9

Distribution of Learners in Developmental Stages by Cluster

	Questions: Highest stage	Possessive determiner stage
Cluster 1	Stage 2: 0	Stage 1-2: 0
(n=15)	Stage 3: 3 (20%)	Stage 3-4: 9 (60%)
	Stage 4: 12 (80%)	Stage 5-7: 6 (40%)
Cluster 2	Stage 2: 2 (11%)	Stage 1-2: 3 (17%)
(n=18)	Stage 3: 6 (33%)	Stage 3-4: 12 (67%)
	Stage 4: 9 (50%)	Stage 5-7: 3 (17%)
	Stage 5: 1 (6%)	
Cluster 3	Stage 2: 0	Stage 1-2: 1 (13%)
(n=8)	Stage 3: 2 (25%)	Stage 3-4: 5 (63%)
	Stage 4: 6 (75%)	Stage 5-7: 2 (25%)
Cluster 4	Stage 2: 1 (20%)	Stage 1-2: 2 (40%)
(n=5)	Stage 3: 1 (20%)	Stage 3-4: 3 (60%)
	Stage 4: 3 (60%)	Stage 5-7: 0
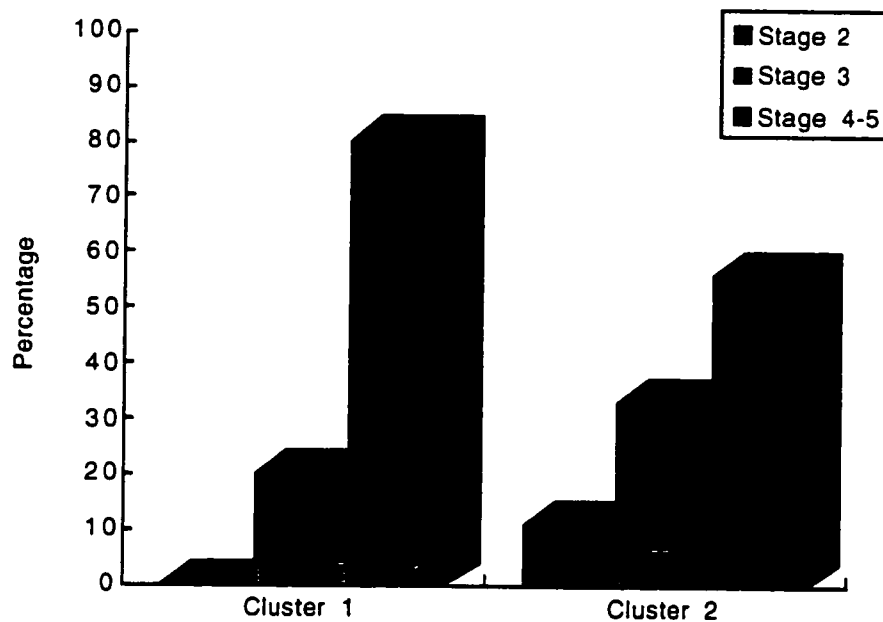
Figure 5.4 Highest question stage by cluster
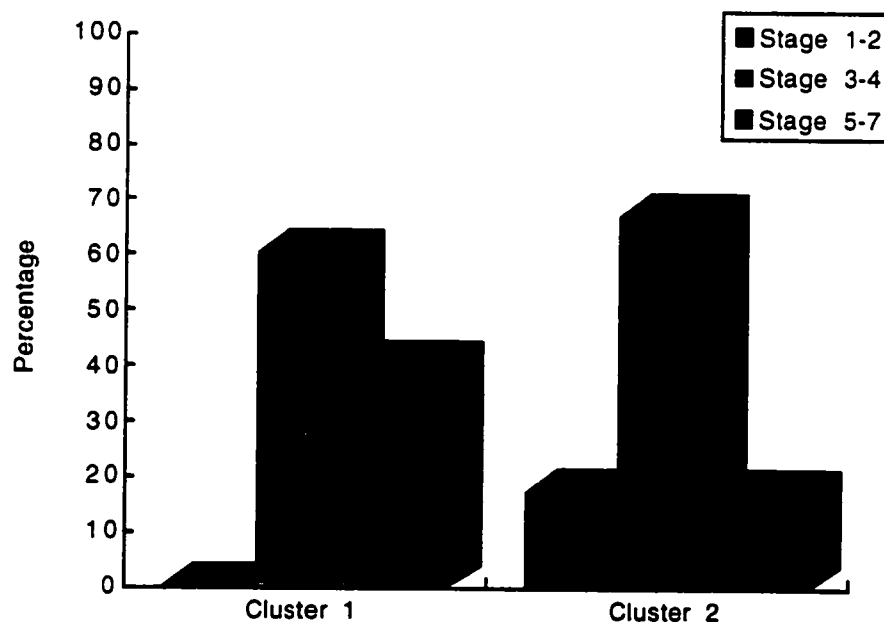
Figure 5.5 Possessive determiner stage by cluster

156

The results from the Cluster 3 and Cluster 4 subsamples must be considered with caution because of their small and unequal cell sizes. Recall that these two clusters were not significantly different from each other on Trouve les Erreurs. We would thus expect that their distribution in developmental stages should be comparable. This is certainly true for the question task where the distribution of learners in the two subsamples is very similar. On the possessive determiner task, however, there is an indication that the Cluster 3 learners are somewhat ahead since more of them are in the Post-emergence stages. But, as this difference reflects the performance of only two individuals, the contrast may be more apparent than real.

Finally, the comparison between Cluster 2 and Cluster 3 does not appear to support a relationship between grammatical sensitivity and interlanguage development. While the learners from Cluster 2 outperformed the learners from Cluster 3 on Trouve les Erreurs, this did not translate into superior performance on the oral tasks. Again, we cannot rule out the possibility that the disparity in the subsample sizes (18 vs. 5) may have obscured a pattern of differences between the two clusters.

Evaluation of Hypothesis #3

The hypothesis related to the analysis of the oral production data is as follows:

There is a relationship between learners' degree of grammatical sensitivity and their stage of L2 grammatical development in oral production such that higher levels of grammatical sensitivity are associated with higher stages of development.

This hypothesis is supported by the analysis of the possessive determiners in the initial analysis when learners were grouped by their GS score; it is also supported by the analysis of the subsample of learners from Clusters 1 and 2 with respect to both the question and possessive determiner forms. A greater percentage of learners demonstrating higher levels

157

of grammatical sensitivity were in the higher stages for both of these developmental features than was the case for learners with lower levels of grammatical sensitivity.

In Chapter 6, these findings will be further discussed in relation to previous research and to theoretical issues presented in earlier chapters.

NOTES

[1] An examination of Appendix H reveals that the reliability coefficients for some of the L2 measures are quite low. This is particularly the case for the Cloze tests. A number of factors are probably responsible for these low coefficients. On the one hand, the Kuder-Richardson 21 formula is a conservative estimate of reliability, which may result in coefficients which seriously underestimate the internal consistency of a test (Brown, 1996, p. 198-199). According to Brown, the likelihood of underestimation is greater on a cloze test than on a multiple-choice test. Two other factors which probably have contributed to the low reliability estimates are the small number of items used in the cloze passage (i.e., 10), and the very restricted range of writing ability of the students who provided the scores upon which the reliability estimate was made (Brown, 1996, p. 209; Schils, van der Poel, & Weltens, 1991). Note that no reliability coefficient is reported for the Yes/No Vocabulary Test due to the fact that there is no accepted method of estimating reliability for this type of task (Randall Halter, personal communication).

[2] Following common usage, the term 'factor' is used here as a synonym for 'component'.

[3] Note that the 'most frequent' question stage is not presented since it proved to have little discriminating value in the initial analysis.

CHAPTER 6

DISCUSSION AND IMPLICATIONS

In this chapter, the findings which were presented in Chapter 5 are discussed in relation to theory and research reviewed in the first three chapters. We begin with a brief summary of the results for each research question. Following the summary, the results from each type of analysis are discussed in turn. In the final part of the chapter, the implications of this research for SLA theory and for teaching in intensive ESL classrooms are discussed.

Summary of the Findings

The first research question concerned whether grammatical sensitivity would be related to performance on measures of L2 knowledge and skill when L2 learning takes place in a communication-oriented instructional setting. The results of the principal components analysis revealed that three independent factors accounted for just over 70% of the variance in scores: 'Post-test English Proficiency', 'Pre-test English Proficiency', and 'Grammatical Sensitivity'. Trouve les Erreurs did not load on the 'Post-test English Proficiency' factor. Furthermore, the L2 metalinguistic task (the Grammaticality Judgment task) did not load on the 'Grammatical Sensitivity' factor. The hypothesis that grammatical sensitivity would be related to performance on form-focussed L2 measures of proficiency was thus not supported.

The second research question concerned whether successful and less successful L2 learners are distinguished by grammatical sensitivity. It was hypothesized that there would be a relationship when success is measured by form-focussed tasks only. Learners were classified into four performance profiles on the basis of a cluster analysis. Cluster 1 contained learners who outperformed the other clusters on all the tests, including the L1 grammatical sensitivity measure. Cluster 3 learners were average on the listening and

159

vocabulary tests but relatively weaker on the grammatical sensitivity measure and on some L2 tests requiring a focus on form. The other two clusters contained the learners who generally performed below the mean on all measures. Cluster 2 differed from Cluster 4, however, by a relatively stronger performance on Trouve les Erreurs. Thus, the level of grammatical sensitivity was associated with degree of success on form-focussed L2 tasks for Clusters 1, 3 and 4 but not for Cluster 2. The hypothesis was thus only partially supported.

The third research question probed the relationship between grammatical sensitivity and grammatical development in oral production. This was tested using qualitative linguistic analyses that focussed on developmental stages for question forms and for third person possessive determiners. Data were collected from a randomly selected sample of 10 learners from each of the five ESL classes. In the initial analysis, learners were assigned to High, Mid and Low GS groups based on their score on Trouve les Erreurs. On the question task, the High GS and the Low GS groups did not differ in the distribution of highest stage or most frequent stage analysis. The analysis of the possessive determiners, however, did suggest a pattern such that High GS learners were more likely to attain higher stages in the production of his/her, and were less likely to be in the earliest stages. A secondary analysis was performed using cluster membership as the grouping variable. The best comparison was between the subsamples from Clusters 1 and 2 which were of comparable size. These groups were statistically different from each other in their performance on the Trouve les Erreurs task. There was a larger percentage of Cluster 1 learners in Stages 4 and 5 for questions and fewer in Stage 2 than was the case for Cluster 2. The same trend was found for the possessive determiners. Altogether, these findings support the hypothesis that higher levels of grammatical sensitivity are associated with higher stages of grammatical development in oral production.

160

Discussion of the Findings

Discussion of the Principal Components Analysis

In this study, the principal components analysis revealed that three independent factors accounted for the performance of intensive ESL students on L2 paper-and-pencil tests. In the last chapter, these factors were labelled by examining the factor loadings. Discussion in this section will focus on explaining these factors by relating them to theoretical accounts of L2 aptitude and metalinguistic skill.

Factor 1: 'Post-test English Proficiency'

Factor 1 was described as representing L2 proficiency gained after five months of intensive ESL instruction since it loaded significantly on all the immediate and delayed post-test measures. This label has the disadvantage that it implies that L2 proficiency is unitary and is available in the performance of all kinds of tasks. We know, however, from the work of Bialystok and Ryan (1985a, 1985b), Cummins (e.g., 1983), and Snow (1991), that language knowledge reveals itself in different ways depending on task demands. A closer examination of the tasks used to measure L2 proficiency in this study helps to explain the nature of this factor.

Three of the post-tests were classified as being form-focussed: the Grammaticality Judgment task, the Cloze and the MEQ Listening/ Reading Comprehension test. On the other hand, the other post-test, the Yes/No Vocabulary test, was not considered to be form-focussed since it tested vocabulary knowledge without requiring a focus on grammatical form. Surprisingly, the Yes/No Vocabulary test loaded together with the form-focussed measures but not with the pre-test vocabulary tests. This suggests that the abilities tapped by the Yes/No test are more metalinguistic than initially supposed. Indeed, the task requires learners to indicate the words which 'they know that they know'.[1]

The question then arises as to why there was no relationship between the measure of the ability to focus on grammatical form in French, and the L2 tasks which required a

focus on form. One explanation is offered by Bialystok's distinction between Analysis and Control. It can be argued that all of the post-tests involve high levels of control of processing. Tasks that are high on Control make demands on the ability to co-ordinate selective attention. This is clearest in the case of the Grammaticality Judgment task and the multiple-choice format of the MEQ test. For the former, learners must shift attention between form and meaning as they read the text and judge whether a given sentence contains an error or not, while the latter requires the learner to manipulate attention from meaning to form, from listening to reading, from one item to the next unrelated item, from the distractors in the multiple-choice to the correct response. With respect to the Yes/No Vocabulary test, we must speculate since it is not known how young, low proficiency L2 learners approach this task. When confronted with a list of words divorced from any context, they must mentally focus on each word one at a time and try to make a connection to their own highly contextualized knowledge of the L2. Since many of the real words are unfamiliar (as are, of course, all of the non-words), they are likely to use some kind of inferencing strategy. The nature of this task suggests that it requires flexibility in attention allocation. Lastly, the Cloze test also makes demands on Control because of the need to co-ordinate meaning and syntactic structure (Bialystok & Ryan, 1985a; Birdsong, 1989, p. 59).

In contrast, Trouve les Erreurs is a task that requires the learner to focus on the written form of disconnected sentences. Comprehending the sentences poses no problem for these native speakers. The difficulty is in performing the necessary grammatical analysis in order to identify and correct errors in grammatical spelling. The task does not require the learners to constantly shift attention between form and meaning as would be the case if they had to make corrections to a coherent text. According to Bialystok's framework, this task makes higher demands on Analysis than on Control; therefore, performance on this task will not necessarily correlate with tasks that require high Control. Bialystok's Analysis/Control framework, thus, accounts for the relationship between all the

L2 post-test measures and the lack of a relationship between these measures and the L1 error correction task.

Working Memory Capacity

Another approach to interpreting Factor 1 is to consider it from the perspective of working memory capacity. Working memory represents the memory processes involved in the simultaneous storage and processing of information in real-time (See Baddeley & Hitch, 1974; Baddeley, 1986, 1990; for L2 applications, Harrington & Sawyer, 1992; McLaughlin, 1990a). There are three components of working memory: the central executive, which is concerned with the allocation of attention, and the phonological loop and the visual-spatial sketchpad which allow for the short-term storage and manipulation of acoustic and visual information, respectively. Working memory is limited in how much information it can hold and how many operations it can perform at one time; information held in this temporary cognitive workspace is subject to rapid decay or loss unless it is rehearsed via the phonological loop or visuo-spatial sketchpad. Given such limitations, the speed and efficiency of working memory will have considerable impact on the success or failure of cognitive operations such as L2 test performance (McLaughlin, 1990).

The test commonly used by cognitive psychologists to measure working memory capacity is the Reading Span test (Daneman & Carpenter, 1980). A subject is asked to read successive sentences in a set while simultaneously remembering the final word of each sentence. At the end of each set the subject is asked to recall all the final words of the sentences in the set. Each new set contains an increasingly larger number of sentences which thus steadily increases the processing load. The number of words correctly recalled becomes the index of the subject's working memory capacity. Parallels are evident here to the demands imposed by the tasks used to measure L2 proficiency in this study.

The MEQ Listening/Reading Comprehension test, for example, certainly taxes working memory. It requires the test taker to perform many operations: he or she must

163

attend carefully to the aural stimulus, read the written multiple-choice response items, and access specific linguistic knowledge from long-term memory to choose the correct answer. This places a load on working memory even for the competent language user. The co-ordination of these operations is much more difficult for the less proficient L2 learner who does not have automatized access to linguistic knowledge. Moreover, if the relevant linguistic representations are lacking or are not in an easily accessible form, the learner has to make use of contextual information or use a strategy to arrive at an answer, and this further burdens attentional resources. The same would be true of the other form-focussed tasks such as the Cloze and the Grammaticality Judgment tasks which require the learner to focus on both form and meaning. This prediction is consistent with the findings of Just and Carpenter (1992) that differences in performance of individuals with high vs.. low working memory capacity emerged when task demands were complex but not when they were simple (Segalowitz, 1997).

The advantage of the working memory explanation is that it is linked to the information-processing model of L2 acquisition proposed by Skehan (1998). Figure 6.1 provides a more detailed schematization of the model than the one presented in Chapter 1 (Figure 1.4). In this model, working memory plays a key role. On the one hand, it is the arena where on-going language comprehension and production occur; on the other, working memory is where conscious metaprocesses such as pattern identification and organization operate on material stored in long-term memory. It has a central function in all three stages of the process of acquisition, from input processing to central processing to output. It follows from the model that a greater working memory capacity will lead to faster L2 learning. (More will be said about the model later.)

high analysis

Analytic
learners [B]

Balanced
learners [A]

low memory

high memory

Memory
oriented
learners [C]

low analysis
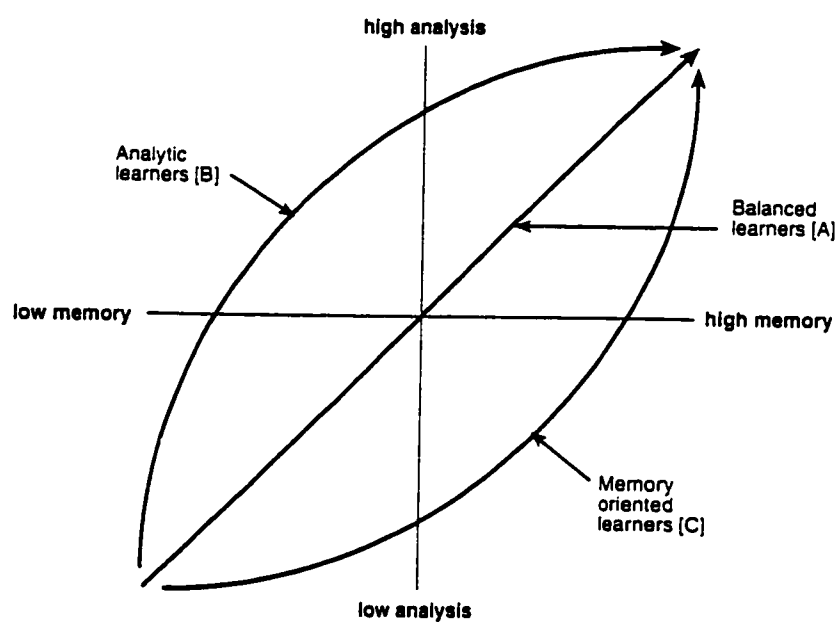
Figure 6.1 Skehan's information-processing model of SLA (Skehan, 1998, p. 57)

Factor 2: 'Pre-test English Proficiency'

The second factor loaded highest on the two September listening tests. These tests require the learner to match isolated words or simple sentences to pictures. They have to recognize the phonetic stimulus forms and then match them against lexical representations in long-term memory before deciding on an appropriate response (i.e., identify the appropriate picture). Thus, fairly low level processing skills are involved, namely, vocabulary recognition and lexical access. The Cloze involves word recognition and lexical access of the written rather than aural form of words. As it involves other skills as well, it is not surprising that its loading is not as high as those of the two listening tests. A low but significant loading is found for the June but not the January MEQ listening/reading comprehension test. The difference may depend on the fact that the additional items on the MEQ-53 version used in June included a subtest which requires the learner to listen to an extended description of the position of objects with reference to a map and then respond to questions concerning the location of these objects.

Since the highest loadings for Factor 2 were the pretest measures, this factor appears to reflect the amount of L2 knowledge, specifically, vocabulary knowledge, that learners had acquired in the limited amount of ESL instruction they had received in the two years prior to their intensive year in grade 6 (i.e., no more than two hours per week). Carroll (1962) proposed that aptitude will be most predictive when little time is available for learning. However, it was not predicted that grammatical sensitivity would be related to performance on these meaning-focussed tasks. How does this factor reflect differences in learner aptitude?

The nature of these vocabulary measures is closely related to the phonetic coding ability component of aptitude which is the "ability to identify distinct sounds, to form associations between those sounds and symbols representing them, and to retain these associations" (Carroll, 1981, p. 105). This ability includes a special type of memory for phonetic material. Carroll proposes that this ability is involved in segmenting speech into

166

words, syllables and phonemes, and in associating such segments with graphemic counterparts. Skehan (1998) describes phonemic coding ability[2] as being important for processing auditory input in real time so that it can be passed on to subsequent stages of processing. It is especially important at the earliest stages of L2 learning in informal learning settings where

> unstructured input, which may vary unhelpfully in acoustic quality, segmentability, and the acoustic salience of its components, will put an even greater premium on input-to-intake conversion. (Skehan, 1998, p. 203)

Studies have shown that phonological short-term memory span, a concept which clearly overlaps with phonemic coding ability, predicts long term acquisition of vocabulary in L1 and L2 (see N. Ellis, 1994 for review). It can be argued that learners in this study who had greater phonemic coding ability were able to learn more vocabulary from the minimal amount of exposure to English that they received in grades 4 and 5. Such an early advantage in vocabulary knowledge is likely to lead to long-term advantages. In order for L2 acquisition to proceed, however, access to vocabulary items in long-term memory has to become automatized through exposure and practice, which makes it possible for limited attentional resources to be allocated to higher levels of processing (Segalowitz, 1997).

Factor 3: 'Grammatical sensitivity'

Factor 3 mainly accounts for the unique variance associated with the L1 Trouve les Erreurs measure. It also loads significantly, but moderately, on the September administration of the Cloze. In both cases, the learner has to engage in some kind of parsing and focus on the appropriate written form of words. It was expected that the two tasks that involved error detection, Trouve les Erreurs and Grammaticality Judgment, would pattern similarly. However, the loading for the Grammaticality Judgment task on Factor 3 was below the cutoff for practical significance (i.e., .4 and above). As we saw in the discussion of Factor 1, this separation is consistent with the view of Bialystok & Ryan

(1985a, 1985b) that metalinguistic ability is not unitary; performance on one type of metalinguistic task does not necessarily predict performance on another.

The limited significance of the 'Grammatical Sensitivity' factor is in contrast to the large body of research on aptitude that has consistently found a relationship between grammatical sensitivity and measures of L2 proficiency. This discrepancy may be due to the way grammatical sensitivity was operationalized in this study -- Trouve les Erreurs and Words in Sentences probably are measuring different things. Here, again, Bialystok's Analysis/Control framework is useful. It was argued earlier that Trouve les Erreurs involves high demands on analysis but low demands on control. The MLAT subtest, on the other hand, makes high demands on both dimensions. It is a challenging task which requires the examinee to move from one pair of decontextualized sentences to another, to parse each sentence, to make mental comparisons between constituents, and choose among several possible responses. This involves control of cognitive processing. Moreover, Words in Sentences requires analyzed knowledge of English since knowledge of grammatical categories has been found to correlate with performance on this aptitude subtest (Carroll, 1979; 1990). Thus the MLAT measure of grammatical sensitivity can be characterized as high analysis, high control. Its predictive power in past research is probably due to the fact that L2 proficiency measures made similar processing demands. It is significant that in a recent factor analytic study by Alderson et al. (1997) Words in Sentences and other measures of metalinguistic knowledge did not load on the same factor as L2 proficiency measures. A similar pattern was found by Sawyer (1992). We can speculate that the reason for this lies with the use of L2 proficiency measures which are pragmatically contextualized and thus make less heavy demands on Analysis and Control.

It is probable that the limited scope of the Grammatical Sensitivity factor is a result of the restricted range of scores on the Trouve les Erreurs measure. Restriction in range attenuates correlations which is the computation underlying principal components analysis.

168

A more discriminating measure might reveal a stronger relationship between grammatical sensitivity and L2 performance on a variety of tasks.

According to Bialystok's framework, performance on Trouve les Erreurs should cluster with L2 tasks that have similar task demands. The only L2 task that required any kind of written response was the Cloze. The clustering of the Cloze with Trouve les Erreurs is consistent with Oller's view of the cloze procedure as reflecting the learner's underlying competence or "grammar of expectancy" (Oller, 1973). The type of grammatical analysis required by the Cloze in this case was quite implicit; the learner had to choose a semantically appropriate lexical item of the appropriate word-class to fit into a paragraph containing a few simple sentences. It is difficult to explain from a processing perspective why the pre-test Cloze loaded on this factor but not the post-test Cloze, and why the post-test Cloze loaded on Factor 1 but not the pre-test Cloze. It must be noted, however, that the January Cloze had a very low reliability coefficient (see Appendix H); it is thus very likely that measurement error has obscured the picture, making interpretation impossible.

Grammatical Sensitivity and the Information-processing Model

How does grammatical sensitivity fit into the information-processing model? We would predict that greater efficiency in processing, especially on demanding tasks, will be possible when linguistic representations are better organized and more accessible. Skehan (1998) proposes that conscious 'metaprocesses' operating in working memory are what cause changes in the representations of the L2 system in long-term memory, leading to rule-based rather than exemplar-based representations. These metaprocesses include matching one's interlanguage forms to corresponding target language forms, appreciating instruction and feedback, and reorganizing L2 material. As we saw in Chapter 2, Part II, Skehan argues that learners with an analytic learning style have a predisposition towards engaging in such metaprocesses. It seems possible to equate grammatical sensitivity with

169

this predisposition to analysis. We would therefore expect to see a relationship between the L1 and the L2 on tasks requiring analyzed knowledge. However, it follows from the processing model that the operation of these metaprocesses presupposes that vocabulary access is sufficiently automatized so that attentional resources required for the metaprocesses are available.

Summary

All three of the factors that emerged from the principal components analysis are interpretable in terms of concepts from the SLA research literature on aptitude, metalinguistic skill and psycholinguistic processing. While this study did not include specific measures of working memory, or of phonetic coding ability, the connection of the test results to these constructs flows from the pattern of factor loadings and the analysis of task demands. The relationship between these different abilities and learning outcomes becomes much clearer from the perspective of the learner profiles which emerged in the cluster analysis.

Discussion of the Cluster Analysis

In this study, four learner profiles emerged from the cluster analysis and were characterized as follows:

Cluster 1
- strong on all L2 pre-tests and post-tests;
- strong on L1 grammatical sensitivity measure

Cluster 2
- weak on all L2 pre-tests and post-tests;
- average on L1 grammatical sensitivity measure

Cluster 3
- average on pre- and post-test measures of L2 vocabulary and listening;
- below average on Grammaticality Judgment, Cloze, and L1 grammatical sensitivity measure

Cluster 4
- weak on all L2 pre-tests and post-tests;
- weak on L1 grammatical sensitivity measure

These profiles can be interpreted by relating them to the discussion of the factors that emerged from the principal components analysis and thus to the concepts of working memory, phonemic coding ability and the analysis of knowledge, which are part of Skehan's information-processing model.

Cluster 1 is characterized by uniformly high levels on all the variables associated with each of the three factors. Their performance reveals above average vocabulary recognition and lexical access at the beginning of instruction, the ability to perform form-focussed tasks, and high levels of grammatical sensitivity. According to the information-processing model, we would infer that their superior test performance is due to strong phonemic coding abilities, adequate working memory capacity, and relatively more analyzed knowledge. Strong phonemic coding ability allows vocabulary to be well encoded into long-term memory while sufficient working memory capacity makes it possible for these learners to engage in analysis of the linguistic material in long-term memory. This in turn leads to greater differentiation and organization of knowledge representations.

171

The Cluster 4 group, in contrast, is weak on all of the three factors. They have great difficulty in accessing vocabulary on the pre-tests, are unable to handle form-focussed tasks, and display low levels of grammatical sensitivity. Their profile reveals that their performance is particularly poor on the two L2 measures administered in June when they had just completed their French academic portion of the year. We would attribute the poor pre-test performance of Cluster 4 learners to difficulties with phonemic coding. Difficulties at this initial stage of processing make it very difficult to retain vocabulary, and then to automatize lexical access or to proceed to further stages of grammatical analysis. The fact that they were more disadvantaged on the delayed post-tests than on the immediate post-tests suggests that they found it very difficult to focus on both English and French at the same time. This may be a reflection of low working memory capacity.

Cluster 2 learners also perform poorly on almost all of the L2 tasks, demonstrating difficulties with vocabulary and form-focussed tasks. However, their average performance on Trouve les Erreurs but weak performance on the Cloze does not fit the pattern of variable loadings for Factor 3. One explanation is that these students were not able to cope with the analytic demands of the Cloze because of their slow L2 vocabulary recognition. Their relatively stronger performance on the French error correction task shows that Cluster 2 learners are capable of dealing with metalinguistic activities given sufficient time and instruction. This explanation thus suggests that the particular weakness of Cluster 2 learners is their phonemic coding ability. As the model predicts, later stages of processing will be affected by inefficient encoding of L2 vocabulary. In comparison to the Cluster 4 group, Cluster 2 learners are better able to handle the post-tests, which may indicate that they have greater working memory capacity.

The profile of Cluster 3 is not as easily accounted for by the three factors. The performance of these learners is average on the two vocabulary pre-tests, and on two of the post-tests (i.e., Yes/No and the MEQ), but the grammatical sensitivity measure, and the Grammaticality Judgment task and the Cloze tests are much more difficult for them. The

information-processing model allows us to define the problem more specifically. Based on their pre-test and post-test performance, Cluster 3 learners appear to have adequate phonemic coding abilities, and working memory capacity. Their difficulties with some of the form-focussed tasks suggests that they have difficulty accessing syntactic knowledge because their representations of the L2 are exemplar-based rather than analyzed. Skehan proposes that the metaprocesses which are responsible for the development of L2 knowledge are conscious. It may be that learners choose not to engage in such processes. Some learners prioritize accuracy, but others favour fluency first. Cluster 3 learners appear to have the latter learner style. It is interesting that Clusters 2 and 3 are not significantly different from each other on the Grammaticality Judgment task. We can speculate that, based on their profiles, the two groups have difficulty with this task for different reasons: reading the extended text is probably difficult for Cluster 2 learners because of their slower processing abilities, while Cluster 3 learners' are likely to have problems focussing on the grammatical form of the sentences.

Validating the Clusters

The above discussion demonstrates how the clusters can be interpreted with respect to Skehan's theoretical model of acquisition. It is important, however, to establish the validity of the cluster solution chosen in this study. Cluster analysis is a technique which describes the sample it is performed upon; it is not an inferential statistic which permits generalizability to a population (Everitt, 1977). Furthermore, the nondeterminacy of the results (i.e., no optimal solution can be determined a priori), means that conclusions should remain tentative (Skehan, 1986a). Everitt (1977, p. 140) suggests that merely naming the clusters is "gratuitous unless the clusters can be shown to be stable and to have some useful purpose outside themselves. One way of demonstrating such properties would be to show that the clusters had predictive value with respect to variables other than those used in

defining them". Two types of data are available to do this: the oral data analysis and the learner questionnaire responses.

Based on the cluster profiles, we would predict that the learners in Cluster 1 would be at a more advanced stage in their production of question forms and of his/her. The findings presented in Chapter 5 with reference to Research Question #3 confirm this prediction (see Table 5.9). Unfortunately, there were only a small number of students for whom oral data were available to represent the two other clusters, so conclusions must be considered tentative. Cluster 3 learners are more likely to be at higher stages of interlanguage development than Clusters 2 and 4, and Cluster 2 is somewhat more advanced than Cluster 4. This pattern is consistent with the cluster profiles, which are based on the paper-and-pencil test scores.

The Learning Preferences Questionnaires, on the other hand, provide further support relevant to all four cluster profiles. Since learners made relatively little use of the low end of the scale, the focus here is on the percentage of students who chose the highest point on the scale (viz, 'beaucoup') indicating that they thought the activity helped them a lot in learning English. These percentages are presented in Table 6.1. We would predict that the cluster differences would be apparent on the activities that make demands on working memory, phonemic coding ability, and grammatical sensitivity. The items 'listening to cassettes' and 'learning and reciting tongue twisters' provide some information on learners' attitudes towards tasks which make demands on phonemic coding ability and working memory. It was noted in Chapter 5 that there were doubts about the reliability of learners' responses to the metalinguistic items 'consulting a grammar book' and 'doing grammar exercises'. This may reflect the fact that these activities were not a part of their ESL classes. The items relating to literacy activities (i.e.,'reading stories' and 'writing texts') do provide some information on learners' attitudes towards form-focussed activities.

Table 6.1

Percentage Of Students Choosing 'Beaucoup' on Learner Preferences Questionnaire Items

	listen to cassettes	recite tongue twisters	write texts	read stories
Cluster 1 (n=53)	43	21	43	72
Cluster 2 (n=41)	34	10	37	58
Cluster 3 (n=28)	43	18	32	46
Cluster 4 (n=13)	46	15	23	39

As expected, Cluster 1 learners are enthusiastic about all types of activities. Cluster 3 learners are more positive than those in Cluster 2 about listening to cassettes and reciting tongue twisters. Conversely, Cluster 3 learners appear to like reading and writing less than the more grammatically sensitive Cluster 2. These differences are small but the trend is consistent with the cluster profiles. Only Cluster 4 is not completely consistent with predictions. We would expect them to pattern more like Cluster 2 on the listening and the tongue twisters items which they do not. They do however, tend to disfavour the two literacy activities, as would be expected. The pattern of responses of Cluster 4 may reflect the smaller cell size of this group relative to the others. It may also be the case that weaker students lack the metacognitive abilities to give reliable self-reports concerning their learning[3]. (See Appendix L for cluster group responses to all items on the questionnaire.)

The oral production data and self-report responses from these learners provide some validation of the learner profiles derived from this population. However, sufficient

175

oral data are available for only two of the clusters, and the differences in questionnaire responses are small. Another form of validation comes from investigations of successful and less successful L2 learners in other instructional settings. Many of these studies are with school-aged learners who typically display a wider range of abilities than adults do. This is because as students get older, they are generally able to make choices about what they study; less successful language learners will choose not to continue with L2 courses. Weak adult L2 learners who persist with language studies later in life usually do so for instrumental motives (e.g., foreign language requirement for a university degree, job opportunities, promotion, etc.).

The Characteristics of Strong vs. Weak Language Learners
Problems with phonemic coding ability.

A number of researchers have pointed to the importance of auditory abilities as a predictor of success in L2 learning in the classroom. A pioneering study of what they called underachievement was conducted by Pimsleur, Sundland, and Comfrey (1964). They concluded that auditory ability was the main variable associated with underachievement in L2 learning. Underachieving students in their first year of French or Spanish in different high schools were significantly lower on two tests of auditory ability. Furthermore, those who reported themselves to be visual learners ('eye-minded') rather than auditory learners ('ear-minded') were more likely to choose to discontinue their L2 studies. Clinical interviews with underachieving adults studying foreign languages at university indicated that difficulty with listening comprehension was the main problem. In addition, these individuals were reported to have "subtle to conspicuous difficulties" with performing such tasks as spelling, or reading aloud in English.

The connection between L1 and L2 difficulties has been documented in other research. For example, Sparks, Ganschow, Javorsky, Pohlman and Patton (1992) studied

176

three groups of American high school students: those who were identified as being low risk for failing their first semester in a foreign language, those who were identified as being at high risk for failure, and those identified as learning disabled. Despite average or above average cognitive ability, high risk learners had problems performing tasks such as correcting spelling and punctuation errors and identifying isolated words in English. There were, however, no significant differences between the three learner groups on tasks involving semantic processing such as reading comprehension. The authors conclude that the difficulties of at-risk students reflect problems with phonological and syntactic processing. Sparks and Ganschow (1991) hypothesize an underlying 'linguistic coding deficit'. While such individuals may be able to compensate for their deficit by using strategies, this is not as effective when faced with learning an unfamiliar linguistic system. According to Sparks and Ganschow, the linguistic coding deficit accounts for why students who are intelligent and motivated may yet experience difficulties learning a L2.

The relevance of phonemic coding ability for learning has also been demonstrated in learners who have access to much more L2 input than students in foreign language classes. Humes-Bartlo (1989) investigated "fast" and "slow" learners who were Spanish-speaking children in grade 5 bilingual classrooms in New York. The fast learners were those who passed an English proficiency test in three years or less, while slow learners were those who failed the test despite having three or more years of schooling in English in the US. The fast learners were distinguished by their superior verbal analogical reasoning, auditory discrimination abilities, and their memory abilities.

Humes-Bartlo comments that the language processing system of the slow learners in her study is adequate for L1 but is overloaded by having to deal with two languages. This description would also fit the learners in Cluster 4 in this study who, unlike their classmates, appeared not to be able to concentrate on both English and French language skills during the February-June academic portion of the intensive program. It is also

177

reminiscent of other learners described in the literature. For example, Juan, a slow learner in Wong Fillmore's study was characterized as having

> a one-step-at-a-time approach to language learning: That is, he tended to work on one kind of construction at a time, and seldom used anything until he was quite certain of its use. (Wong Fillmore, 1979, p. 227).

Similarly, in the adult study by Pimsleur et al. (1964, p. 150):

> instead of simply listening, they anxiously struggled to translate mentally into English as they went, thereby losing large segments of the aural material. One said "I don't like to go on until I know what this word is.

Another example is provided by Wesche (1981) who describes adult students in the Canadian public service language program who scored poorly on the Phonetic Coding subtests of the MLAT as follows:

> Classroom manifestations of difficulties in this area observed in our program include the jumping of syllables or words, and slowness in putting sentences together when speaking the target language; momentary forgetting of the meanings of familiar French words, and production of 'one-by one' word strings without 'integrating' them into phrases. (p. 131)

Finally, Mangubhai (1991) cites a think-aloud protocol from Eric, a slow learner in his study of adults learning Hindi through Total Physical Response approach:

> What I'm doing is trying to remember the sounds and then I'm just rehearsing the sounds to recall, then recalling it, and translating it and recalling it in English and then remembering that and then going on with the rest and then putting them together ... so I guess it's like a straight translation process. (p.584)

We might attribute the difficulties faced by these learners as being due to limited working memory capacity. However, even with adequate working memory, L2 learning is seriously hindered when learners have a problem with phonemic coding. If learners have

difficulty in encoding lexical items, and storing them in long-term memory, then they will be unable to automatize access to vocabulary items. When vocabulary recognition is inadequate, learners use contextual cues for comprehension (Hulstijn, in press; Stanovich, 1980) or strategies such as on-line translation for production. Such activities monopolize attentional resources available in working memory, and thus make progress to syntactic processing of the L2 impossible or extremely difficult. [4]

Problems with decontextualized language skills.

Researchers have also identified decontextualized language ability as a distinguishing characteristic of strong and weak L2 learners. For example, on the Verbal Analogies test used by Humes-Bartlo (1989), test takers must respond to items like 'Mother is to father as sister is to ___ '. This test clearly makes demands on both working memory and linguistic knowledge. Not surprisingly, it was the strongest predictor in the discriminant analysis of fast vs. slow learners.

The measure of decontextualized language skills used by Snow (1991) and her colleagues was more of an analytic task. Subjects were asked to give definitions of common nouns; responses were coded according to whether they provided a formal dictionary-type definition or used informal descriptive or narrative strategies. The learners were primary school children (grades 2-5) at the United Nations International School in New York. Low-English proficient students who make up about half of the school population are first placed in an ESL class and then in a transition program. Once these students are in the regular program, they receive instruction in French as a L3 which is offered at three different levels at each grade. Snow describes three learner profiles: fast learners of both English and French, slow learners in both English and French, and a third group of children who learned English within a year or two, but had difficulty learning French, remaining in introductory French classes for several successive years. Snow

argues that this third group lacked the decontextualization skills as demonstrated by their difficulty in giving formal definitions. Presumably instruction in the French program at this school is oriented towards linguistic form rather than communicative competence. Tasks that call upon analyzed knowledge put this group of students who were able to acquire English through informal contact at a disadvantage in the French class. This description is similar to that of Cluster 3 learners who were disadvantaged on form-focussed tasks relative to their performance on the other L2 measures.

Empirical evidence from different research contexts supports the interpretation of the clusters found in this study. We can also relate the clusters to the learner style framework proposed by Skehan (1998). As was discussed in Chapter 2, Part II, learner styles are viewed as varying on two independent axes: degree of analysis and amount of memory (see Figure 6.2.) The clusters from this study can be 'plotted' on the two dimensions, assuming that the Vocabulary Recognition task reflects learners' memory abilities: Cluster 1 learners are balanced learners who are high in analysis and high in memory; Cluster 3 learners are average in memory, but low in analysis; Cluster 2 is low in memory, but average in analysis; and Cluster 4 is low in memory and low in analysis. It is interesting that the last 'learner style' is logically possible from the framework, but Skehan does not mention it. This framework gives us a static picture of the four clusters. The previous discussion of the role of phonemic coding, working memory and analysis in L2 learning allows us to account for the underlying processing problem of the low memory learners.
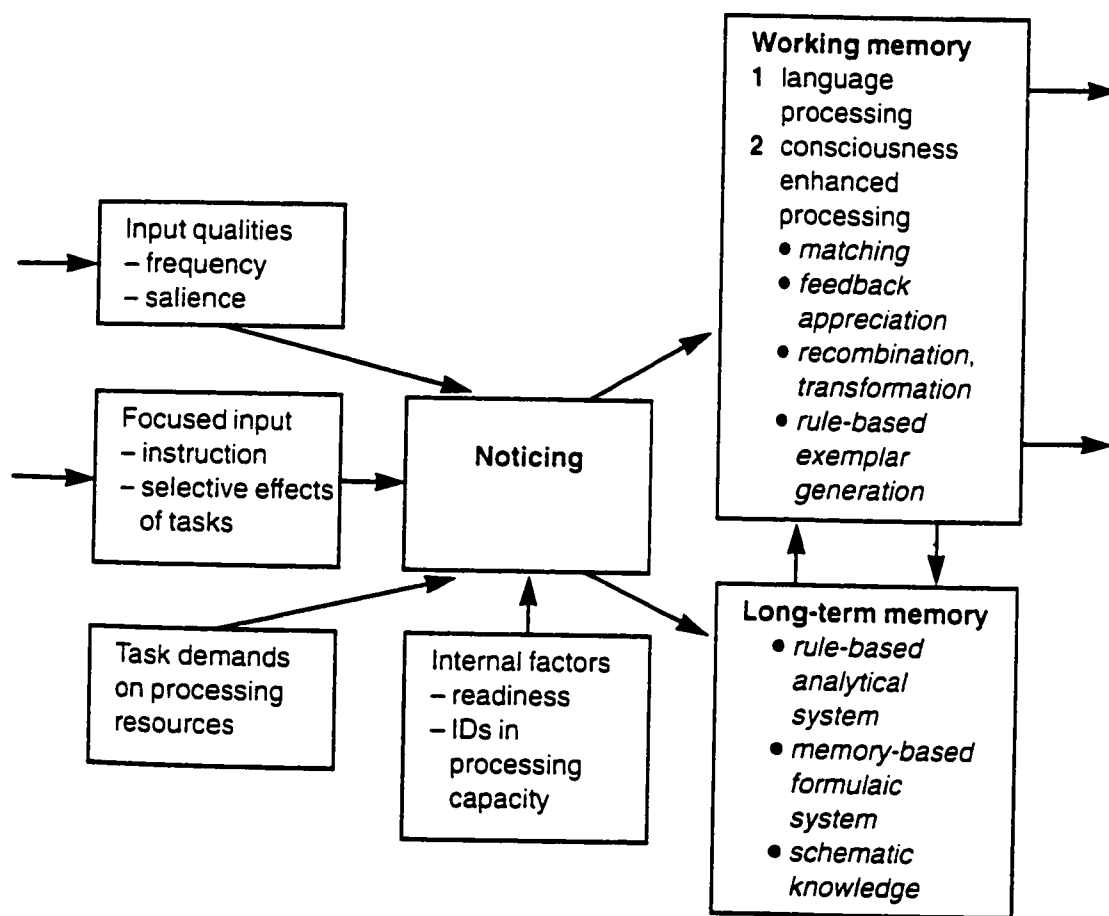
Figure 6.2 Learner styles and paths of L2 development (Skehan, 1998, p. 270)

Discussion of the Interlanguage Analysis

The analysis of the test data provides information concerning the ability of these beginner L2 learners to perform different kinds of language tasks. Many factors other than knowledge of the L2 (e.g., strategies, test-wiseness, environmental conditions, etc.) may contribute to an individual's score on any given test. Oral production data, on the other hand, allow us to see in a more direct way, what the learner knows about the L2 (see Gass & Selinker, 1994 for a discussion of production vs. judgment data). In this study, the oral production tasks targeted specific aspects of the grammatical system of the L2 known to have a developmental profile. This contrasts with the measurement approaches used in other studies reviewed in Chapter 3.

It was hypothesized that learners with greater grammatical sensitivity would be better able to notice structures in the input and to integrate them into their interlanguage. The cluster analysis revealed that the learners with high scores on the measure of grammatical sensitivity (i.e., Cluster 1) were strong on all the measures. They can thus be considered to be high aptitude learners. The question and the possessive determiner stage analyses confirmed the test results: the high aptitude learners were also faster in their development of the ability to use certain grammatical structures in oral production. This finding suggests a quantitative effect: the high aptitude learners learn more quickly and therefore know more English than learners with lower aptitude. However, it can be argued that the high aptitude learners are also qualitatively different because their knowledge of the L2 is represented in a more analyzed form.

Bialystok (1994b) posits that the development of language proficiency involves the process of analysis whereby knowledge representations "become more explicit, more structured, and more accessible to introspection" (p. 561). Skehan (1998) suggests that some learners will develop proficiency in this way while others will choose to prioritize fluency. Analytic learners develop differentiated, organized and rule-based representations

of language, but memory-oriented learners develop communicative ability through highly accessible lexicalized exemplars.

Using a developmental stage framework facilitates the comparison of one group of learners to another. Unfortunately, much information is lost in the reduction of linguistic data into a single number. A qualitative examination of the question transcripts provides some further evidence that the representations of high aptitude learners is more analyzed. Consider for example, the yes/no questions produced by Learner A and Learner B:

Learner A (Low GS/Cluster 4)	Learner B (High GS/Cluster 1)
Is it a shoe under uh on the floor?	Do, uh, is there a garbage?
The dog is it uh all white?	Is it this card?
On the picture is it a garbage?	The sun, is uh, it's in uh, is it on the right
Is it a cat in the picture?	corner?
The cat is it at right or at left?	The little girl that is running, is she on the
The cat is uh on the bus?	sidewalk?
The sun is at left at the right?	The girl, is she on the grass?
Do you have a bank on your picture?	Do the dog have some dots?
	Do the dog is on the bus?
	Do the bicycle rider is holding his bike on
	the sidewalk?
	Do two woman are talking together on the
	corner of the page?

On a superficial level, Learner A is more accurate than Learner B (82% vs. 62% correctly formed questions, respectively). This is because of the invariant use of well-formed 'is it' by the weaker learner. But there are many features of Learner B's interlanguage which reveal a deeper knowledge of English. Learner B uses 'is it' in contrast to 'is she', and produces 'do' fronted questions in a systematic way with noun phrase subjects which are difficult for francophone learners (Spada & Lightbown, 1999; White et al., 1991); there is an absence of SVO (Stage 2) forms; a variety of lexical verbs are used; and greater morphological and syntactic complexity is shown by the use of the progressive and a relative clause.

In many ways, the possessive determiner stages are more discriminating than the question stages. It is very difficult to determine whether a learner has the ability to correctly invert the subject and auxiliary in questions or whether their well-formed question is an unanalyzed chunk. The question elicitation task lends itself to the use of formulaic questions such as "Do you have X in your picture?" or "Where is X?". When learners make use of formulae, the stage score overestimates their knowledge of the complex rules of question formation in English. For example, an examination of Table 5.9 reveals that the distribution of learners in the higher question stages (3 to 5) is quite similar for Clusters 2, 3 and 4. For any particular learner, it is impossible to know whether the analysis is overestimating or underestimating his or her knowledge of English question formation.

Possessive determiners, on the other hand, do not lend themselves to formulaic use, or rather it is usually clear when this is the case because it will lead to errors. The correct choice of determiner has to be 'computed' each time it is used. Furthermore, the stage descriptions themselves are more stringent because stage assignment takes into account, to some extent, both correct and incorrect uses of the target forms. Thus the fact that Cluster 1 learners are more likely to be in the Post-emergence stages (see Figure 5.5) indicates their greater knowledge of the agreement rule for his/her.

One of the characteristics of explicit knowledge is that it is available for introspection (Bialystok, 1994b). In this study, it is possible to examine the relationship between the use of his/her in oral production and metalinguistic judgments of errors involving these forms. Cluster 1 learners, who as a group tend to be at higher stages in their use of the possessive determiners, are also significantly better than the other three clusters on the possessive determiner items on the Grammaticality Judgment task (see Table 5.6). While the unbalanced sample size makes it difficult to compare the subsamples from Clusters 1 and 3, a comparison between a pair of individual learners is revealing. Learner B (from above) was classified as being at stage 6 in her use of his/her. On the his/her items on the Grammaticality Judgment task, she scored 18/21 and demonstrated by her responses

to the kin-different items that she knew the agreement rule. In contrast, Learner C from Cluster 3 was also classified as a stage 6 in oral production of <u>his</u> and <u>her</u> but scored 5/21 on the Grammaticality Judgment task. Typical of the overall profile for Cluster 3, Learner C had a low score on Trouve les Erreurs. These observations provide support for the claim that high aptitude learners, characterized by relatively higher levels of grammatical sensitivity, have a more analyzed and accessible representation of the agreement rule for the possessive determiners in English than the lower aptitude learners.

The findings of the interlanguage analysis suggest that the high aptitude learners are more efficient in noticing patterns in the input, and matching these patterns with their own output. They appear to 'notice the gap' between the target language and their interlanguage (Krashen, 1983; Schmidt & Frota, 1986). It must be remembered that the instructional input in these classrooms can be characterized as being meaning-focussed rather than form-focussed since there is no explicit grammar teaching, and almost no translation. Teachers teach vocabulary through extralinguistic cues, and through songs, tongue-twisters, and activities. It is highly unlikely that any of the teachers ever explained what the word <u>do</u> was or its role in forming questions. On the other hand, questions are a constantly occurring feature of teacher talk and functionally very important. Certain useful questions such as 'How do you say X in English? Can I go to the washroom?' were taught and learned as formulas. Very little in the instructional input was designed to increase the saliency of grammatical structures. Thus, the high aptitude learners appear to be able to create 'internal saliency', as Sharwood Smith (1991b) puts it, and this drives acquisition forward. Given the earlier discussion of the effect of processing capacity, we might speculate that internally created saliency is the result of processing capacity becoming available, through the automatization of lower level processes. In this study, the real puzzle is Cluster 3 -- they appear to have sufficient working memory capacity (based on their performance on some decontextualized tasks), but have difficulty with deriving structure from the input. One

plausible hypothesis is that they need externally created saliency in order to be able to analyze the L2.

Implications for SLA Theory

While the role of grammatical sensitivity could not be isolated as being a major predictor of L2 development in this study, the findings are consistent with other research that shows learner aptitude to be related to speed of learning. Furthermore, the interpretation of the learner profiles confirms the usefulness of viewing aptitude as componential rather than unitary. This study extends the empirical evidence concerning aptitude to a context not previously studied, that of older children studying in ESL classes where the focus was on interpersonal communication-skills rather than academic skills. In contrast to some other studies, aptitude was found to be associated with performance on both paper-and-pencil tests and oral production tasks.

From the perspective of SLA theory, this study demonstrates the viability of the L2 information-processing model proposed by Skehan (1998). This model was based on the premise that L2 learning in older learners proceeds like all other cognitive skills. The three stages of input processing, central processing and output are derived from theory and research in cognitive psychology and a relatively thin strand of SLA research. The learning profiles of the pre-adolescent children in a communication-oriented instructional environment can be explained with reference to this model. In particular, the model clarifies the importance of phonemic coding ability needed for vocabulary acquisition and the more limited role of grammatical sensitivity for these learners who are at the early stages of their learning of English.

It was argued in Chapter 1 that SLA theory needed to find ways of integrating individual differences into models of the L2 acquisition process. A first attempt of this sort can be found in Carroll (1981) where the four components of aptitude are related to the cognitive processing model current at that time. In an updated view of both aptitude and

186

information-processing, Skehan (1998, p. 203) makes the link between the three components of aptitude and the three stages of language processing. Thus, phonemic coding ability is related to noticing that takes place at the input processing stage, language analytic ability is related to the metaprocesses that take place at the central processing stage and memory is related to the output stage where knowledge in long-term memory is retrieved.

Skehan's schematization neatly links aptitude with the processes of acquisition. The paradox is that by viewing the aptitude factors from the perspective of information-processing, the whole notion of learner aptitude as some kind of entity disappears. It is not the case that some individuals have some 'thing' which sets them apart from the 'normal' population. Rather, some learners are able to do more with the same cognitive 'hardware' that everybody else has because they are more efficient in processing information. A better understanding of the nature of this efficiency will lead to ways of helping those who lack it.[5] To this end, the evidence from this study suggests that the concepts of working memory capacity and phonemic coding ability merit further research. Adopting an information-processing model, as Skehan has done, also leads to 'de-reification' with respect to learner styles. The analytic vs. memory dichotomy is replaced by a framework where individuals are viewed as having a tendency to emphasize one aspect of the complex process of L2 learning over another.

Assuming that the good language learner is one who has a larger working memory capacity and can thus direct more attention to relevant features of the input, there still remains the question of what 'pushes' some learners forward in their analysis of the L2. In Skehan's model, the central processing stage plays a crucial role in the development of grammatical competence in L2. Skehan follows Schmidt (1990, 1994a) in assuming that conscious processing of the input is necessary. However, those scholars who hold that form-focussed instruction plays a role only in the earlier, selective attention stage of acquisition (e.g., Gass, 1991) would dispute this claim. Is it possible that we must look to

187

motivational factors to account for continued grammatical development once processing capacity is sufficient? We know that aptitude and motivation have independent but interactive effects on L2 learning. Interestingly, motivation is now also being studied from a processing perspective by Dörnyei (1998). Further exploration of the interaction between processing capacity (i.e., aptitude) and motivation along the lines of Carroll's (1962) Model of School Learning offers a promising avenue for research.

Future Research

The learners in this study were at the early stages of their L2 acquisition. A follow up of these learners at a later stage in their L2 learning would shed light on the third stage of Skehan's model where memory is predicted to play a more important role. Longitudinal data would provide insights into the issue of whether aptitude-related differences in L2 proficiency persist in the long term and to what extent learners' motivation can compensate for limitations in processing capacity.

Another logical follow-up to this study is to examine the effects of instructional interventions which aim to help learners with phonological coding of the L2 such as the "multisensory structured language approach" to language teaching advocated by Sparks and Ganschow (Sparks, Ganschow, Kenneweg, & Miller, 1991).

Limitations of this Study

The main limitations of this study are as follows:

1. The findings relate to the early stages of learning, and thus deal with rate of acquisition rather than ultimate attainment. The factors that were found to be associated with performance on measures of L2 proficiency may not be relevant for older, and/or more advanced learners.

188

2. It is also very important to consider the limitations of the L2 measures which sampled from a fairly narrow range of language abilities. Notably absent are measures of oral communication skills. This is significant since the program is geared towards the communicative use of the L2. It is conceivable that a student might be judged successful by her teacher in the context of this program and yet fare poorly on tests administered in this study. Furthermore, this study did not include measures of the ability to read and write texts in English, and thus does not address the issue of how grammatical sensitivity relates to functional literacy skills in the L2.

3. While the error correction task used to measure grammatical sensitivity had high face validity, it was not a powerful discriminator of analytic ability among these learners, especially those in the middle range. It is thus possible that the limited role attributed to grammatical sensitivity may be due to measurement error rather than a reflection of the way L2 acquisition actually proceeds.

Implications for L2 Teaching in Intensive ESL

What are the implications of this study for L2 teaching in this particular context? It was found that when learners are given proficiency tests that are more form-focussed than those that they have experienced in their ESL classes, about a third of the students are able to cope on the basis of their own abilities. Since such activities are not part of the goals of the ESL program, it can be argued that this is not a matter of concern. Of course, the fact that these students may need to develop decontextualized or form-focussed L2 skills later in their academic careers would be an argument for including more form-focussed instruction in the ESL program. Furthermore, such instruction may be necessary for some learners to be able to utilize their contextualized knowledge of the L2 for the performance of

decontextualized tasks (see Segalowitz, 1997 for discussion of 'transfer-appropriate learning').

As we saw in Chapter 1, there is evidence that learners in communicative L2 settings benefit from instruction which draws their attention to the non-salient aspects of the L2 as is the case when the L1 and L2 systems share some features but not others (e.g., adverb placement and inversion in English and French). However, this study highlights the fact that many students have difficulty acquiring basic vocabulary. For learners without automatized access to lexical items, further interlanguage development is hampered since processing the meaning of utterances uses up too many attentional resources. Instruction in this context should include ways of helping learners with the process of coding and remembering vocabulary and formulae. Research on memory in cognitive psychology suggests that codability and retention of vocabulary is enhanced by increased frequency. Hulstijn (in press) recommends the use of classroom activities which recycle vocabulary items from earlier lessons. Similarly, Gatbonton and Segalowitz (1988) recommend that L2 teaching should provide opportunities for repetition and practice within a communicative context. Recycling language material is a concept that every teacher endorses, and yet there is little explicit recycling of vocabulary in many of the L2 materials used in Quebec (Marlise Horst, personal communication). Indeed, in a study of the frequency characteristics of teacher talk in seven intensive ESL classrooms, Lightbown, Meara and Halter (in press) found that very little repetition of new words took place.

It is perhaps fitting, in conclusion, to consider the words of Corder written some thirty years ago,

> We have been reminded of von Humboldt's statement that we cannot really teach language, we can only create conditions in which it will develop spontaneously in the mind in its own way. We shall never improve our ability to create such favourable conditions until we learn more about the way a learner learns and what his built-in syllabus is. (Corder, 1967, p. 170)

190

The field of SLA has certainly made strides over the years in documenting what learners learn and how they go about it. Studies of individual differences in learning, such as this one, reveal that many learners need help with the 'spontaneous' process of L2 acquisition. Let us hope that the research efforts of the next thirty years will provide the basis for creating favourable conditions for all learners.

NOTES

[1] The metalinguistic nature of the task would explain the high correlations between the Yes/No Vocabulary test and standardized proficiency tests such as the Michigan Test of English (Loring, 1995).

[2] Skehan always refers to this factor as 'phonemic coding ability' rather than using Carroll's 'phonetic coding ability'. No explanation is given for the change in terminology. I will use the term 'phonemic' unless I am referring to Carroll directly.

[3] Stone, Stone and Gueutal (1990) found that examinees with lower cognitive abilities produced questionnaires with lower reliability coefficients.

[4] It is interesting that Mangubhai discusses the behaviour of his subject, Eric, as if it were a conscious choice not to focus on grammatical structure.

[5] Gould (1996) writes very eloquently about the negative consequences of the reification of intelligence.

REFERENCES

Ackerman, P. L. (1989). Individual differences and skill acquisition. In P. L. Ackerman, R. J. Sternberg, & R. Glaser (Eds.), Learning and individual differences (pp. 165-217). NY: Freeman.

Ackerman, P., & Schneider, W. (1985). Individual differences in automatic and controlled information-processing. In R. Dillon (Ed.), Individual differences in cognition. Vol. 2 (pp. 35-66). NY: Academic Press.

Alderson, C., Clapham, C., & Steel, D. (1997). Metalinguistic knowledge, language aptitude and language proficiency. Language Teaching Research, 1, 93-121.

Anderson, J.R. (1983). The architecture of cognition. Cambridge, MA/London: Harvard University Press.

Anderson, J.R. (1990). Cognitive psychology and its implications (3rd ed.). NY: W. H. Freeman & Co.

Bachman, L. (1990). Fundamental considerations in language testing. Oxford: Oxford University Press.

Baddeley, A. D. (1986). Working memory. Oxford: Clarendon Press.

Baddeley, A. D. (1990). Human memory: Theory and practice. Hove: Lawrence Erlbaum.

Baddeley, A. D., & Hitch, G. (1974). Working memory. In G.A. Bower (Ed.), The psychology of learning and motivation.Vol. 8 (pp. 47-90). New York: Academic Press.

Bates, E., Bretherton, I., & Snyder, L. (1988). From first words to grammar: Individual differences and dissociable mechanisms. Cambridge: Cambridge University Press.

Bates, E., Dale, P., & Thal, D. (1995). Individual differences and their implications for theories of language development. In P. Fletcher & B. MacWhinney (Eds.), The handbook of child language (pp. 98-151). Oxford: Blackwell.

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

Berko, J. (1958). The child's learning of English morphology. Word, 14, 150-177.

Bialystok, E. (1978). A theoretical model of second language learning. Language Learning, 28, 69-83.

Bialystok, E. (1979). Explicit and implicit judgements of L2 grammaticality. Language Learning, 29, 81-103.

Bialystok, E. (1986). Factors in the growth of linguistic awareness. Child Development, 57, 498-510.

Bialystok, E. (1987). Influences of bilingualism on metalinguistic development. Second Language Research, 3, 154-166.

Bialystok, E. (1988). Aspects of linguistic awareness and reading comprehension. Applied Psycholinguistics, 9, 123-139.

Bialystok, E. (1990). The dangers of dichotomy: A reply to Hulstijn. Applied Linguistics, 11, 46-51.

Bialystok, E. (1991). Metalinguistic dimensions of bilingual language proficiency. In E. Bialystok (Ed.), Language processing in bilingual children (pp. 113-140). Cambridge: Cambridge University Press.

Bialystok, E. (1992). Attentional control in children's metalinguistic performance and measures of field independence. Developmental Psychology, 28, 654-664.

Bialystok, E. (1994a). Analysis and control in the development of second language proficiency. Studies in Second Language Acquisition, 16, 157-168.

Bialystok, E. (1994b). Representation and ways of knowing: Three issues in second language acquisition. In N. Ellis (Ed.), Implicit and explicit learning of languages (pp. 549-569). London: Academic Press.

Bialystok, E. & Fröhlich, M. (1978). Variables of classroom achievement in second language learning. Modern Language Journal, 62, 327-336.

Bialystok, E., & Hakuta, K. (1994). In other words: The science and psychology of second-language acquisition. NY: Basic Books.

Bialystok, E., & Ryan, E.B. (1985a). A metacognitive framework for the development of first and second language skills. In D. Forrest-Pressley, G. MacKinnon, & T. Waller (Eds.), Metacognition, cognition, and human performance Vol. 1 (pp. 207-252). Orlando, FA: Academic Press.

Bialystok, E., & Ryan, E.B. (1985b). Towards a definition of metalinguistic skill. Merrill-Palmer Quarterly, 31, 229-264.

Bialystok, E., & Sharwood Smith, M. (1985). Interlanguage is not a state of mind: An evaluation of the construct for second-language acquisition. Applied Linguistics, 6, 101-117.

Birdsong, D. (1989). Metalinguistic performance and interlinguistic competence. Berlin: Springer-Verlag.

Bloom, L., Hood, L., & Lightbown, P.M. (1974). Imitation in language development: If, when, and why? Cognitive Psychology, 6, 380-420.

Bloom, L., Lightbown, P.M., & Hood, L. (1975). Structure and variation in child language. Monographs for the Society for Research in Child Development, 40(2, Serial No. 160).

Brown, H.D. (1994) Teaching by principles: An interactive approach to language pedagogy. Englewood Cliffs, NJ: Prentice Hall Regents.

Brown, H.D. (1997). Review of the book The good language learner. Studies in Second Language Acquisition, 19, 510-11.

Brown, J.D. (1996). Testing in language programs. Upper Saddle River, NJ: Prentice Hall Regents.

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

Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. Applied Linguistics, 1, 1-47.

Carroll, J. B. (1962). The prediction of success in intensive foreign language training. In R. Glaser (Ed.), Training research and education (pp. 87-136). New York: John Wiley.

Carroll, J. B. (1973). Implications of aptitude test research and psycholinguistic theory for foreign language teaching. International Journal of Psycholinguistics, 2, 5-14.

Carroll, J. B. (1979). Psychometric approaches to the study of language abilities. In C. Fillmore, D. Kempler, & W. Wang (Eds.), Individual differences in language ability and language behavior (pp. 13-31). NY: Academic Press.

Carroll, J. B. (1981). Twenty-five years of research on foreign language aptitude. In K. Diller (Ed.), Universals in language learning aptitude (pp. 83-118). Rowley, MA: Newbury House.

Carroll, J. B. (1985). Second-language abilities. In R. Sternberg (Ed.), Human abilities (pp. 83-101). NY: W.H. Freeman & Co.

Carroll, J. B. (1990). Cognitive abilities in foreign language aptitude: Then and now. In T. Parry & C. Stansfield (Eds.), Language aptitude reconsidered (pp. 11-29). Englewood Cliffs, NJ: Prentice Hall Regents.

Carroll, J. B., & Sapon, S. (1958). The Modern Language Aptitude Test - Form A. NY: The Psychological Corporation.

Cattell, R. B. (1987). Intelligence: Its structure, growth and action. Amsterdam: North-Holland.

Chapelle, C. (1988). Field independence: A source of language test variance. Language Testing, 5, 62-82.

Chapelle, C. (1992). Disembedding 'Disembedded figures in the landscape ...': An appraisal of Griffiths and Sheen's 'Reappraisal of L2 research on field dependence/ independence'. Applied Linguistics, 13, 375-384.

Chapelle, C., & Green, P. (1992). Field independence/dependence in second-language acquisition research. Language Learning, 42, 47-83.

Chapelle, C. & Roberts, C. (1986). Ambiguity tolerance and field independence as predictors of proficiency in English as a second language. Language Learning, 36, 27-45.

Clahsen, H., Meisel, J., & Pienemann, M. (1983). Deutsch als Zweitsprache: Der Sprachwerb ausländischer Arbeiter. Tübingen: Gunter Narr.

194

Clark, E. (1978). Awareness of language: Some evidence from what children say and do. In A. Sinclair, R. Jarvella, & W. Levelt (Eds.), The child's conception of language (pp. 17- 43). Berlin: Springer-Verlag.

Clyne, M. (1985). Medium or object- Different contexts of (school-based) second language acquisition. In K. Hyltenstam & M. Pienemann (Eds.), Modelling and assessing second language acquisition (pp. 197-212). Clevedon, UK: Multilingual Matters.

Cook, V.J. (1988). Chomsky's Universal Grammar: An introduction. Oxford: Basil Blackwell.

Cook, V. (1991). Second language learning and language teaching. London: Edward Arnold.

Cook, V. (1996). Second language learning and language teaching (2cd ed.). London: Edward Arnold.

Corder, S.P. (1967). The significance of learners' errors. IRAL, 5, 161-170.

Cronbach, L. (1975). Beyond the two disciplines of scientific psychology. American Psychologist, 30, 116-127.

Cummins, J. (1980). The cross-lingual dimensions of language proficiency: Implications for bilingual education and the optimal age issue. TESOL Quarterly, 14, 175-187.

Cummins, J. (1983). Language proficiency and academic achievement. In J. Oller (Ed.), Current issues in language testing research (pp. 108-129). Rowley, MA: Newbury House.

Cummins, J. (1987). Bilingualism, language proficiency, and metalinguistic development. In P. Homel, M. Palij, & D. Aaronson (Eds.), Childhood bilingualism (pp. 57-73). Hillsdale, NJ: Lawrence Erlbaum.

Cummins, J. (1991). Interdependence of first-and second-language proficiency in bilingual children. In E. Bialystok (Ed.), Language processing in bilingual children (pp. 70-89). Cambridge: Cambridge University Press.

Cummins, J. & Mulcahy, R. (1978). Orientation to language in Ukranian-English bilingual children. Child Development, 49, 1239-1242.

Daneman, M., & Carpenter, P. (1980). Individual differences in working memory and reading. Journal of Verbal Learning and Verbal Behavior, 19, 450-46.

Davies, A. (1971). Language aptitude in the first year of the U.K. secondary school. RELC Journal, 2, 4-19.

DeKeyser, R. (1995). Learning second language grammar rules: An experiment with a miniature linguistic system. Studies in Second Language Acquisition, 17, 379-410.

de Graaff, R. (1997). The eXperanto experiment: Effects of explicit instruction on second language acquisition. Studies in Second Language Acquisition, 19, 249-275.

de Villiers, J., & de Villiers, P. (1978). Language acquisition. Cambridge, MA/London, UK: Harvard University Press.

195

Dörnyei, Z. (1998, March). What is L2 motivation? Paper presented at the annual meeting of the American Association of Applied Linguistics, Seattle, WA.

Dulay, H., Burt,M., & Krashen, S. (1982). Language two. NY: Oxford University Press.

Ehrman, M., & Oxford, R. (1995). Cognition plus: Correlates of language learning success. Modern Language Journal, 79, 67-89.

Eisenstein, M. (1980). Childhood bilingualism and adult language learning aptitude. International Review of Applied Psychology, 29, 159-174.

Ellis, N. (1993). Rules and instances in foreign language learning: Interactions of explicit and implicit knowledge. European Journal of Cognitive Psychology, 5, 289-318.

Ellis, N. (1994). Vocabulary acquisition: The implicit ins and outs of explicit cognitive mediation. In N. Ellis (Ed.), Implicit and explicit learning of languages (pp. 211-282). London: Academic Press.

Ellis, N., & Schmidt, R. (1997). Morphology and longer distance dependencies: Laboratory research illuminating the A in SLA. Studies in Second Language Acquisition, 19, 145-171.

Ellis, R. (1994). The study of second language acquisition. Oxford: Oxford University Press.

Emmorey, K., Bellugi, U., Friederici, A., & Horn, P. (1995). Effects of age of acquisition on grammatical sensitivity: Evidence from on-line and off-line tasks. Applied Psycholinguistics, 16, 1-23.

Everitt, B. (1977). Cluster analysis and miscellaneous techniques. In A. Maxwell (Ed.), Multivariate analysis in behavioural research. London: Chapman and Hall.

Felix, S. (1981). The effect of formal instruction on second language acquisition. Language Learning, 31, 87-112.

Felix, S., & Hahn, A. (1985). Natural processes in classroom second-language learning. Applied Linguistics, 6, 223-238.

Gagné, E. D. (1985). The cognitive psychology of school learning. Boston, MA: Little Brown.

Gardner, R. (1960). Motivational variables in second-language acquisition. Unpublished doctoral dissertation, McGill University, Montreal.

Gardner, R., & Lambert, W. (1972). Attitudes and motivation in second-language learning. Rowley, MA: Newbury House.

Gardner, R., & MacIntyre, P. (1992). A student's contribution to second language learning. Part I: Cognitive variables. Language Teaching, 25, 211-220.

Gardner, R., & MacIntyre, P. (1993). A student's contribution to second language learning. Part II: Affective variables. Language Teaching, 26, 1-11.

Gardner, R., Smythe, P., Clément, R., & Gliksman, L. (1976). Second-language learning: A social psychological perspective. Canadian Modern Language Review, 32, 198-213.

Gass, S. (1988). Integrating research areas: A framework for second language studies. Applied Linguistics, 9, 198-217.

Gass, S. (1991). Grammar instruction, selective attention and learning processes. In R. Phillipson, E. Kellerman, L. Selinker, M. Sharwood Smith, & M. Swain (Eds.), Foreign/second language pedagogy research: A commemorative volume for Claes Faerch (pp.134-141). Clevedon, UK: Multilingual Matters.

Gass, S., & Selinker, L. (1994). Second language acquisition: An introductory course. Hillsdale, NJ: Lawrence Erlbaum.

Gatbonton, E., & Segalowitz, N. (1988). Creative automatization: Principles for promoting fluency within a communicative framework. TESOL Quarterly, 22, 473-492.

Genesee, F. (1976). The role of intelligence in second language learning. Language Learning, 26, 267-280.

Genesee, F. (1987). Learning through two languages. NY: Newbury House.

Genesee, F., & Hamayan, E. (1980). Individual differences in second language learning. Applied Psycholinguistics, 1, 95-110.

Geva, E., & Ryan, E. (1993). Linguistic and cognitive correlates of academic skills in first and second languages. Language Learning, 43, 5-42.

Gleitman, L., & Gleitman, H. (1970). Phrase and paraphrase. NY: W. W. Norton & Co.

Gleitman, H & Gleitman, L (1979). Language use and language judgment. In C. Fillmore, D. Kempler & W. Wang (Eds.), Individual differences in language ability and language behavior. New York : Academic Press.

Gleitman, L.R., Gleitman, H. & Shipley, E. (1972). The emergence of the child as grammarian. Cognition, 1, 137-164.

Goldfield, B., & Snow, C. (1989). Individual differences in language acquisition. In J. Berko Gleason (Ed.), The development of language (2cd ed., pp. 303-325). Columbus, OH: Merrill Publishing.

Gould, S. J. (1996). The mismeasure of man (Rev. ed.). New York: W.W. Norton.

Gouvernement du Québec, Ministère de l'Education. (1981). Programme d'études, primaire: anglais, langue seconde. Quebec: Direction de la formation générale.

Green, P. (1974/75). Aptitude testing: An on-going experiment. Audio-Visual Journal, 12, 205-210.

Griffiths, R., & Sheen, R. (1992). Disembedded figures in the landscape: A reappraisal of L2 research on field dependence/independence. Applied Linguistics, 13, 133-148.

197

Grimm, L., & Yarnold, P. (1995). Reading and understanding multivariate statistics. Washington, DC: American Psychological Association.

Harley, B., & Hart, D. (1997). Language aptitude and second language proficiency in classroom learners of different starting ages. Studies in Second Language Acquisition, 19, 379-400.

Harley, B., & Swain, M. (1984). The interlanguage of immersion and its implications for second language teaching. In A. Davies, C. Criper, & A. Howatt (Eds.), Interlanguage (pp. 291-311). Edinburgh: Edinburgh University Press.

Harrington, M., & Sawyer, M. (1992). L2 working memory capacity and L2 reading skill. Studies in Second Language Acquisition, 14, 25-38.

Hatch, E. (1974). Second language learning -- universals? Working Papers on Bilingualism, 3, 1-17.

Horgan, D. (1978). How to answer questions when you've got nothing to say. Journal of Child Language, 5, 159-165.

Horn, J. L. (1989). Models of intelligence. In R. Linn (Ed.), Intelligence (pp. 29-73). Urbana, IL: University of Illinois Press.

Horwitz, E. (1987). Linguistic and communicative competence: Reassessing foreign language aptitude. In B. VanPatten, T. Dvorak, & J. Lee (Eds.), Foreign language learning (pp. 146-157). Cambridge, MA: Newbury House.

Hudson, T. (1993). Nothing does not equal zero: Problems with applying developmental sequence findings to assessment and pedagogy. Studies in Second Language Acquisition, 15, 461-493.

Hulstijn, J. (1990). A comparison between the information-processing and the Analysis/Control approaches to language learning. Applied Linguistics, 11, 30-45.

Hulstijn, J. (in press). Intentional and incidental second-language vocabulary learning: A reappraisal of elaboration, rehearsal and automaticity. In P. Robinson (Ed.), Cognition and second language instruction. Cambridge University Press.

Humes-Bartlo, M. (1989). Variation in children's ability to learn second languages. In K. Hyltenstam & L. Obler (Eds.), Bilingualism across the lifespan (pp. 41-54). Cambridge: Cambridge University Press.

Hunt, D., Butler, L., Noy, J., & Rosser, M. (1978). Assessing conceptual level by the paragraph completion method. Toronto: OISE.

Ianco-Worrall, A. (1972). Bilingualism and cognitive development. Child Development, 43, 1390-1400.

Ioup, G., Boustagui, E., El Tigi, M., & Moselle, M. (1994). Reexamining the Critical Period Hypothesis: A case study of successful adult SLA in a naturalistic environment. Studies in Second Language Acquisition, 16, 73-98.

Johnson, J., & Rosano, T. (1993). Relation of cognitive style to metaphor interpretation and second language proficiency. Applied Psycholinguistics, 14, 159-175.

Johnston, L. (1977). Hi Mom! Hi Dad! NY: Meadowbrook.

Johnston, L. (1985). Do they ever grow up? Toronto: Stoddart.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. Psychological Review, 99, 122-149.

Karanth, P., & Suchitra, M. (1993). Literacy acquisition and grammaticality judgments in children. In R. Scholes (Ed.), Literacy and language analysis (pp. 143-156). Hillsdale, NJ: Lawrence Erlbaum.

Kessel, F. (1970). The role of syntax in children's comprehension from ages six to twelve. Monographs of the Society for Research in Child Development, 35(6, Serial No. 139).

Krashen, S. (1981). Aptitude and attitude in relation to second language acquisition and learning. In K. Diller (Ed.), Universals in language learning aptitude (pp. 155-175). Rowley, MA: Newbury House.

Krashen, S. (1982). Principles and practice in second language acquisition. Oxford: Pergamon.

Krashen, S. (1983). Newmark's 'ignorance hypothesis' and current second language acquisition theory. In S. Gass & L. Selinker (Eds.), Language transfer in language learning (pp. 135-153). Rowley, MA: Newbury House.

Krashen, S. (1985). The input hypothesis: Issues and implications. London: Longman.

Krashen, S., & Terrell, T. (1983). The Natural Approach: Language acquisition in the classroom. Oxford and San Francisco: Pergamon/Alemany.

Larsen-Freeman, D., & Long, M. (1991). An introduction to second language acquisition research. London/NY: Longman.

Leopold, W.F. (1939-49). Speech development of a bilingual child: A linguist's record. Evanston, IL: Northwestern University Press.

Lightbown, P. M. (1983). Acquiring English L2 in Quebec classrooms. In S. Felix & H. Wode (Eds.), Language development at the crossroads (pp. 101-120). Tübingen: Gunter Narr Verlag.

Lightbown, P. M. (1985). Great expectations: Second-language acquisition research and classroom teaching. Applied Linguistics, 6, 173-189.

Lightbown, P. M. (1992). Can they do it themselves? A comprehension-based ESL course for young children. In R. Courchêne, J. St. John, C. Thérien, & J. Glidden (Eds.), Comprehension-based language teaching: Current trends. Ottawa: University of Ottawa Press.

Lightbown, P.M., & Spada, N. (1990). Focus-on-form and corrective feedback in communicative language teaching. Studies in Second Language Acquisition, 12, 429-448.

Lightbown, P.M., & Spada, N. (1993). How languages are learned. Oxford: Oxford University Press.

Lightbown, P.M., & Spada, N. (1994). An innovative program for primary ESL in Quebec. TESOL Quarterly, 28, 563-579.

Lightbown, P.M., & Spada, N. (1997). Learning English as a second language in a special school in Québec. Canadian Modern Language Review, 53, 315-355.

Lightbown, P.M., & White, J. (1995, March). Learning by osmosis? Paper presented at SLRF '95, Pittsburgh, PA.

Lightbown, P.M., Meara, P., & Halter, R.H. (in press). Contrasting patterns in classroom lexical environments. In D. Albrechtsen, B. Henriksen, & I. Mees (Eds.), Festschrift for Kirsten Haastrup. Odense, Denmark: Odense University Press.

Loman, D. (1989). Human intelligence: An introduction to advances in theory and research. Review of Educational Research, 59, 333-373.

Long, M. H. (1988). Instructed interlanguage development. In L. Beebe (Ed.), Issues in second language acquisition (pp. 115-141).NY: Newbury House.

Long, M. H. (1990). The least a second language acquisition theory needs to explain. TESOL Quarterly, 24, 649-666.

Long, M. H., & Robinson, P. (1998). Focus on form: Theory, research, and practice. In C. Doughty & J. Williams (Eds.), Focus on form in classroom second language acquisition (pp. 15-41). Cambridge: Cambridge University Press.

Loring, T. (1995). A yes/no vocabulary test in a university placement setting. Unpublished master's thesis, Concordia University, Montreal.

Lorr, M. (1983). Cluster analysis for the social sciences. San Francisco, CA: Jossey-Bass.

Lyster, R. (1987). Testing for language transfer and fossilization in the immersion interlanguage. Unpublished manuscript, OISE, Toronto.

Lyster, R., & Ranta, L. (1997). Corrective feedback and learner uptake: Negotiation of form in communicative classrooms. Studies in Second Language Acquisition, 19, 37-66.

MacWhinney, B. (1989). Competition and connectionism. In B. MacWhinney & E. Bates (Eds.), The crosslinguistic study of sentence processing (pp. 422-457). Cambridge: Cambridge University Press.

Marr, D. (1982). Vision. San Francisco: W.H.Freeman.

Martens, M. (1988). Recognition and production of pronouns by francophone learners of English as a second language. Unpublished master's thesis, Concordia University, Montreal.

Masny, D. (1987). The role of language and cognition in second language metalinguistic awareness. In J. Lantolf & A. Labarca (Eds.), Research in second language learning: Focus on the classroom (pp. 59-73). Norwood, NJ: Ablex Publishing.

Masny, D., & d'Anglejan, A. (1985). Language, cognition and second language grammaticality judgements. Journal of Psycholinguistic Research, 14, 175-97.

McLaughlin, B. (1987). Theories of second language learning. London: Edward Arnold.

McLaughlin, B. (1990a, October). Another look at aptitude. Paper presented at Ohio State University.

McLaughlin, B. (1990b). Restructuring. Applied Linguistics, 11, 113-128.

McLaughlin, B. (1990c). The relationship between first and second languages: Language proficiency and language aptitude. In B. Harley, P. Allen, J. Cummins, & M. Swain (Eds.), The development of second language proficiency (pp. 158-174). Cambridge: Cambridge University Press.

McLaughlin, B. (1995). Aptitude from an information-processing perspective. Language Testing, 12, 370-387.

McLaughlin, B., & Harrington, M. (1989). Second language acquisition. Annual Review of Applied Linguistics, 10, 122-134.

McLaughlin, B., & Heredia, R. (1996). Information-processing approaches to research on second language acquisition and use. In W. Ritchie & T. Bhatia (Eds.), Handbook of second language acquisition (pp. 213-228). NY: Academic Press.

Meara, P. (1994). LLAT: Lingua Language Aptitude Test (v. 1.3). Swansea: Centre for Applied Languge Studies, Swansea University of Wales.

Meara, P., & Buxton, B. (1987). An alternative to multiple-choice vocabulary tests. Language Testing, 4, 142-154.

Meara, P. & Jones, G. (1989). Eurocentres vocabulary test 10 KA. Zurich: Eurocentres.

Meara, P., Lightbown, P.M. & Halter, R.H. (1994). The effect of cognates on the applicability of yes/no vocabulary tests. Canadian Modern Language Review, 50, 296-311.

Meier, S. (1994). The chronic crisis in psychological measurement and assessment: A historical survey. San Diego, CA: Academic Press.

Meisel, J. (1983). Strategies of second language acquisition: More than one kind of simplification. In R. Andersen (Ed.), Pidginization and creolization as language acquisition (pp. 120-57). Rowley, MA: Newbury House.

Meisel, J., Clahsen, H., & Pienemann, M. (1981). On determining developmental stages in natural second language acquisition. Studies in Second Language Acquisition, 3, 109-135.

Mungubhai, F. (1991). The processing behaviours of adult L2 learners and their relationship to L2 proficiency. Applied Linguistics, 12, 268-298.

Naiman, N., Fröhlich, M., Stern, H., & Todesco, A. (1977). The good language learner. Toronto: OISE. Reprinted in 1996, Clevedon: Multilingual Matters.

Nation, R., & McLaughlin, B. (1986). Language learning in multilingual subjects: An information-processing point of view. In V. Cook (Ed.), Experimental approaches to second language acquisition (pp. 41-53). Oxford: Pergamon.

Nayak, N., Hansen, N., Krueger, N., & McLaughlin, B. (1990). Language-learning strategies in monolingual and multilingual adults. Language Learning, 40, 221-244.

Nelson, K. (1973). Structure and strategy in learning to talk. Monographs of the Society for Research in Child Development, 38(1-2, Serial No. 149).

Nelson, K. (1975). The nominal shift in semantic-syntactic development. Cognitive Psychology, 7, 461-479.

Nicholas, H. (1987). A comparative study of the acquisition of German as a first and as a second language. Unpublished doctoral dissertation, Monash University, Melbourne.

Norusis, M. (1985). SPSS-X advanced statistics guide. Chicago: SPSS.

Novoa, L., Fein, D., & Obler, L. (1988). Talent in foreign languages: A case study. In L. Obler & D. Fein (Eds.), The exceptional brain (pp. 294-302). NY: Guildford Press.

O'Malley, J. M., & Chamot, A. U. (1987). The Cognitive Academic Language Learning Approach: A bridge to the mainstream. TESOL Quarterly, 21, 227-249.

O'Malley, J., & Chamot, A. (1990). Learning strategies in second language acquisition. Cambridge: Cambridge University Press.

O'Malley, J.M., & Chamot, A.U. (1993). Learner characteristics in second-language acquisition. In S. Omaggio Hadley (Ed.), Research in language learning: Principles, processes, and prospects (pp. 96- 123). Lincolnwood, IL: National Textbook Co.

Obler, L. (1989). Exceptional second language learners. In S. Gass, C. Madden, D. Preston, & L. Selinker (Eds.), Variation in second language acquisition vol. 2: Psycholinguistic issues (pp. 141-159). Clevedon & Philadelphia: Multilingual Matters.

Oller, J. (1973). Cloze tests of second language proficiency and what they measure. Language Learning, 23, 105-118.

Olson, D. (1994). The world on paper: The conceptual and cognitive implications of writing and reading. Cambridge, UK: Cambridge University Press.

Olson, D. (1996). Towards a psychology of literacy: On the relations between speech and writing. Cognition, 60, 83-104.

Oxford, R. (1990). Styles, strategies, and aptitude: Connections for language learning. In T. Parry & C.Stansfield (Eds.), Language aptitude reconsidered (pp. 67-125). Englewood Cliffs, NJ: Prentice-Hall Regents.

Parry, T., & Child, J. (1990). Preliminary investigation of the relationship between VORD, MLAT, and language proficiency. In T. Parry & C. Stansfield (Eds.), Language aptitude reconsidered (pp. 30-66). Englewood Cliffs, NJ: Prentice Hall Regents.

Peters, A. (1981). Language learning strategies: Does the whole equal the sum of the parts? In K. Diller (Ed.), Universals in language learning aptitude (pp. 37-52). Rowley, MA: Newbury House.

Petersen, C., & Al-Haik, A. (1976). The development of the Defense Language Aptitude Battery (DLAB). Educational and Psychological Measurement, 36, 369-380.

Pienemann, M., Johnston, M., & Brindley, G. (1988). Constructing an acquisition-based procedure for assessing second language acquisition. Studies in Second Language Acquisition, 10, 217-243.

Pienemann, M., Johnston, M., & Meisel, J. (1993). The multidimensional model, linguistic profiling, and related issues: A reply to Hudson. Studies in Second Language Acquisition, 15, 495-503.

Pimsleur, P. (1966). The Pimsleur language aptitude battery. NY: Harcourt Brace Jovanovich.

Pimsleur, P., Sundland, D, & Comfrey, R. (1964). Underachievement in foreign language learning. IRAL, 2, 113-150.

Politzer, R. (1970). On the use of aptitude variables in research on foreign language teaching. IRAL, 8, 333-340.

Ramer, A. (1976). Syntactic styles in emerging language. Journal of Child Language, 3, 49-62.

Ravem, R. (1973). Language acquisition in a second language environment. In J. Oller & J. Richards (Eds.), Focus on the learner (pp. 136-144). Rowley, MA: Newbury House.

Read, C., Zhang, Y., Nie, H., & Ding, B. (1986). The ability to manipulate speech sounds depends on knowing alphabetic reading. Cognition, 24, 31-44.

Reber, A. (1989). Implicit learning and tacit knowledge. Journal of Experimental Psychology: General, 118, 219-235.

Reid, J. (1995). (Ed.). Learning styles in the ESL/EFL classroom. NY: Heinle & Heinle.

Reves, T. (1983). What makes a good language learner? Unpublished doctoral dissertation, Hebrew University of Jerusalem.

Ricciardelli, L. (1993). Two components of metalinguistic awareness: Control of linguistic processing and analysis of linguistic knowledge. Applied Psycholinguistics, 14, 349-367.

Robinson, P. (1995). Attention, memory, and the 'noticing' hypothesis. Language Learning, 45, 283-331.

Robinson, P. (1996). Learning simple and complex second language rules under implicit, incidental, rule-search, and instructed conditions. Studies in Second Language Acquisition, 18, 27-67.

Robinson, P. (1997). Individual differences and the fundamental similarity of implicit and explicit adult second language learning. Language Learning, 47, 45-99.

Romesburg, H.C. (1984). Cluster analysis for researchers. Belmont, CA: Lifetime Learning Publications.

Rutherford, W., & Sharwood Smith, M. (1985). Consciousness raising and Universal Grammar. Applied Linguistics, 6, 274-282.

Ryan, E.B. & Ledger, G. (1984) Learning to attend to sentence structure: Links between metalinguistic development and reading. In J. Downing & R. Valtin (Eds.) Language awareness and learning to read (pp. 151-171). New York: Springer-Verlag.

Sasaki, M. (1993a). Relationships among second language proficiency, foreign language aptitude, and intelligence: A protocol analysis. Language Learning, 43, 469-505.

Sasaki, M. (1993b). Relationships among second language proficiency, foreign language aptitude, and intelligence: A structural equation modeling approach. Language Learning, 43, 313-344.

Sawyer, M. (1992). Language aptitude and language experience: Are they related? The Language Programs of the International University of Japan Working Papers, 3, 27-45.

Saywitz, K. & Wilkinson, L. (1982). Age-related differences in metalinguistic awareness. In S. Kuscaj II (Ed.), Language development: Volume II. Language, thought, and culture (pp. 229-250). Hillsdale, NJ: Lawrence Erlbaum.

Schils, E., van der Poel, M., & Weltens, B. (1991). The reliability ritual. Language Testing, 8, 125-138.

Schmidt, R. (1990). The role of consciousness in second language learning. Applied Linguistics, 11, 129-158.

Schmidt, R. (1994a). Deconstructing consciousness: In search of useful definitions for Applied Linguistics. AILA Review, 11, 11-26.

Schmidt, R. (1994b). Implicit learning and the cognitive unconscious: Of artificial grammars and SLA. In N. Ellis (Ed.), Implicit and explicit learning of languages (pp. 165-209). London: Academic Press.

Schmidt, R. & Frota, S. (1986). Developing basic conversational ability in a second language: A case study of an adult learner of Portuguese. In R. Day (Ed.), Talking to learn: Conversation in second language acquisition (pp. 237-322). Rowley, MA: Newbury House.

Schneiderman, E., & Desmarais, C. (1988). A neuropsychological substrate for talent in second-language acquisition. In L. Obler & D. Fein (Eds.), The exceptional brain (pp. 103-126). NY/London: The Guildford Press.

Segalowitz, N. (1997). Individual differences in second language acquisition. In A. M. B. de Groot & J. F. Kroll (Eds.) , Tutorials in bilingualism: Psycholinguistic perspectives (pp. 85-112). Hillsdale, NJ: Lawrence Erlbaum.

204

Segalowitz, N. & Segalowitz, S. (1993). Skilled performance, practice and the differentiation of speed-up from automatization effects: Evidence from second language word recognition. Applied Psychololinguistics, 14, 369-385.

Selinker, L. (1971). Interlanguage. IRAL, 10, 209-231.

Sharwood Smith, M. (1981). Consciousness-raising and the second language learner. Applied Linguistics, 11, 159-168.

Sharwood Smith, M. (1991a). Language modules and bilingual processing. In E. Bialystok (Ed.), Language processing in bilingual children (pp. 10-24). Cambridge, UK: Cambridge University Press.

Sharwood Smith, M. (1991b). Speaking to many minds. Second Language Research, 7, 118-132.

Sharwood Smith, M. (1993). Input enhancement in instructed SLA: Theoretical bases. Studies in Second Language Acquisition, 15, 165-179.

Sharwood Smith, M. (1994). Second language learning: Theoretical foundations. London: Longman.

Sheen, R. (1993). A rebuttal to Chapelle's response to Griffiths and Sheen. Applied Linguistics, 14, 98-100.

Shiffrin, R., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. Psychological Review, 84, 127-190.

Skehan, P. (1982). Memory and motivation in language aptitude testing. Unpublished doctoral dissertation, University of London.

Skehan. P. (1986a). Cluster analysis and the identification of learner types. In V. Cook (Ed.), Experimental approaches to second language acquisition (pp. 81-94). Oxford: Pergamon.

Skehan, P. (1986b). The role of foreign language aptitude in a model of school learning. Language Testing, 3, 188-221.

Skehan, P. (1989). Individual differences in second-language learning. London: Edward Arnold.

Skehan, P. (1990). The relationship between native and foreign language learning ability: Educational and linguistic factors. In H. Dechert (Ed.), Current trends in European second language acquisition research (pp. 83-106). Clevedon, UK: Multilingual Matters.

Skehan, P. (1991). Individual differences in second language learning. Studies in Second Language Acquisition, 13, 275-298.

Skehan, P. (1992). Second language acquisition strategies and task-based learning. Thames Valley University Working Papers in English Language Teaching, 1, 178-208.

Skehan, P. (1998). A cognitive approach to language learning. Oxford: Oxford University Press.

Smith, C.L. & Tager-Flusberg, H. (1982). Metalinguistic awareness and language development. Journal of Experimental Child Psychology, 34, 449-468.

Snow, C. (1991). Language proficiency: Towards a definition. In G. Appel & H. Dechert (Eds.), A case for psycholinguistic cases (pp. 63-89). Amsterdam/ Philadelphia: John Benjamins.

Snow, C., & Hoefnagel-Höhle, M. (1979). Individual differences in second-language ability: A factor-analytic study. Language and Speech, 22, 151-162.

Snyder, L., Bates, E., & Bretherton, I. (1981). Content and context in early lexical development. Journal of Child Language, 8, 565-582.

Spada, N. (1979). Formal and informal learning: Some aspects of the ESL development of francophone students who do or do not have out-of-classroom exposure to English. Unpublished master's thesis, Concordia University, Montreal.

Spada, N. (1997). Form-focussed instruction and second language acquisition: A review of classroom and laboratory research. Language Teaching, 30, 73-87.

Spada, N., & Fröhlich, M. (1995). The Communicative Orientation of Language Teaching (COLT) observation scheme: Applications for research and teacher education. Sydney: National Centre for English Language Teaching and Research (NCELTR).

Spada, N., & Lightbown, P.M. (1989). Intensive ESL programs in Quebec primary schools. TESL Canada Journal, 7, 11-32.

Spada, N., & Lightbown, P.M. (1993). Instruction and the development of questions in L2 classrooms. Studies in Second Language Acquisition, 15, 205-224.

Spada, N., & Lightbown, P.M. (1999). Instruction, L1 influence and developmental readiness in second language acquisition. Modern Language Journal, 83.

Spada, N., Ranta, L., & Lightbown, P.M. (1996). Working with teachers in second language acquisition research. In J. Schachter & S. Gass (Eds.), Second language classroom research: Issues and opportunities (pp. 31-44). Mahwah, NJ: Lawrence Erlbaum.

Sparks, R., & Ganschow, L. (1991). Foreign language learning differences: Affective or native language aptitude differences? Modern Language Journal, 75, 3-16.

Sparks, R., & Ganschow, L. (1993). Searching for the cognitive locus of foreign language learning difficulties: Linking first and second language learning. Modern Language Journal, 77, 289-302.

Sparks, R., Ganschow, L., Fluharty, K., & Little, S. (1995). An exploratory study on the effects of Latin on the native language skills and foreign language aptitude of students with and without disabilities. The Classical Journal, 91, 165-184.

Sparks, R., Ganschow, L., Javorsky, J., Pohlman, J., & Patton, J. (1992). Test comparisons among students identified as high-risk, low-risk, and learning disabled in high school foreign language courses. Modern Language Journal, 76, 142-159.

Sparks, R., Ganschow, L., Kenneweg, S., & Miller, K. (1991). Use of an Orton-Gillingham approach to teaching a foreign language to dyslexic/learning disabled students: Explicit teaching of phonology in a second language. Annals of Dyslexia, 41, 96-118.

Stanovich, K. (1980). Toward an interactive-compensatory model of individual differences in the development of reading fluency. Reading Research Quarterly, 16, 360-406.

Stansfield, C., & Hansen, J., (1983). Field dependence-interdependence as a variable in second language cloze test performance. TESOL Quarterly, 17, 29-38.

Stevens, J. (1986). Applied multivariate statistics for the social sciences. Hillsdale, NJ: Lawrence Erlbaum.

Stone, E., Stone, D., & Gueutal, H. (1990). Influence of cognitive ability on responses to questionnaire measures: Measurement precision and missing response problems. Journal of Applied Psychology, 75, 418-427.

Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), Input in second language acquisition (pp. 235-253). Rowley, MA: Newbury House.

Swain, M. (1993). The output hypothesis: Just speaking and writing aren't enough. Canadian Modern Language Review, 50, 158-164.

Thomas, J. (1988). The role played by metalinguistic awareness in second and third language learning. Journal of Multilingual and Multicultural Development, 9, 235-246.

Truscott, J. (1998). Noticing in second language acquisition: A critical review. Second Language Research, 14, 103-135.

Tucker, R., Hamayan, E., & Genesee, F. (1976). Affective, cognitive and social in second language acquisition. Canadian Modern Language Review, 23, 214-226.

Tudor, I. (1996). Learner-centredness as language education. Cambridge: Cambridge University Press.

Tunmer, W., Pratt, C. & Herriman, M. (Eds.) (1984). Metalinguistic awareness in children. Berlin: Springer-Verlag.

Upshur, J. (1983). Measurement of individual differences and explanation in the language sciences. Language Learning, 33, [Special issue, 'An epistemology for the language sciences'], 99-140.

Upshur, J., & Turner, C. (1995). Constructing rating scales for second language tests. English Language Teaching Journal, 49, 3-12.

Van Kleeck, A. (1982). The emergence of linguistic awareness: A cognitive framework. Merrill-Palmer Quarterly, 28, 237-265.

207

Van Kleeck, A. (1984). Metalinguistic skills: Cutting across spoken and written language and problem-solving activities. In G. Wallach & K. Butler (Eds.), Language learning disabilities in school-age children (pp. 128-153). Baltimore: Williams & Wilkins.

VanPatten, B. (1990). Attending to form and content in the input: An experiment in consciousness. Studies in Second Language Acquisition, 12, 287-301.

Ward, J. H. (1963). Hierarchical grouping to optimize an objective function. Journal of the American Statistical Association, 58, 236-244.

Watts, W., & Snow, S. (1993). L'anglais intensif au Québec 1976-1993. Montreal: SPEAQ.

Welkowitz, J., Ewen, R., & Cohen, J. (1991). Introductory statistics for the behavioral sciences. San Diego: Harcourt Brace Jovanovich.

Wells, G. (1985). Language development in the pre-school years. Cambridge: Cambridge University Press.

Wells, G. (1986). Variation in child language. In P. Fletcher & M. Garman (Eds.), Language acquisition (2cd ed., pp. 109-139). Cambridge: Cambridge University Press.

Wesche, M.B. (1981). Language aptitude measures in streaming, matching students with methods, and diagnosis of learning problems. In K. Diller (Ed.), Individual differences and universals in language learning aptitude (pp. 119-154). Rowley, MA: Newbury House.

Wesche, M.B., Edwards, H., & Wells, W. (1982). Foreign language aptitude and intelligence. Applied Psycholinguistics, 3, 127-140.

White, J. (1996). An input enhancement study with ESL children: Effects on the acquisition of possessive determiners. Unpublished doctoral dissertation, McGill University, Montreal.

White, J. (1998). Getting learners' attention: A typographical input enhancement study. In C. Doughty & J. Williams (Eds.), Focus on form in classroom second language acquisition (pp. 85-113). Cambridge: Cambridge University Press.

White, L. (1987). Against comprehensible input: The input hypothesis and the development of second language competence. Applied Linguistics, 8, 95-110.

White, L. (1989). Universal Grammar and second language acquisition. Amsterdam: John Benjamins.

White, L. (1991). Adverb placement in second language research: Some effects of positive and negative evidence in the classroom. Second Language Research, 7, 133-161.

White, L., Spada, N., Lightbown, P.M., & Ranta, L. (1991). Input enhancement and L2 question formation. Applied Linguistics, 12, 416-432.

Wilkinson, L., Hill, M., & Vang, E. (1992). SYSTAT: Statistics, Version 5.2 Edition. Evanston, IL: SYSTAT Inc.

Willing, K. (1989). Teaching how to learn: Learning strategies in ESL. Sydney: National Centre for English Language Teachng and Research, Macquarie University.

Willows, D., & Ryan, E. B. (1986). The development of grammatical sensitivity and its relationship to early reading achievement. Reading Research Quarterly, 21, 253-266.

Witkin, H., & Goodenough, D. (1981). Cognitive styles: Essence and origins -- Field dependence and field independence. Psychological Issues Monograph, 51.

Witkin, H., Moore, C., Goodenough, D., & Cox, P. (1977). Field-dependent and field-independent cognitive styles and their educational implications. Review of Educational Research, 47, 1-64.

Wode, H. (1978). The L1 vs. L2 acquisition of English interrogation. Working Papers in Bilingualism, 11, 1-13.

Wong Fillmore, L. (1979). Individual differences in second language acquisition. In C. Fillmore, D. Kempler, & W. Wang (Eds.), Individual differences in language ability and language behavior (pp. 203-228). NY: Academic Press.

Woods, A. (1983). Principal components and factor analysis in the investigation of the structure of language proficiency. In A. Hughes & D. Porter (Eds.), Current developments in language testing (pp. 43-52). London: Academic Press.

Woods, A., Fletcher, P., & Hughes, A. (1986). Statistics in language studies. Cambridge, UK: University of Cambridge Press.

Yelland, G., Pollard, J., & Mercuri, A. (1993). The metalinguistic benefits of limited contact with a second language. Applied Psycholinguistics, 14, 423-444.

Zobl, H. (1984). The wave model of linguistic change and interlanguage systems. Studies in Second Language Acquisition, 6, 160-185.

Zobl, H. (1985). Grammars in search of input and intake. In S. Gass & C. Madden (Eds.), Input in second language acquisition (pp. 32-44). Rowley, MA: Newbury House.

Appendix A

Instructions and Sample Items from Words in Sentences subtest of MLAT (Carroll & Sapon, 1958).

Instructions:

This is a test of your ability to understand the function of words and phrases in sentences.

Look at the following sample item:

LONDON is the capital of England.

He liked to go fishing in Maine.
A   B   C   D   E

In the first sentence, which we will call the **key** sentence, LONDON is printed in capital letters. Which word in the second sentence does the same thing in **that** sentence as LONDON does in the **key** sentence? The right answer is the word "he," because the key sentence is about "London," and the second sentence is about "he."

Sample items:

1.   Jill fell down AND Jack came tumbling after.

Now, you may wait out there, or you may come back on Friday if you wish.
  A   B   C   D   E

45. The child hurt himself.

Although I myself would do that by myself, Mary gained herself the help of some of
              A   B   C   D
her classmates.
    E

Appendix B

Logically possible learner types (Skehan, 1989, p. 35)

	Phonemic coding ability	Language analytic ability	Memory	
1	+	+	+	General, even high aptitude
2	+	+	-	Good auditory and analytic ability, poor memory
3	+	-	+	Good auditory and memory abilities; Poor analysis
4	+	-	-	Good auditory ability; Poor otherwise
5	-	+	+	Poor auditory; Good analysis and memory
6	-	+	-	Poor auditory and memory; Good analysis
7	-	-	+	Poor auditory and analysis; Good memory
8	-	-	-	Evenly poor aptitude

Appendix C

Grammaticality Sensitivity measure

Nom de l'élève: _____

Trouve les Erreurs

Lis chaque phrase attentivement. Si tu trouves une erreur dans une phrase, encercle l'erreur et corrige-la dans l'espace prévu à cet effet. S'il n'y a pas d'erreur, écris "pas d'erreur". Chaque phrase incorrecte ne contient qu'une seule erreur alors que quelques phrases n'ont aucune erreur.

1. Les légumes frais sont plus bons pour la santé que les bonbons.

   _____

2. Toutes les enfants prennent l'autobus à 15 heures.

   _____

3. Grand-père leurs a raconté une histoire.

   _____

4. L'hiver dernier je fais du ski tous les jours.

   _____

5. Nous nous sommes beaucoup amusés à Calgary.

   _____

6. Mes parents ont allés à Toronto.

   _____

7. Sa mère l'a prêté de l'argent.

   _____

8. Nos voisins ont achetaient une nouvelle voiture.

   _____

9. Il faut que je vais chez le dentiste.

   _____

10. Les joueur de hockey gagnent trop d'argent.

_____

11. Hier soir elles sont sorties pour aller au théâtre.

_____

12. Il a passé les plus beaux vacances de sa vie!

_____

13. Richard sent va à la campagne.

_____

14. Je ne sais pas très bien nagé.

_____

15. Les premières deux personnes ont eu un prix.

_____

16. Ils nous avons offert un repas délicieux.

_____

17. Après nager pendant une heure, on était très fatigué.

_____

18. Lorsque mon oncle a ouvri la porte, il a eu une grande surprise.

_____

19. Nous avons complètement oublié de la téléphoner.

_____

20. J'ai besoin d'une nouvelle bicyclette.

_____

Appendix D

Learning Preferences questionnaire

Nom de l'élève: _____

Ma façon d'apprendre l'anglais

Réfléchis sur ta façon d'apprendre l'anglais. Ta façon n'est pas nécessairement la même que celle de tes amis. Parmi les activités suivantes, choisis celles qui, à ton avis, conviennent le plus à ta façon d'apprendre.

Ce qui m'aide à apprendre l'anglais:

1. écrire des textes en anglais
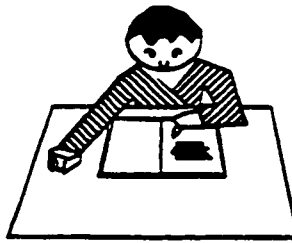
Ça m'aide: beaucoup   un peu   très peu   pas du tout

2. faire des jeux
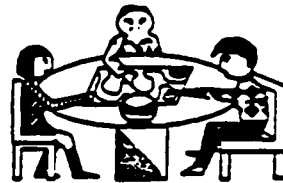
Ça m'aide: beaucoup   un peu   très peu   pas du tout

3. chercher des mots dans le dictionnaire

Ça m'aide: beaucoup   un peu   très peu   pas du tout

214

4. écouter des cassettes en anglais
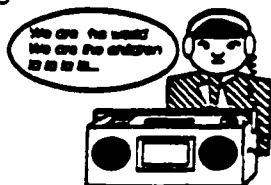
Ça m'aide: beaucoup   un peu   très peu   pas du tout

5. lire des histoires en anglais

Ça m'aide: beaucoup   un peu   très peu   pas du tout

6. faire des activités orales
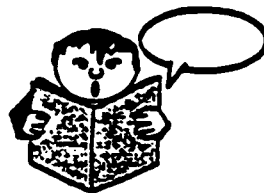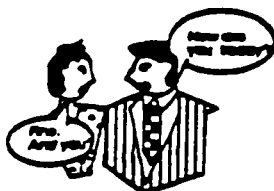

Ça m'aide: beaucoup   un peu   très peu   pas du tout

· 7. apprendre et réciter une poésie
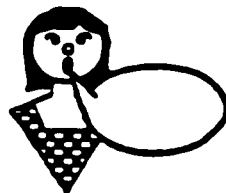
Ça m'aide: beaucoup   un peu   très peu   pas du tout

8. consulter un livre de grammaire
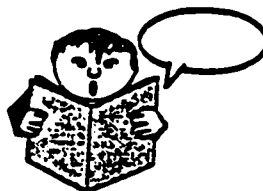
Ça m'aide: beaucoup   un peu   très peu   pas du tout

9. réviser des listes de vocabulaire

Ça m'aide: beaucoup   un peu   très peu   pas du tout

10. écouter des explications données par le professeur
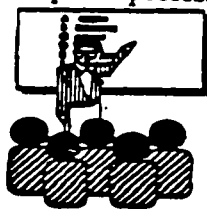
Ça m'aide: beaucoup   un peu   très peu   pas du tout

11. écouter des videos en anglais
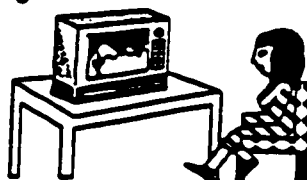
Ça m'aide: beaucoup   un peu   très peu   pas du tout

12. faire des exercices de grammaire

Ça m'aide: beaucoup   un peu   très peu   pas du tout

13. me faire corriger par le professeur

Ça m'aide: beaucoup   un peu   très peu   pas du tout

14. créer et jouer des sketches
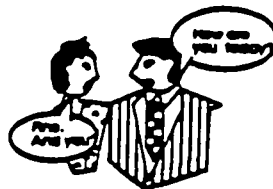
Ça m'aide: beaucoup   un peu   très peu   pas du tout

15. écouter la télévision en anglais
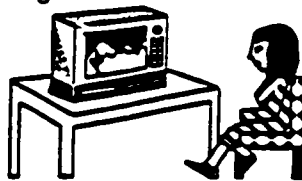
Ça m'aide: beaucoup   un peu   très peu   pas du tout

16.   Est-ce qu'il y a d'autres activités qui t'aident beaucoup à apprendre l'anglais?

_____

_____

_____

_____

_____

Appendix E
Measures of Engish Proficiency
Paper-and-Pencil Tasks

.

Appendix E-1

Aural Vocabulary Recognition Test (AVR)

Sample page

PAGE 1
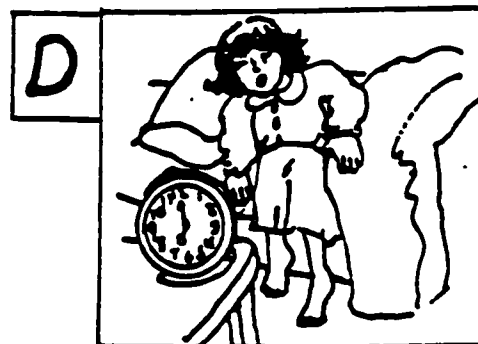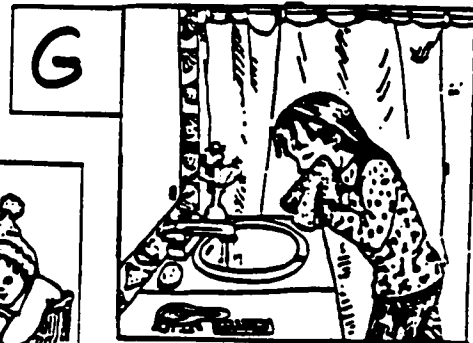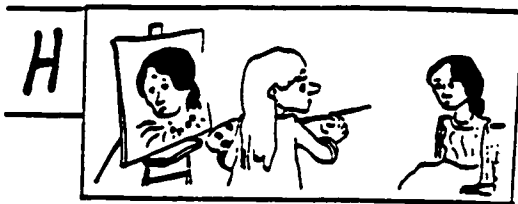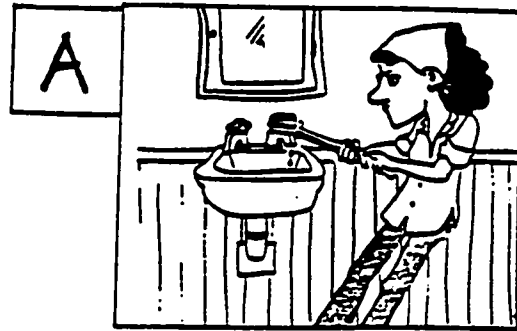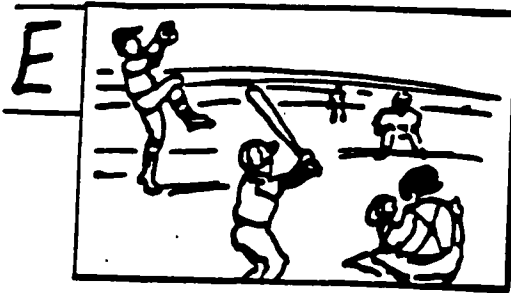
Vocabulary items on the AVR:

Page 1   paint, wake up, play, carry, buy, write, wash, fix

Page 2   push, cook, chase, break, burn, pour, hold, crash

Page 3   deer, rabbit, octopus, owl, bee, butterfly, squirrel, goat

Page 4   teacher, schoolbus, ruler, schoolbag, computer, lunchbox, notebook, globe

Page 5   cookies, napkins, peanut butter, forks, cakes, pineapple, tea, spoons

Page 6   church, school, truck, playground, store, stoplight, puddle, sidewalk

Page 7   yard, toothbrush, soap, newspaper, needle, basket, fire, tie

Page 8   bottle, logs, boat, hammer, cage, rope, nails, nest

Listening Comprehension Test (Sept.)

PARTIE I – COMPRÉHENSION AUDITIVE:

Tu entrendras 10 phrases en anglais. Chaque phrase sera prononcée deux fois. Choisis pour chaque phrase le dessin qui représente le mieux la phrase. Sur ta feuille de réponses tu inscris la lettre qui correspond à ce dessin à côté du numéro de la phrase. Pour commencer nous allons faire un exemple ensemble.
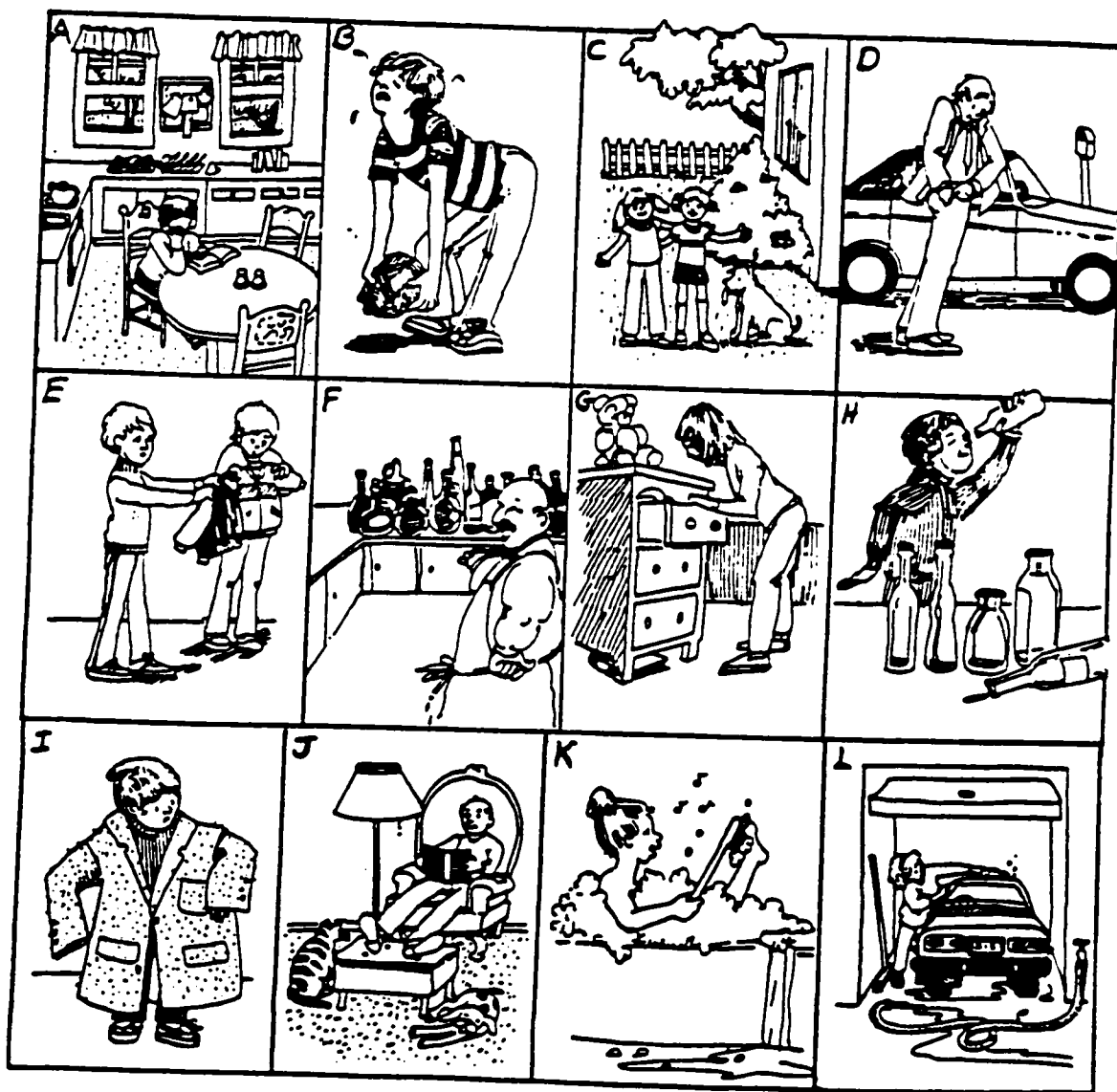
Sentences on Listening Comprehension Test (Sept.):

1. We're in the yard.

2. I'm in the kitchen reading.

3. She's looking for her book.

4. They're too small.

5. He's looking for his key.

6. It's too heavy.

7. It's too big.

8. She's washing her feet.

9. He's reading in the living room.

10. She's in the garage washing the car.

Cloze Test

Lecture

Complète ce texte en écrivant dans les espaces les mots anglais qui manquent.

I have two friends. Their _____
are Tom and Mary. They _____
not live on _____ same street
as I do, but they _____ to the
same school. They go to _____
by bus, but I _____ my bike.
Tom and Mary _____ eleven;
I _____ only ten. We are all in
the _____ class at school. We
play together every day after
_____.

Appendix E-4

MEQ Listening/Reading Comprehension Test

Sample Item

DIRECTIVES: Choisis parmi les quatre phrases suggérées celle qui correspond le mieux à celle que tu auras entendue.

Exemple: (The books are ours.)

    (A) They belong to us.

    (B) They belong to her.

    (C) They are my books.

    (D) They are difficult.

La réponse est:

A	B	C	D	E
●	○	○	○	○

Nous continuons.

16. (I can't sell my old bike.)

    (A) He needs to buy an old bike.

    (B) He sold his old bike to a friend.

    (C) He'd like to sell his old bike.

    (D) He found someone with an old bike to sell.

Grammaticality Judgment Task

'The Birthday Party'

Today is David Ryan's birthday. He is twelve years old. He's having a party with his friends and family. Much people are invited. David's grandfather can't come because he moved to Florida last winter. David is happy to have a party but he misses her grandfather. He hopes everyone else can come.

Susan is going at David's party. She's carrying a present in his arm. It's a game. Her mother helped Susan to choose it. Susan is happy to be invited to the party and to wear his pretty new dress. She don't want to be late.

Marc is David's best friend. He lives just next door. Marc and David is in the same class at school. He has a special present for David in her hand. It's a baseball bat. Marc and her father picked it out. David loves baseball and plays every summer. Last year his team won the city championship.

David's two cousins are also invited to the party. Annie is ten. She's tall and thin and his hair is blond. Her brother Eric is tall and thin too, but her hair is black. He's twelve. Now David and his cousin are the same age. Annie is bringing her cousin a model plane and Eric is give him a book. He hopes David will enjoy reading her new book.

David is excited to see your friends. His younger sister Diane is excited too. She has a party hat on his head. She has balloons in his hand. Diane thinks birthday parties are super. Her birthday is on July. She will be six year old.

There is a baby sister too. Her name is Carole. Carole is holding balloons that his sister Diane gave her. His mother is showing her the candles on David's birthday cake. It's chocolate. David's mother prepared the cake for his son's party. Diane prepared the lemonade. Everyone likes chocolate cake and lemonade!

The children are all having the good time. David is playing with her new plane. Diane is sitting near her baby sister. She's looking at his brother's birthday cards. Susan and Marc are playing a game with Eric and her sister. It's a great parties.

David's father, Mr. Ryan, is home from work early. He has a package for her son. It's big but it's not heavy. David opens the box. Another box is inside. David opens five boxes! In the last box is an envelope. It's a plane ticket for David to visit his grandfather in Florida! He is so happy!

Mr. Ryan have a new camera. He want to take many pictures of his son's party. First, he'll take a picture of David with her family and friends. Then he'll photograph her wife and the children. Finally, Mrs. Ryan will take a picture of his husband with David. The pictures will help everyone remember David's wonderful party.

Appendix E-6

Yes/No Vocabulary Test

SAMPLE

Read through the lists of words carefully.

For each word:

- if you know what it means, place a check mark (✓) beside the word.

- if you don't know what it means, or if you are not sure, leave the box blank.

Do not guess. Put a check mark beside the word **only** if you are sure you know the word in English. **Some of these words are not real English words.**

Here are some examples in French.

1. (✓) chat   2. (✓) école   3. ( ) couquir

You know what **chat** means. You know what **école** means. You do not know what **couquir** means. **Couquir** looks like a French word, but it is not a real French word.

1. ( ) wallage	21. ( ) condick	41. ( ) rather
2. ( ) foot	22. ( ) look after	42. ( ) trust
3. ( ) hallett	23. ( ) hear	43. ( ) pocock
4. ( ) west	24. ( ) pencil	44. ( ) churchlow
5. ( ) hospite	25. ( ) roof	45. ( ) group

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

227

Appendix F

Oral Production Tasks

(Questions)

.

.

Appendix F-1

'Five Questions' Picture Card Game

Sample picture set
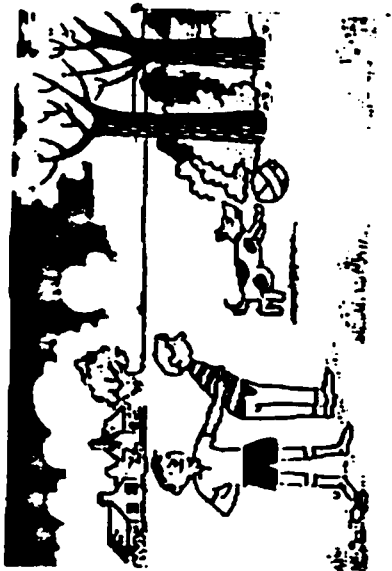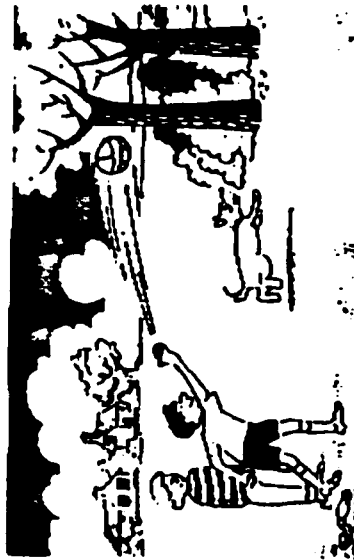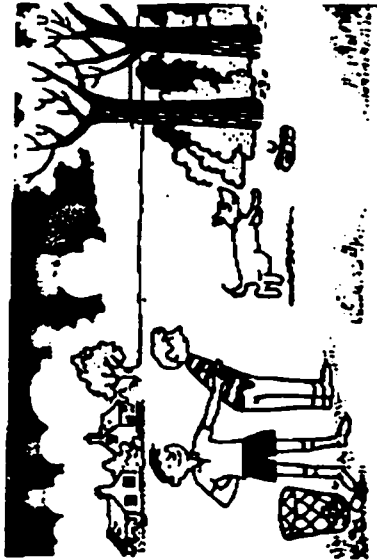
Appendix F-2

'Five Questions' Picture Card Game

Sample Transcript

Note: + indicates the student's speech; = indicates the interviewer's speech

Questions: Boys and dog

+   Okay uh, do you, there, the dog there (laughs) is it the ball who the dog go to catch?
=   No.
+   No? The shoes it's on the floor or in the in the air?
=   In the air.
+   In the air? Five question! (laughs)
=   Yes! (laughs)
+   How many trees there was there?
=   Two.
+   Two. There there is some paper in uh next to the garbage?
=   Yes.
+   Four question now?
=   Four, right.
+   Okay. Which is the colour of the, the shirt of the, the boy? Boy?
=   Uh, which boy?
+   Just one of the shirt?
=   One boy has a white shirt.
+   Is uh it's it's the pants? This one, the pant, I think?
=   Oh, oh, the pants, okay, he has white uh he has a black and white shirt.
+   Okay.

Questions: Children & school bus

+   The -- what what side the sun the sun is uh is on your paper?
=   On the left.
+   On the left. Okay. How many students in your your card [la], in your card?
=   Students?
+   Yes, students.
=   Uh, I would say one, two, three, four, five...I'd say seven.
+   Seven, okay. There was animals?
=   Yes. Three, I think.
+   Yes, I -- uh, how how many it's on the bicycle? Sorry.
=   Two.
+   Two. How many have the cap?
=   Two.
+   Okay! The, you have this one.

Questions: Busy downtown street

+   In your picture uh do you have a, a boy with her bicycle who want uh, uh...who want uh, pass the the street [mais], how do you say [traverser]?

=   Cross.

+   Cross the street?

=   Um...I guess -- n-, uh, you mean is he crossing the street?

+   Yes.

=   No.

+   No? Okay. Okay uh do you have a boy with her bicycle in your picture?

=   Yes.

+   Yes? Okay. Uh, uh how many policemen?

=   Two.

+   Two.

=   Can you maybe ask me a question that begins with 'where'?

+   'Where'? 'Where' 'where' 'where', mmhm. Where is your boy with her bicycle?

=   He's on the sidewalk.

+   Oh, okay. There there is there there is some woman in the middle of the street in the bottom of the picture?

=   Yes.

+   Yes? It's this one.

Appendix G

Oral Production Tasks

(Possessive Determiners)

Appendix G-1

Possessive Determiner Elicitation Task

Sample Picture

What's the matter, Mama... don't you like my haircut?

Appendix G-2

Sample Transcript

Note: + indicates the student's speech; = indicates the interviewer's speech

Pronouns: Haircut

+   Uh, okay, it's uh the the little girl \cup\ her hair and the woman saw that she saw her lit- and uh her Mom saw, saw her go with the, with the *\haircup\ -cut*, uh, the hair it's on the floor and she she uh, she crawled.

=   Okay, uh..why is the mother crying?

+   Because uh the the hair it's not very beautiful uh how the the little girl cut.

=   Okay, very good.

Pronouns: Band-aids

+   Okay uh, the, if the mother to said at her uh her husband uh the, the, *your her* girl have a, a little accident, and uh and when he see *the her* girl uh, he can uh, he can uh see uh she have many plaster.

=   Mmhm.

+   And uh maybe it's not a true accident because uh, she have s- she have too too many I think.

=   Where are the plasters?

+   On her stomach, in her in her, arm legs, knees, ankle.

=   Okay, yeah, all over. Okay!

Pronouns: Make-up mess

+   Okay, it's uh the the little girl take the, the make-up of her mother, and uh she she want to uh, to to to make herself more beautiful, and uh the all of the make-up it's on the floor and her face it's not very beautiful and uh she uh her sweater it's all covered with the the make-up, and *you can saw on the you can see on the* floor uh, the the the the \beautee\ \beautee\ cream uh, all uh, uh, how do you say [dégoutter]?

=   Running over.

+   All running over.

=   Okay.

Pronouns: Bicycle accident

+   Okay, it's the boy, uh, it's with her uh moto- it's her tricycle okay? And he fell down, so uh, he uh he standing up and go at her home to uh, to talk at her mother. Her mother uh, take take care of her flower, and after the the her mother don't know what he uh happen, and uh the the, the boys crawl and uh, jump on *her the* hand of her mother and said "I fell down".

=   Okay, very good.

Appendix H

Table H-1

Test reliability estimates*

Measure	Method	Coefficient
Aural Vocabulary Recognition-pre (Sept.)	K-R 21	.77
Listening Test (Sept.)	K-R 21	.49
Cloze-pre (Sept.)	K-R 21	.51
Aural Vocabulary Recognition-post (Jan.)	K-R 21	.92
Cloze-post	K-R 21	.13
MEQ-32 items (Jan.)	K-R 21	.63
Grammaticality Judgment (Jan.)	K-R 21	.89
MEQ-53 (June)	K-R 21	.80
Yes/No Vocabulary Test (June)	-	-
Trouve les Erreurs	Cronbach alpha	.62

Note: K-R 21 = Kuder-Richardson, formula 21

* See discussion of the reliabilities in Chapter 5, Endnote #2, p.158.

Appendix I

Principal Components Analysis

20:25:23   SPSS Open VMS Development SYS   on VAX2::   VMS V6.1

- F A C T O R   A N A L Y S I S

Analysis number 1   Listwise deletion of cases with missing values

Extraction   1 for analysis   1, Principal Components Analysis (PC)

Initial Statistics:

Variable	Communality	*	Factor	Eigenvalue	Pct of Var	Cum Pct
VAR4	1.00000	*	1	4.65251	51.7	51.7
VAR5	1.00000	*	2	1.04262	11.6	63.3
VAR7	1.00000	*	3	.79250	8.8	72.1
VAR12	1.00000	*	4	.57486	6.4	78.5
VAR8	1.00000	*	5	.56203	6.2	84.7
VAR15	1.00000	*	6	.45173	5.0	89.7
VAR24	1.00000	*	7	.35536	3.9	93.7
VAR25	1.00000	*	8	.30845	3.4	97.1
DMTOT	1.00000	*	9	.25993	2.9	100.0

- F A C T O R   A N A L Y S I S

Eigenvalue (Scree) Plot

```
  4.653 +    *

E
I
G
E
N
V
A
L
U
E
S

  1.043 +       *
   .793 +          *
   .452 +             *    *    *
   .260 +                          *    *    *
   .000 +---+---+---+---+---+---+---+---+---+
           1   2   3   4   5   6   7   8   9
```

FACTORS

PC   extracted   3 factors.

Factor Matrix:

	Factor 1	Factor 2	Factor 3
VAR4	.76099	-.36277	.23839
VAR5	.62547	-.60287	.05158
VAR7	.68972	-.09447	.49318
VAR12	.77402	.09035	.02773
VAR8	.75168	.01372	-.30267
VAR15	.83248	-.05939	-.21585
VAR24	.50264	.63447	.44580
VAR25	.76306	.26877	-.19836
DMTOT	.71724	.22796	-.33575

- F A C T O R   A N A L Y S I S

Final Statistics:

Variable	Communality	*	Factor	Eigenvalue	Pct of Var	Cum Pct
VAR4	.76753	*	1	4.65251	51.7	51.7
VAR5	.75733	*	2	1.04262	11.6	63.3
VAR7	.72787	*	3	.79250	8.8	72.1
VAR12	.60803	*				
VAR8	.65682	*				
VAR15	.74314	*				
VAR24	.85394	*				
VAR25	.69384	*				
DMTOT	.67913	*				

VARIMAX   rotation   1 for extraction   1 in analysis   1 - Kaiser Normalization.

VARIMAX converged in 7 iterations.

Appendix J

Table J-1

One-Way ANOVA for Aural Vocabulary Recognition Test by Cluster

Source	SS	df	MS	F
Between groups	4588.33	3	1529.44	45.62*
Within groups	4392.07	131	33.53	

* $p < .001$

Table J-2

One-Way ANOVA for Listening Comprehension Test (Sept.) by Cluster

Source	SS	df	MS	F
Between groups	188.97	3	63.00	23.63*
Within groups	349.18	131	2.67	

* $p < .001$

Table J-3

One-Way ANOVA for Cloze (Sept.) by Cluster

Source	SS	df	MS	F
Between groups	142.21	3	47.40	21.16*
Within groups	293.53	131	2.24	

* $p < .001$

Table J-4

One-Way ANOVA for MEQ-32 (Jan.) by Cluster

Source	SS	df	MS	F
Between groups	6959.45	3	2319.82	31.39*
Within groups	9680.45	131	73.90	

* $p < .001$

Table J-5

One-Way ANOVA for MEQ-53 (June) by Cluster

Source	SS	df	MS	F
Between groups	10098.52	3	3366.17	73.01*
Within groups	6040.08	131	46.11	

* $p < .001$

Table J-6

One-Way ANOVA for Trouve les Erreurs by Cluster

Source	SS	df	MS	F
Between groups	457.72	3	152.57	28.92*
Within groups	691.14	131	5.28	

* $p < .001$

240

Table J-7

One-Way ANOVA for Grammaticality Judgment Task by Cluster

Source	SS	df	MS	F
Between groups	3594.19	3	1198.06	42.14*
Within groups	3724.01	131	28.43	

* $p < .001$

Table J-8

One-Way ANOVA for Yes/No Vocabulary Test by Cluster

Source	SS	df	MS	F
Between groups	1.7790	3	.5930	64.45*
Within groups	1.2053	131	.0092	

* $p < .001$

Table J-9

One-Way ANOVA for Cloze (January) by Cluster

Source	SS	df	MS	F
Between groups	159.02	3	53.01	47.79*
Within groups	145.30	131	1.11	

* $p < .001$

Appendix K

Classification of Students who performed the Oral Production Tasks (n=46)

Student Code	GS Group	Cluster
BELSI	high	1
BERJA	low	2
BOUAN	mid	3
CLANA	mid	2
CODMA	mid	2
CONGU	low	3
CYRLI	mid	1
DESAL	mid	2
DESFR	mid	2
DESJO	mid	4
DESMB	low	3
DESNB	mid	1
DUBAL	mid	2
DURMA	high	1
FORSA	low	4
GAGAN	mid	2
GREIS	mid	2
HOTAL	mid	2
HOUSY	high	1
HOUMA	low	3
LALPH	high	1
LAVKA	low	2
LEBMI	high	1
LEFGU	mid	2
LEGJE	low	3
LORJE	mid	1
MAIMA	high	1
MARCY	high	1
MARER	high	2
METSI	high	2
MONSU	mid	4
NOLAM	low	1
OUEOL	high	1
PERCH	mid	4
PERKA	mid	2
POURO	mid	1
PRUJE	mid	2
REGCA	low	3
RENPI	high	1
ROUJE	mid	3
SEVMA	mid	2
SIRGU	low	3
TASMI	low	4
VACMA	mid	2
VAIBE	mid	2
VIGIS	mid	1

242

Appendix L

Distribution of Responses to Items on Learner Preferences Questionnaire by Cluster (in Percentages)

1. écrire des textes en anglais ('write texts in English')

	'beaucoup'	'un peu'	'tres peu'	'pas du tout'
Cluster 1 (n = 53)	43	43	13	0
Cluster 2 (n = 41)	37	56	2	5
Cluster 3 (n=28)	32	57	7	4
Cluster 4 (n=13)	23	46	15	15

2. faire des jeux ('play games')

	'beaucoup'	'un peu'	'tres peu'	'pas du tout'
Cluster 1 (n = 53)	49	36	9	6
Cluster 2 (n = 41)	54	37	5	5
Cluster 3 (n=28)	54	39	0	7
Cluster 4 (n=13)	38	46	15	0

3. chercher des mots dans le dictionnaire ('look up words in the dictionary')

	'beaucoup'	'un peu'	'tres peu'	'pas du tout'
Cluster 1 (n = 53)	43	30	19	8
Cluster 2 (n = 41)	44	34	17	5
Cluster 3 (n=28)	54	39	7	0
Cluster 4 (n=13)	31	46	23	0

4. écouter des cassettes en anglais ('listen to tapes in English')

	'beaucoup'	'un peu'	'tres peu'	'pas du tout'
Cluster 1 (n = 53)	43	42	13	2
Cluster 2 (n = 41)	34	37	24	5
Cluster 3 (n=28)	43	25	25	7
Cluster 4 (n=13)	46	31	23	0

5. lire des histoires en anglais ('read stories in English')

	'beaucoup'	'un peu'	'tres peu'	'pas du tout'
Cluster 1 (n = 53)	72	21	6	2
Cluster 2 (n = 40)	58	38	5	0
Cluster 3 (n=28)	46	46	7	0
Cluster 4 (n=13)	38	38	15	8

6. faire des activités orales ('do oral activities')

	'beaucoup'	'un peu'	'tres peu'	'pas du tout'
Cluster 1 (n = 53)	81	15	2	2
Cluster 2 (n = 41)	68	24	5	2
Cluster 3 (n=28)	75	21	0	4
Cluster 4 (n=13)	62	23	15	0

7. apprendre et réciter une poésie ('learn and recite a tongue-twister')

	'beaucoup'	'un peu'	'tres peu'	'pas du tout'
Cluster 1 (n = 53)	21	38	26	15
Cluster 2 (n = 41)	10	37	24	29
Cluster 3 (n=28)	18	32	32	18
Cluster 4 (n=13)	15	31	38	15

8. consulter un livre de grammaire ('consult a grammar book')

	'beaucoup'	'un peu'	'tres peu'	'pas du tout'
Cluster 1 (n = 53)	11	38	26	25
Cluster 2 (n = 41)	15	20	34	32
Cluster 3 (n=28)	11	32	39	18
Cluster 4 (n=13)	23	8	38	31

9. réviser des listes de vocabulaire ('study vocabulary lists')

	'beaucoup'	'un peu'	'tres peu'	'pas du tout'
Cluster 1 (n = 53)	32	45	19	4
Cluster 2 (n = 41)	46	41	12	0
Cluster 3 (n=28)	25	57	11	7
Cluster 4 (n=13)	54	31	0	15

10. écouter des explications données par le professeur ('listen to the teacher's explanations')

	'beaucoup'	'un peu'	'tres peu'	'pas du tout'
Cluster 1 (n = 53)	74	19	4	4
Cluster 2 (n = 41)	76	17	2	5
Cluster 3 (n=28)	61	36	4	0
Cluster 4 (n=13)	85	8	0	8

11. écouter des vidéos en anglais ('watch videos in English')

	'beaucoup'	'un peu'	'tres peu'	'pas du tout'
Cluster 1 (n = 53)	90	8	0	2
Cluster 2 (n = 41)	80	15	5	0
Cluster 3 (n=28)	89	11	0	0
Cluster 4 (n=13)	62	38	0	0

12. faire des exercices de grammaire ('do grammar exercises')

	'beaucoup'	'un peu'	'tres peu'	'pas du tout'
Cluster 1 (n = 53)	19	43	23	15
Cluster 2 (n = 41)	5	58	25	12
Cluster 3 (n=28)	18	36	25	21
Cluster 4 (n=13)	15	38	38	8

245

13. me faire corriger par le professeur ('get corrected by the teacher')

	'beaucoup'	'un peu'	'tres peu'	'pas du tout'
Cluster 1 (n = 53)	45	40	11	4
Cluster 2 (n = 41)	43	28	17	12
Cluster 3 (n=28)	46	43	11	0
Cluster 4 (n=13)	69	23	8	0

14. créer et jouer des sketches ('create and perform skits')

	'beaucoup'	'un peu'	'tres peu'	'pas du tout'
Cluster 1 (n = 53)	68	23	6	4
Cluster 2 (n = 41)	68	28	2	2
Cluster 3 (n=28)	67	33	0	0
Cluster 4 (n=13)	62	31	8	0

15. écouter la télévision en anglais ('watch TV in English')

	'beaucoup'	'un peu'	'tres peu'	'pas du tout'
Cluster 1 (n = 53)	94	4	2	0
Cluster 2 (n = 41)	83	18	0	0
Cluster 3 (n=28)	82	14	0	4
Cluster 4 (n=13)	69	23	8	0