THE RELATIONSHIP BETWEEN FOREIGN LANGUAGE APTITUDE AND ENGLISH LANGUAGE PROFICIENCY AMONG SAUDI LEARNERS OF ENGLISH AS A FOREIGN LANGUAGE IN SAUDI UNIVERSITIES

Merzin Awdah Alshahrani

BA (English) MLitt (Applied linguistics)

A thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy in Applied Linguistics

The University of Newcastle, Australia

February, 2011
Statement of Originality

This thesis contains no material which has been accepted for the award of any other degree or diploma in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. I give consent to this copy of my thesis, when deposited in the University Library, being made available for loan and photocopying subject to the provisions of the Copyright Act 1968.

Signed

Merzin Awdah Alshahrani
Acknowledgments

After four years of my PhD study in Newcastle University in Australia, the time comes to end this journey with this work. Undoubtedly, it is impossible to achieve this goal without sacrifice, hard work, great commitment, and without the support of great people. It is difficult to express my great thanks in just a few lines. It is my pleasure to express my greatest thanks for all of the following people.

First and foremost, it is my pleasure to thank my supervisors who supported me endlessly in my research journey in the University of Newcastle. It was a privilege and an honour to have these three great supervisors: Dr Christo Moskovsky, my principal supervisor who supported me from the first day in my PhD study in the university with his intellectual insights, patience, guidance and kindness; Dr. Stefania Paolini with her exceptional expertise in statistical analyses and her great attention to detail, ensured the outcomes of this thesis exceeded my initial expectations; and Dr. Silvia Ratcheva with her humour, kindness and support in the difficult times of my research. Without their assistance, encouragement, their great academic input, and suggestions, a project of this magnitude could not be achieved.

Secondly, I would like to extend my thanks to King Khalid University in Saudi Arabia for offering me the financial support to complete my study in Australia. Also, I wish to thank the students, the class teachers and the College of Languages and Translation for offering me their support and ready participation in this project.

I owe a great debt to my inspirational professors, Dr. Abdulllah Abu Eishi, Dr. Abdullah Bin Hadi and Dr. Mohammad Zaid for their endless support and encouraging during my academic life. With their extraordinary support and encouraging for more than ten years, I can say now, I have achieved the dream.
I give a special acknowledgment to the best friends in Australia: Dr. Saeed Alsurf and Majed Alqahtani, it was a great blessing to have wonderful friends like them during my study. Their friendly support, generous help and their great advice were the best motivation for completing my study. I learned from them the meaning of the real friendship.

My special thanks and gratitude go to Dianne Allen whose commitment and outstanding intellectual expertise enabled this project to be in the highest quality. Her editorial support was exceptional during all phases of this project.

Finally, I could never overstate my thanks to my parents, for their love, supplications, patience and their admirable support. Also, I want to thank my great family: my loyal wife, Monerah, for her love and support and for being patient with me for being away from the house for long hours. Without her patience, support and taking care of the kids, this project would not have been completed. Many thanks go to my beloved son, Abdulullah and for my lovely two daughters, Amal and Afnan for their love and support.
TABLE OF CONTENTS

Statement of Originality ......................................................... i
Acknowledgments ....................................................................... ii
Table of Contents ....................................................................... iv
LIST OF FIGURES ........................................................................ vi
Glossary and Abbreviations ......................................................... xi

Chapter 1 .................................................................................. 1
INTRODUCTION ........................................................................... 1
1.1 Statement of the problem ......................................................... 1
1.2 Key concepts in the research ..................................................... 2
1.3 Research aim and research questions ......................................... 3
1.4 Significance of the study .......................................................... 4
1.5 Teaching English in Saudi Arabia .............................................. 7
1.6 FL aptitude research in the Saudi context ..................................... 13
1.7 Organisation of the thesis ......................................................... 14

Chapter 2 .................................................................................. 17
LITERATURE REVIEW ............................................................... 17
2.1 FL aptitude concept ................................................................. 17
2.2 Historical review of L2 aptitude testing from 1920 to the present ... 19
2.3 FL aptitude components .......................................................... 25
2.4 Foreign Language aptitude testing .............................................. 27
2.5 Various issues related to foreign language aptitude ......................... 35
2.6 Uses of the FL aptitude tests ..................................................... 47
2.7 Methodological approaches in FL aptitude research ........................ 51
2.8 Recent theoretical explanations for FL aptitude and L2 proficiency .... 55
2.9 Summary .................................................................................. 63

Chapter 3 .................................................................................. 66
METHODS .................................................................................. 66
3.0 Introduction ............................................................................ 66
3.1 Research questions and hypotheses .......................................... 66
3.2 Study design ............................................................................ 67
3.3 Study setting ............................................................................ 70
3.4 Participants ............................................................................. 72
3.5 Foreign language aptitude test instrument ................................... 75
3.6 Other instruments ................................................................. 83
3.7 Data collection procedures ...................................................... 87
3.8 Data preparation ................................................................. 90
3.9 Data processing ........................................................................ 96
3.10 Discussion of limitations ....................................................... 101
3.11 Summary ............................................................................. 102

Chapter 4 .................................................................................. 103
FINDINGS .................................................................................. 103
4.0 Overview ............................................................................. 103
4.1 Preliminary statistical analyses: Descriptive statistics ..................... 104
4.2 FL aptitude testing instrument – Test of psychometric properties ....... 109
4.3 FL aptitude predicting English proficiency .................................... 112
4.4 Development of English proficiency from seven-month intensive English course ................................................................. 126
4.5 Examining the relationship between FL aptitude measures and English progress ................................................................. 129
4.6 The FL aptitude and TOEFL tests as predictors of achievement......133
4.7 Summary of study findings ................................................................. 135

Chapter 5 .......................................................................................................... 137
DISCUSSION ........................................................................................................ 137
5.0 Overview ........................................................................................................ 137
5.1 The nature of aptitude: a stable or dynamic construct ..................... 139
5.2 FL aptitude as a predictor of English proficiency ................................ 143
5.3 Development of English proficiency ......................................................... 146
5.4 The relationship between FL aptitude measures and English progress 149
5.5 The FL aptitude and TOEFL tests as predictors of achievement......151
5.6 Subtests analyses ....................................................................................... 153
5.7 Summary ..................................................................................................... 160

Chapter 6 .......................................................................................................... 162
CONCLUSION .................................................................................................... 162
6.1 Research purpose and questions ............................................................. 162
6.2 Empirical findings .................................................................................... 163
6.3 Theoretical considerations ...................................................................... 165
6.4 Limitations of the current study and recommendations for future research in the same area ......................................................... 169
6.5 Implications for FL aptitude and SLA research ...................................... 166
6.6 Practical implications of the current study ............................................. 167
6.7 Final remarks ............................................................................................ 169
BIBLIOGRAPHY .............................................................................................. 174

Appendix 1: A summary of the previous studies in the current study .....185
Appendix 2: Foreign language aptitude test ................................................. 188
Appendix 3: A translated copy of foreign language aptitude test .......... 197
Appendix 4: The English proficiency test (adapted TOEFL) ................. 211
Appendix 5: Attrition data analyse .............................................................. 222
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Hierarchical model of cognitive abilities for FL aptitude (derived from</td>
<td>61</td>
</tr>
<tr>
<td></td>
<td>Robinson, 2002, p.119</td>
<td></td>
</tr>
<tr>
<td>2.2</td>
<td>A combined study design for FL aptitude and proficiency measured on</td>
<td>70</td>
</tr>
<tr>
<td></td>
<td>two sequential occasions</td>
<td></td>
</tr>
<tr>
<td>4.1</td>
<td>Distribution of Foreign language aptitude test scores at T1 (N=56)</td>
<td>106</td>
</tr>
<tr>
<td>4.2</td>
<td>Distribution of Foreign language aptitude test scores at T2 (N=56)</td>
<td>107</td>
</tr>
<tr>
<td>4.3</td>
<td>Distribution of English Proficiency scores at Time 1 (N=56)</td>
<td>107</td>
</tr>
<tr>
<td>4.4</td>
<td>Distribution of English Proficiency scores at Time 2 (N=56)</td>
<td>108</td>
</tr>
<tr>
<td>4.5</td>
<td>Visual representation of contributions of subtests to the foreign language</td>
<td>108</td>
</tr>
<tr>
<td></td>
<td>aptitude as represented by ranges (Minimum and Maximum)</td>
<td></td>
</tr>
<tr>
<td>4.6</td>
<td>Visual representation of contributions of subtests to English proficiency</td>
<td>109</td>
</tr>
<tr>
<td></td>
<td>as represented by ranges (Minimum and Maximum)</td>
<td></td>
</tr>
<tr>
<td>4.7</td>
<td>Mean differences comparison, diagrammatically</td>
<td>110</td>
</tr>
<tr>
<td>4.8</td>
<td>Correlation of aptitude over time, diagrammatically</td>
<td>111</td>
</tr>
<tr>
<td>4.9</td>
<td>Cross-sectional correlations between FL aptitude and English proficiency</td>
<td>113</td>
</tr>
<tr>
<td></td>
<td>(T1 and T2)</td>
<td></td>
</tr>
<tr>
<td>4.10</td>
<td>Longitudinal correlations of FL aptitude measures and English proficiency</td>
<td>122</td>
</tr>
<tr>
<td></td>
<td>measures</td>
<td></td>
</tr>
<tr>
<td>4.11</td>
<td>Mean differences for English proficiency measures</td>
<td>126</td>
</tr>
<tr>
<td>4.12</td>
<td>Correlation of distribution of English proficiency measures</td>
<td>128</td>
</tr>
<tr>
<td>5.1</td>
<td>The S-curve for learning</td>
<td>147</td>
</tr>
</tbody>
</table>
LIST OF TABLES

Table                       Page
2.1 A summary table of the cross-sectional studies of the relationships   53
   between FL aptitude and L2 proficiency at Time 1
2.2 Summary of cross-sectional relationships of FL aptitude and L2
   proficiency at Time 2
2.3 Summary of the longitudinal relationship between FL aptitude at Time 1
   and L2 proficiency at Time 2
2.4 SLA processing stages and potential components (Skehan, 2002: p. 90)
3.1 MLAT and Foreign language aptitude test tasks
3.2 Correlations of Foreign Language Aptitude components (N=10)
3.3 Summary of the components of the Foreign Language Aptitude test
3.4 Summary of the components of the English proficiency test
3.5 Summary of data-collection procedures
3.6 Outliers for all study variables on both occasions, Time 1 and Time 2,
   for (N=56)
4.1 Descriptive statistics of the Foreign Language Aptitude scores and
   English proficiency scores at T1 and T2 (N=56)
4.2 Mean differences for aptitude measures before and after the seven-
   month course (N=56)
4.3 Correlations of aptitude measures between T1 and T2 measurements
   (N = 56)
4.4 Zero-order correlations between FL aptitude components with English
   proficiency subtests and total score across the whole sample at Time 1
   (N = 56)
4.5 Multiple regression between aptitude components at T1 and English
   proficiency total scores at T1 (N= 56)
4.6 Zero–order correlations between Foreign Language Aptitude
   components with English proficiency subtests for the whole sample
   (N=56) in T2
4.7 Multiple regression between Aptitude components at Time 2 and
   English proficiency total scores at Time 2 (N= 56)
4.8 Zero-order correlations of FL aptitude Components at T1 with English
   proficiency Subtests at T2 for the whole sample (N=56)
4.9 Multiple regression between Foreign Language Aptitude components at
   Time 1 and English proficiency total scores at Time 2 (N= 56)
4.10 Mean differences for English proficiency measures before and after the
    seven-month course (N=56)
4.11 Correlations of English proficiency measures between T1 and T2
    measurements (N = 56)
4.12 Correlations between Foreign Language Aptitude and English
    proficiency progress by using the absolute value formula, (N = 56)
4.13 Correlations between Foreign Language Aptitude variables at Time 1
    and the ratio of English progress along the seven-month English course
    (N=56)
Table

4.14 Correlations between Foreign Language Aptitude variables and the ratio of English progress along the seven-month English course as per the third formula (N=56) ........................................................................................................... 132
4.15 Multiple regression between the two main variables at Time 1 and English proficiency at Time 2 (N= 56) ..................................................................................................................... 133
4.16 Multiple regression between all variables at Time 1 and English proficiency at Time 2 (N= 56) 134
5.1 Mean differences for aptitude total and its subtest components and English proficiency and its subtest components (N=56) ........................................ 154
5.2 Correlation coefficients (r) of regressions between aptitude and its subtests and English proficiency and its subtests at Time 1 and Time 2 (N=56) .................................................................................................................................. 155
ABSTRACT

The present study aims to explore whether or not there is a relationship between two multifaceted variables: foreign language aptitude and second language proficiency. 56 male students at King Khalid University in Saudi Arabia participated in the current study. The researcher examined how well the participants’ score in a foreign language aptitude test predicted their success in a foreign language proficiency test (TOEFL) before and after a seven-month English course. In contrast to traditional studies in the field, this study uses not only the cross-sectional design, but also the longitudinal one to answer its questions. This combined design provides more opportunities to triangulate not only the findings of the past studies, but also to examine various hypotheses which were set by previous studies without empirical testing. The innovative design of this study allows the researcher to explore the relationship between foreign language aptitude and second language proficiency in greater depth than has been possible in previous research.

The study addresses a range of issues in FL research concerning the concept of FL aptitude which have been the subject of ongoing debate, such as the nature of FL aptitude, the role of FL aptitude in predicting L2 proficiency, the relationship between FL aptitude and English progress. The current study involved piloting the study’s measures and the initial adaptation of the foreign language aptitude test for Arab native speakers, followed by two phases of data collection (Time 1 and Time 2).
The data gathered during the experimental phase of the project have been subjected to a range of different statistical procedures yielding the following main findings:

1) Foreign language aptitude as measured in the study’s population sample manifested itself as a stable trait.

2) There was a positive significant relationship between foreign language aptitude and English proficiency cross-sectionally and longitudinally.

3) There was a statistically significant change in the levels of English proficiency among the participants, but this change was relatively small.

4) There was no significant relationship between the foreign language aptitude test and English progress.

The current study can be credited as contributing to the field of applied linguistics in general and to the FL aptitude research in Saudi Arabia in particular. Among its most significant contributions are the development of a foreign language aptitude test for Arab native speakers, as well as the use of the cross-sectional and its longitudinal design which has enabled to examine the extent to which foreign language aptitude is correlated with the progress in the attainment of the target language. What is also unique to this study is the ability to examine the interactions between all of the subcategories of the foreign language aptitude test and all the subcategories of the English proficiency test in a cross-sectional and longitudinal design. Such an examination built a detailed picture of how, and to what degree foreign language aptitude can predict English proficiency among Saudi students in the target sample. The findings have significant implications for language teaching and learning practices, as well as for theories of Second Language Acquisition.
### Glossary and Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description/ Glossary</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAT</td>
<td>Arabic aptitude test</td>
</tr>
<tr>
<td></td>
<td>SPECIAL NOTE: Throughout this thesis the term ‘Arabic aptitude test’ has been used to refer to the adaptation of MLAT for Arabic native speakers; It is an instrument for measuring the foreign language aptitude of Arabic native speakers (see details at Chapter 3.5).</td>
</tr>
<tr>
<td>BA</td>
<td>Bachelor of Arts</td>
</tr>
<tr>
<td>CANAL-FT</td>
<td>Cognitive Ability for Novelty in Acquisition of Foreign Language</td>
</tr>
<tr>
<td>EFL</td>
<td>English as a Foreign Language – English being learned as a Foreign Language where there is little or no access to English in the community at large of the learner.</td>
</tr>
<tr>
<td>EPT</td>
<td>English proficiency test</td>
</tr>
<tr>
<td>ESL</td>
<td>English as a Second Language – English being learned by student with another L1, within a community where English is the general L1</td>
</tr>
<tr>
<td>FL</td>
<td>Foreign language. In this study, this term refers to any language other than the native language (L1).</td>
</tr>
<tr>
<td>FSI</td>
<td>Foreign Service Institute</td>
</tr>
<tr>
<td>GPAs</td>
<td>Grade Point Averages</td>
</tr>
<tr>
<td>GS</td>
<td>Grammar Sensitivity (measure in AAT)</td>
</tr>
<tr>
<td>HUNLAT</td>
<td>The Standard Hungarian Language Aptitude Test</td>
</tr>
<tr>
<td>Ids</td>
<td>Individual Differences</td>
</tr>
<tr>
<td>L1</td>
<td>First Language or Native Language</td>
</tr>
<tr>
<td>L2</td>
<td>Second Language. It refers to any language learned after native language (L1)</td>
</tr>
<tr>
<td>LABJ</td>
<td>Language Aptitude Battery for Japanese</td>
</tr>
<tr>
<td>LAT</td>
<td>Language Aptitude Test</td>
</tr>
<tr>
<td>LCDH</td>
<td>Linguistics Coding Differences Hypothesis</td>
</tr>
<tr>
<td>LLAMA</td>
<td>Swansea Language Aptitude Test (Meara, 2005).</td>
</tr>
<tr>
<td>MCS</td>
<td>Memory for Contingent Speech</td>
</tr>
<tr>
<td>MCT</td>
<td>Memory for Contingent Text</td>
</tr>
<tr>
<td>MD</td>
<td>Mean Difference</td>
</tr>
<tr>
<td>MEM</td>
<td>Memory (measure in AAT)</td>
</tr>
<tr>
<td>MLAT</td>
<td>Modern Language Aptitude Test – Carroll and Sapon’s 1959 instrument in its most recent form, 2002.</td>
</tr>
<tr>
<td>P</td>
<td>Significance</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description/ Glossary</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------------------</td>
</tr>
<tr>
<td>PLAB</td>
<td>Pimsleur Language Aptitude Battery – Pimsleur’s (1966) instrument</td>
</tr>
<tr>
<td>PMA</td>
<td>The Primary Mental Abilities Test</td>
</tr>
<tr>
<td>$R$</td>
<td>Co-efficient of correlation, (Pearson’s)</td>
</tr>
<tr>
<td>S-curve</td>
<td>Learning curve (Jovanovic and Lach, 1989)</td>
</tr>
<tr>
<td>SD</td>
<td>Standard Deviation</td>
</tr>
<tr>
<td>SLA</td>
<td>Second Language Acquisition</td>
</tr>
<tr>
<td>SPSS</td>
<td>Statistical Package for the Social Sciences Version 17, 2008</td>
</tr>
<tr>
<td>$T_0$</td>
<td>Time before any L2 learning</td>
</tr>
<tr>
<td>$T_1$</td>
<td>Time 1 (before start of seven-month intensive English course)</td>
</tr>
<tr>
<td>$T_2$</td>
<td>Time 2 (after completion of seven-month intensive English course)</td>
</tr>
<tr>
<td>TOEFL</td>
<td>Test of English as a Foreign Language – measures the ability of individuals to use and understand English in an academic setting.</td>
</tr>
<tr>
<td>TUNJO</td>
<td>Test of Aptitude for the Learning of Foreign Languages (<em>Test Uzdolnień do 8auki Języków Obcych – TUNJO</em>)</td>
</tr>
<tr>
<td>VORD</td>
<td>Aptitude battery by Parry and Child (1989)</td>
</tr>
<tr>
<td>VWC</td>
<td>Vocabulary and Written Coding (measure in foreign language aptitude test)</td>
</tr>
</tbody>
</table>
Chapter 1

INTRODUCTION

1.1 Statement of the problem

It is clear that there has been intense focus, in the field of second language acquisition research, on identifying the common or universal approach or methods to acquire a second language (L2). These studies attempt to identify a common sequence or process that all learners follow to acquire a new language, such as: the same developmental sequence of L2 learning, the linguistic transfer from a person’s native language (L1) to L2, overgeneralisation, fossilisation, etc. (Skehan, 1991). In contrast, the study of individual differences especially those related to cognitive factors (e.g. aptitude) among language learners has not attracted enough theoretical or empirical interest. It is difficult to ignore the clear variations that have been observed, of the impact of learners’ characteristics on second language acquisition (SLA) and more generally on the language learning process. It is time, then, to re-address this relatively neglected area of individual differences (IDs).

The individual factors related to second language learning can be divided into various categories: affective factors (e.g., motivation, attitude, and personality) and cognitive factors (e.g., intelligence, aptitude). A considerable volume of IDs research has examined the impact of affective factors. Far less research has investigated the impact of cognitive variables on second language acquisition. Among these cognitive variables is FL aptitude which has a consistent high predictability of L2 success compared to other variables
Given this, a major aim of this study is to investigate the relationship between foreign language (FL) aptitude and second language proficiency. The investigation involves developing a foreign language aptitude test for Arabic native speakers, adapted from the Carroll Modern Language Aptitude Test (MLAT). The FL aptitude-L2 proficiency relationship is then tested, in both a longitudinal and a cross-sectional design, among Saudi students in their first-year of university studies.

This study also aims to shed some light on the role of FL aptitude in second language success in a distinctive context (English as a FL), since very little research has been conducted to investigate the impact of FL aptitude on learners’ success in English as a FL. Further, none of the previous studies that have been conducted in various cultural contexts have considered the Saudi context. That is to say that this Saudi study, which broadens EFL context studies, also broadens the studies that have been conducted of aptitude. Comparing the findings of this study with other previous studies, in different learning contexts, will be part of that process.

1.2 Key concepts in the research

*Foreign language aptitude*

The concept ‘language aptitude’ has various names such as ‘talent’, ‘knack’. The term ‘foreign’ with ‘language aptitude’ applies whether the language being learned is ‘foreign’, ‘second’, ‘third’, etc. It is making a distinction between the first and native/parental language and any other language. It applies most accurately when the other language is
not common to the general community surrounding the language learner. In bilingual or multi-lingual communities different terminology needs to be used.

\textit{FL vs. L2 learning context}

Both foreign language (FL) and second language (L2) contexts refer to the L2 learning contexts. For this study the foreign language being learned is English, and it is being learned in a context where Arabic is the predominant native language (L1), and learned by attendance at classes designed to teach English as a foreign language.

\textbf{1.3 Research aim and research questions}

The main aim of this study is to shed some light on the role of FL aptitude in second language success in a distinctive context (English as a FL). Not much research has been conducted to investigate the impact of FL aptitude on learners’ success in English as a FL.

More specifically, this study intends to fill part of that gap, and by investigating the relationship between the foreign language aptitude test and the English proficiency in King Khalid University in Saudi Arabia. Furthermore, this study also endeavors to enrich Saudi educators and researchers with more ideas about the effect of FL aptitude on language success in their universities, which in turn will have a positive influence on English learning in Saudi Arabia.

This study aims to assess whether the foreign language aptitude test instrument predicts English proficiency during seven-months of L2 English learning in the class, and by using a cross-sectional and longitudinal design that provides a firm basis for confident predictive inferences.
The main goal of the current study is achieved by answering the following research questions:

1. Is the foreign language aptitude testing instrument in the current study a valid test, at a level similar to other equivalent instruments?
2. Do the foreign language aptitude test scores at the beginning of the English course (Time 1) differ from the foreign language aptitude test scores at the end of the seven-month English course (Time 2)?
3. Does the foreign language aptitude test predict English-language learning proficiency to a similar standard as other aptitude tests?
4. Do the foreign language aptitude test scores correlate significantly and positively with English proficiency progress after a seven-month intensive English course?

1.4 Significance of the study

This study is significant to the field of FL aptitude research because it addresses a number of issues: (1) the absence of an Arabic aptitude instrument; (2) adding longitudinal design to the usual cross-sectional correlation studies; (3) endeavouring to compare L2 progress with aptitude differences; (4) the specifics of the learning context in which L2 proficiency is being assessed; (5) adding to the research of individual differences; (6) exploring the educational implications of FL aptitude measures.

Adaptation of the foreign language aptitude test from MLAT

Most Arab linguists have not been interested to investigate the impact of FL aptitude in L2 proficiency. The little interest in FL aptitude among the Arab native speakers has been due to the lack of a valid, reliable and standard foreign language aptitude test. This study provides a lead in ongoing Arabic language learning research by developing a foreign
language aptitude test for native speakers of Arabic. The new foreign language aptitude test was adapted from the Modern Language Aptitude Test (MLAT) developed by Carroll and Sapon (1959, 2002) for native English speakers. The MLAT is used in many academic articles and still widely used in many universities and institutions as the most valid measure for FL aptitude.

**Design of the study**

Most FL studies have been mono-designed to investigate the relationship between FL aptitude and L2 proficiency, either cross-sectionally or longitudinally. This study has been designed to be a combination of the two designs, cross-sectional and longitudinal. This way of design allows the researcher to compare the results of the two designs for the same sample.

**English Progress**

Most previous studies have based their findings on the relationship between the total score of FL aptitude test and the total score of L2 proficiency test. Few studies have considered the relationship between FL aptitude and L2 progress, and those studies that have considered this relationship could be criticised for calculating the progress by the difference between the two measures of L2: before and after the course. The current study is novel in investigating the relationship between FL aptitude and the ratio of progress instead of using the differences between the two proficiency measures.
Learning contexts

The past 15 years have witnessed an increased interest in the role of FL aptitude in SLA. Few studies have investigated the relationship between FL aptitude and L2 success in a context where English is the FL. This study aims to investigate the impact of FL aptitude in English proficiency among Saudi students in King Khalid University. The results of this study should therefore make some contribution to the field of second language acquisition by re-examining FL aptitude in a different language context – Arabic. The variations in findings of FL aptitude studies in FL/L2 contexts will contribute to a wider understanding of the process of second language acquisition and also expand our knowledge of the impact of aptitude on learners’ L2 success in various contexts.

The lack of FL aptitude research

In the last few decades, a considerable amount of research has investigated the impact of various individual differences (IDs), such as age, motivation, personality, on L2 success. In contrast, far too little attention has been paid to FL aptitude. In this study, I examine the impact of FL aptitude on learners’ language success among the target population.

Educational Implications

Any better understanding of the role of the FL aptitude in SLA, developed in the current study would have implications for the educators and the course designers. The FL aptitude test scores could be used as a diagnostic tool. Variations of the aptitude scores could be taken into account in using new teaching methods or in designing new materials to be more appropriate for learners’ needs.
1.5 Teaching English in Saudi Arabia

A considerable amount of research has explored the impact of L2 learning context in SLA, and in aptitude-L2 proficiency relationships. In the following sections, detailed descriptions of the status of English in Saudi Arabia are provided to allow the reader to have more understanding of the Saudi context and to offer the detail necessary from which explanations of the findings of the current study may need to be drawn.

The focus will be on the educational and cultural aspects that provide the context for learning EFL, in Saudi Arabia. Later, detail concerning the FL aptitude research context will be presented.

1.5.1 Saudi Arabian English context

Saudi Arabia is located in the Southwest of Asia. It is the biggest country in the Middle East. It occupies more than two thirds of the Arabia Peninsula. Its area is two million and quarter square kilometres. Two long coasts extend on its borders; the Red Sea along the western coast and the Arabia Gulf along the eastern border. Saudi Arabia is surrounded by various Arab countries. Some gulf countries such as United Arab Emirate, Bahrain and Qatar are to the west; Yemen and Oman are to the South; Iraq, Jordan, Syria are to the North. According to the last statistics (2008), the Saudi population is approximately 23 million citizens and about 7 million non-citizens. This seven million are from different parts of the world. The majority are from non-English speaking countries such as India, Pakistan, and Philippines, and from different cultural, linguistic and intellectual backgrounds. Most do not know how to communicate in Arabic as the official language in the country. As a result, English is used as a lingua franca among Saudis and this multi-lingual, ethnically diverse community as a tool of communication. As Traynham (2006)
has noted, English, as a lingua franca, has advantages and disadvantages. It creates a good atmosphere for Saudis to practise their English with this multi-lingual people, but, unfortunately most English-speakers from outside in the country are non-native speakers of English and their English is poor. In the cases of belief and the formal language, Islam is the religion of the country and Arabic is the mother tongue of about 98% of the country. Saudi society is well known for its loyalty to Arabic as the language of religion, culture, media and education.

The interest in introducing English in Saudi Arabia started after the establishment of the Kingdom of Saudi Arabia in 1932 by King Abdul-Aziz Al-Saud. The King had realised the significance of English as a means of communication with other countries around the world and with people from outside, visiting or working in Saudi Arabia. For instance, after the discovery of oil, many English teaching programs were introduced by the Saudi government to equip their citizens with English to enable them to communicate with experts and workers from outside (Alam, 1986). In another example, about two million non-Arabic Muslims from all over the world come to Makkah (a holy city in the west of Saudi Arabia) annually to perform Hajj (pilgrimage). Therefore, there is a necessity to teach English, in public education and in various government sectors, to facilitate communication with those people from all over the world (Al-Seghayer, 2005). These needs crystallised in the efforts of the Education Ministry to introduce teaching English as a compulsory subject in the Saudi curriculum in 1927 in all public schools around the country.
1.5.2 Status of English in educational setting

In public schools English is taught in intermediate and high schools, and recently it has been introduced gradually to primary schools since 2003.

English is widely used in many universities, schools, hospitals and in most of the government organisations, but Arabic is still the official language strongly associated with cultural and intellectual dimensions.

Teaching English in public schools was initially established in 1927 (Jan, 1984). At present Saudi students in public education study English for eight years, starting from the 5th grade in primary school until Year 12th, the last year in high school. After finishing high school, a good level of English proficiency is required for entrance to tertiary level studies in most majors such as medicine, engineering, medical sciences or to join any technical or vocational colleges around the country. Furthermore, most Saudis who are interested in working in banks, companies, business or industry should exhibit a good level of English fluency and accuracy.

English is taught in private schools in Saudi Arabia. These schools are owned by private parties. Curriculum, textbooks, teaching materials, programs are supervised by the Saudi Ministry of Education and Learning. In these private schools, students have a greater chance of exposure to English than students in public schools.

English instruction in private schools begins from the first year in primary school or in the 4th grade in some private schools. The greater length of exposure to English seems to advantage these students more than their counterparts in public schools. But this kind of schooling is very expensive for the population. It is limited to rich families or those who have good positions in the government. The parents’ aim for sending their children to
private schools is to improve their children’s level of English proficiency by exposing them to more English instruction at an early age (Zaid, 1993).

In tertiary education, English is used as the language of instruction in many colleges and universities in Saudi Arabia. For instance, in King Fahd University of Petroleum and Minerals (KFUPM), a good level of English proficiency is required to enrol in all the university’s colleges and departments. In most medical, technical and industrial colleges around the country, all courses are conducted in English. In Arabic and Islamic colleges where English is not the language of instruction, introductory English courses are required in the foundation year. In English departments in Education colleges, Art colleges, and Languages and Translation colleges, English is taught as a core of instruction for four years. All English courses such as morphology, syntax, semantic, phonetics, teaching methods and language acquisition theories are taught as core subjects for students majoring in English to fulfil the requirements of BA in English. After graduation, most graduated students mainly work as an English teacher in public or private schools. Those who are not interested in teaching may work as a translator in hospitals, companies or in any other government sectors. Graduated students with high grades in BA are chosen to teach in universities or colleges as teaching assistants.

The role of government in introducing English can also be noted through the mass media. One of the Saudi local television channels is English Channel 2 which broadcasts all program and news in English. Furthermore, three of the daily newspapers are published in English, the Arab News, the Daily Riyadh, and the Saudi Gazette (Al-Seghayer, 2005).
1.5.3 English curriculum

Saudi Public education consists of three levels: elementary for six years (Grade 1 to Grade 6), intermediate for three years (Grade 7 to Grade 9) and secondary for three years (Grade 10 to Grade 12). English is taught, starting from Grade 5 up to 12 for eight years as the only foreign language in the Saudi public curriculum. The textbooks and the teaching materials were developed by the Ministry of Education to be the same for boys and girls. All English teachers in every grade are required to cover the same syllabus according to predetermined guidelines and deadlines (Al-Seghayer, 2005). Class numbers and workloads, from this researcher’s experience as a high school teacher, mean that time to find and provide extra materials to suit particular student learning needs is very tight, and the standardised curriculum and assessment mean that the development of diverse support materials is not encouraged, and their availability is limited.

English, in Saudi Arabia, is taught for its functional use in some situations. One of those uses is as the language of instruction in different stages in the Saudi educational system. It is limited in its use outside the classroom (Al-Seghayer, 2005). As Nunan and Carter (2001) have noted in countries where English is a foreign language, there is not much exposure to English and less opportunity to use it in daily life.

1.5.4 English teaching methods in Saudi Arabia

Teaching English in Saudi Arabia is mainly conducted using two classic teaching methods: Audio-lingual teaching method and Grammar translation method. The former is the most common teaching method in spite of the lack of language laboratories which are necessary to apply this method (Zaid, 1993). Teaching grammar and vocabulary by memorisation and repetition are common features of Saudi classrooms (Al-Ahaydib, 1986). It is noted that most English teachers use Arabic instead of English and rely on
translation inside classrooms (Abu-Ghararah, 1990; Almalki, 2008). In other words, English as the ‘language of instruction’ is an idea, but not the actual practice, and this applies as much to English lessons as it does to lessons in other subjects.

1.5.5 Reasons behind the lack of development of English proficiency in the Saudi educational system

One of the major problems facing the policy makers and the curriculum designers in Saudi Arabia is the noticeable lack of English proficiency among high school graduates. Zaid, in his comprehensive analysis of the Saudi Curriculum, noted that the main reason behind the English weakness among Saudi students is related to focusing on the content of language itself not the use of language as a communication vehicle (Zaid, 1993). To solve this problem, the Ministry of Education decided to introduce English to primary schools as a core subject starting from 2003. This decision aims to increase the length of exposure to English to include Grades 4, 5 and 6. The decision makers believe that increasing English teaching exposure (3 years more) to younger learners could improve English in Saudi Arabia. Unfortunately, this decision is not strongly supported by many empirical studies in the field of second language acquisition, especially in countries where English is a foreign language such as Saudi Arabia (Alshahrani, 2007). Also, the implementation of this decision created a new problem which is the need to recruit more English teachers from abroad.

The researcher, as a high school teacher for seven years in Saudi Arabia, agrees with Zaid’s (1993) observations: most Saudi students have some misconceptions about the significance of learning English. Students generally believe that there is no necessity for learning English. They claim that after finishing schooling they will join universities or obtain jobs where English is not required. Also, they don’t pay serious attention to
English as a tool of communication in their daily life. Instead, they deal with it as any subject in the curriculum: one that has to be studied, and sometimes just enough to achieve a pass mark. Based on these misconceptions, students just memorise the grammatical rules, written passages and some vocabulary to get the required score to pass the exam.

1.6 FL aptitude research in the Saudi context

After reviewing the many publications and academic dissertations concerning English in Saudi Arabia, the researcher noted that most of previous research has been devoted to the study of the English curriculum and various practices of teaching English in Saudi Arabia as a example of teaching English in a foreign context (Alkeaid 2004; Zaid, 1993; Almalki 2008). A small number of studies have been devoted to studying the impact of individual differences (IDs) in SLA: motivation and attitude (Al-Ajmi, 2001; AlMaiman, 2005; Semmar, 2005), learning strategies (Al-Otaibi, 2004; AL Qahtani, 2005). But no single study has investigated the impact of FL aptitude on English proficiency in the Saudi context where English is a FL.

The lack of interest in FL Aptitude research among Saudi Applied Linguists was a result of the lack of a valid, reliable foreign language aptitude test. Therefore, the researcher in the current study decided to develop a FL aptitude test for Arabic native speakers. The new test was adapted from the most well-known and widely used FL aptitude test, The Modern Language Aptitude Test (MLAT) which was designed by Carroll and Sapon (1959, 2002). This test has been used in many articles in various well-known journals in the field of applied linguistics. Furthermore, it is considered to be the most valid and reliable FL aptitude test. Many educational and institutions use this test as a tool of
selection and for diagnostic purposes. More information about the adaptation of MLAT and developing of the foreign language aptitude test will be provided in Chapter 3.

From the above, the researcher decided to develop the new adapted foreign language aptitude test used in this study. It fills a significant gap in current research in the Arabic language context, something that will have a wider application than only the Saudi Arabian system. The development and testing of this instrument will support Saudi and Arabic language applied linguists and educators. Also, it will be used with Arabic native speakers to measure their aptitude to learn a new language.

1.7 Organisation of the thesis

This thesis has been divided into six chapters. The first chapter introduces the main problem of this research that many studies have been conducted to investigate various individual differences among FL learners but the role of FL aptitude has been neglected. This study aims to examine the impact of FL aptitude on the English competences of Saudi territory students. Because of the importance of L2 learning context in SLA, this chapter provides a detailed description of English in Saudi Arabia. Also, this chapter addresses the practical advantages of this study for the language teachers and for FL aptitude researchers.

Chapter 2

After reviewing the main issues related to the main aims and practical and contextual significance of the present study, various theoretical and experimental issues related to FL aptitude are presented in Chapter Two. These issues relate to various topics, including the definition and the conceptualisation of FL aptitude, the history of aptitude testing and the basic components of the measuring instruments such as MLAT and PLAB and
methodological approaches used in previous studies, and the aim is to provide a historical and theoretical background for FL aptitude research up to the date of this study. It is important to have a clear understanding of FL aptitude before presenting the methodological procedure related to this study.

Chapter 3
A comprehensive description of all issues in relation to the research protocol deployed in this project will be laid. In order to accomplish this, the main dimensions of the methodology such as the general design of the study, participant recruitment procedure as well as details about the demographics of the sample, research instruments and data collection procedures are considered. The chapter concludes with a brief statement of the study’s limitations. These practical procedures were then conducted to collect the related data from the target sample. All the data collected were then prepared for the statistical analyses in the next chapter.

Chapter 4
Chapter Four reports the findings of the testing, starting with preliminary descriptive statistics for the selected sample, and then proceeding to report the material relevant to four research steps: foreign language aptitude testing instrument – its empirical testing; FL aptitude predicting English proficiency; Development of English proficiency from a seven-month intensive English course; any relationship between FL aptitude measures and English progress.

Chapter 5
While the facts (empirical findings) have been presented in Chapter 4, the question answered in this chapter addresses what these facts mean and imply for the field. To
determine the relevance of the findings to the other aptitude researchers and theorists, the findings need to be evaluated and interpreted. One such evaluation is a comparison with other relevant research findings. For what these findings mean, explanations need to be developed, based on either practical or empirical contexts, or related to explanatory theories.

Chapter six

This chapter begins by reporting the main findings of the present study. Then, it sheds some light of the limitations of this study combined with the resultant recommendations for future studies. At the end of the chapter, more theoretical implications of these findings are provided for others in FL aptitude research and some pedagogical implications are directed to those applied linguists who work in the Saudi context.
Chapter 2  
LITERATURE REVIEW

2.1 FL aptitude concept

It is recognised that people in general vary in their abilities in daily life; it is impossible to find two persons with exactly the same abilities. The first attempt to test the ability of human beings occurred 4000 years ago, when the Chinese designed the first ability tests to select candidates to work in government sectors or to enrol in educational programs (Cooper, 2002). For a long time, and for the same purposes, intelligence tests, as the main example of cognitive ability testing, were associated with learning success in general. By 1905, the best known of these, the Binet-Simon Intelligence Scale, was established as the first generally used intelligence test. It aimed to identify students who did not have the intellectual ability to study in public schools. During the 1900s, most studies of individual differences associated intelligence with language-learning success (Dörnyei, 2005).

In the 1920s and 1930s, the concept of foreign language (FL) aptitude was separated out, in order to distinguish this aspect of language-learning from the main stream of general learning ability. In a later, pioneering work in FL testing, Carroll proposed that “specialized abilities beyond general intelligence play an important role in learning a foreign language” (Carroll, 1981, p. 27).

In the second-language (L2) learning domain, FL aptitude can be defined as a set of relatively fixed characteristics or talents not affected by previous learning experience. These talents enable some students with high talent or aptitude to master a new language faster and easier than less-talented students (Carroll & Sapon, 1959; Gardner &
MacIntyre, 1992). Based on the rate and the ease of learning, it is noted that individuals with good aptitude abilities learn foreign languages faster and more easily than learners with poor aptitude abilities (Carroll, 1981). In accordance with Carroll’s definition, Skehan (2002) defines aptitude as “the ability to learn a new language quickly and to a high degree of proficiency” (p. 119). Skehan insists that Carroll’s conceptualisation of aptitude should be clearly understood as the learner’s capacity for learning a new language with a sufficient level of proficiency in the limited time of a language program. He adds that many researchers such as Stansfield and Reed (2004) believe that any learner has the ability to master a new language but that learners show discrepancies in the rate and the ease of FL learning abilities. It is clear that Carroll’s aptitude concept does not directly aim to diagnose learning disabilities. This was emphasised by Carroll himself in an interview when he was asked about the relationship between L2 aptitude and learning disabilities (Skehan, 1998).

In summary, FL aptitude conceptualisation is based on certain assumptions: 1) FL aptitude is relatively stable and is not affected by previous learning experience (Carroll, 1981). 2) FL aptitude is independent of general cognitive ability such as intelligence (Carroll & Sapon, 1959). 3) FL aptitude is independent of other non-cognitive factors such as motivation, personality and learning strategies (Carroll, 1981). 4) FL aptitude varies between learners (Skehan, 1998).

The issues related to the relation between aptitude stability and the learner’s proficiency in their primary language (L1) as well as the relation between FL aptitude and intelligence have been heatedly contested issues in foreign language research for the last few decades.
2.2 Historical review of L2 aptitude testing from 1920 to the present

Having reviewed the definition of foreign language aptitude, it is crucial to provide some historical background; it here extends back to the 1920s and 1930s, the time of the emergence of FL aptitude tests. The increasing interest in developing foreign language aptitude measures stretches to the twenty-first century. In this part of the study, the researcher will discuss, in chronological order, the recognised periods of aptitude-testing development and the variety of FL aptitude tests associated with those periods. Later, more detail will be provided and varieties of FL aptitude tests; more details will be provided for the best-known aptitude test: the modern language aptitude test (MLAT) since it is the measure of FL aptitude that will be used in this study.

2.2.1 The beginning of L2 aptitude testing (1920s-1930s)

The 1920s and 1930s witnessed the first unsuccessful attempts to develop foreign language aptitude tests in the United States. A number of American schools and universities developed aptitude tests called Prognosis tests. These were designed for various purposes: to diagnose learners’ weakness in acquiring foreign languages, to identify untalented foreign language learners enrolling in FL programmes in government schools, and to economise on foreign-language learning costs (Dörnyei, 2006). As mentioned in Sparks and Ganschow (2001), in 1925, Stoddard and Vander Beke developed a test to measure ability for English grammar and Esperanto lexis learning. Three years later, Laria and Orleans designed a new prognosis test based on vocabulary tasks and grammar translation for two languages (Spanish and French). Another example of these tests was constructed by Symonds (Kaulfers, 1931). In his test, he hypothesised that the ability to learn L1, intelligence, and “quick learning” were three significant types of aptitude. Kaulfers (1931) argued that the prognosis tests of the 1920s and 1930s were
relatively poor predictors of foreign-language learning. He proposed that English grades and intelligence scores were better predictors than prognosis tests. He also suggested that different learning contexts could affect FL learning (Kaufers, 1939).

These ineffective prognosis tests led to more aptitude-testing research. Before World War II, linguists and psychologists from the US Army attempted to devise an aptitude battery for predicting success in learning foreign languages. The best method at that time for selecting learners for foreign language training was to offer a short trial course. After finishing the trial, learners would be tested. This method was effective with those students with high aptitude abilities, but not for the majority of the participants in the trial. This trial method was also very costly for the military and so the need for more adequate aptitude measures was recognised (Carroll, 1981).

2.2.2 Aptitude tests during and immediately after World War II

During World War II, the necessity for mastering foreign languages increased rapidly in the military. As a result, many language-training programs were introduced. At that time the selection criteria for the intensive language programs were based on the previous education of the applicants (Sparks & Ganschow, 2001). In 1953, the US government funded a project for FL aptitude testing at the University of California, Los Angeles (UCLA). It was a paper/pencil aptitude test that measured ten aptitude skills. Unfortunately this project was also unsuccessful in predicting learners’ achievements in learning a foreign language (Carroll, 1993). These unsatisfactory results, however, did lead to more aptitude testing research.
2.2.3 Aptitude testing in Carroll’s era (1950s-1960s)

The main aim of the initial attempts at developing FL aptitude testing in the 1920s-1930s and 30 years later, was to minimise the cost of language education. In the 1950s and 1960s, a new scientific trend emerged in aptitude-testing research in the USA. This period is considered to be the golden era in the history of FL aptitude testing. In that time, many empirical studies were conducted and a number of psychometric aptitude tests were developed. As a result of these efforts, two of the most popular scientific aptitude tests were developed: the Modern Language Aptitude Test, (MLAT, by Carroll and Sapon, 1959) and the Pimsleur Language Aptitude Test (PLAB, Pimsleur, 1966). These two batteries were designed to measure learners’ aptitude abilities to master a foreign language from age fourteen and up, in a classroom setting (Carroll, 1981; Rees, 2000). Other tests were designed later, but these two batteries are the most widely used tests in most of aptitude-testing research because of their predictive power. Thus, these two batteries will be discussed in more detail later.

2.2.4 Aptitude research after Carroll’s era (1960s -1990s)

Between the 1960s and the late 1990s, aptitude research did not make any major breakthroughs. Carroll (1990) states:

Since 1959, the publication date of the MLAT, there has been considerable research that throws light on the components of foreign language aptitude and that provides information that might be useful in revising this and other batteries of foreign language aptitude test. For the most part, this research has not suggested any major change in the components of foreign language aptitude that have been recognised from the start. (p. 14)

There is common agreement that there is a noticeable lack of interest in FL aptitude research in this era, except for several aptitude tests which were developed after MLAT and PLAB. These tests were designed for the same traditional purposes, such as
predicting foreign language success (Ehrman, 1998) and selecting candidates to enrol in
government or language learning program. Instances are: the Defense Language Aptitude
Battery (Peterson & Al-Haik, 1976), and VORD (Parry & Child, 1990). None of these
tests showed any kind of any superiority above MLAT (Dörnyei, 2005). In VORD, Parry
and Child compared the predictive validity of various aptitude tests with their newly
developed battery which measured only grammatical ability. They found that MLAT was
“the best overall instrument for predicting language-learning success” (p. 52). Almost all
FL aptitude tests developed in this period show no more predictive power than MLAT in
L2 learning achievement (Dörnyei, 2005).

After the golden Carrollian period until the 1990s, aptitude has received relatively little
attention, except for some studies and articles discussing the relationships between
aptitude, intelligence and individual differences such as motivation, personality and age
(Diller, 1981; Genesee & Hamayan, 1980). The best description of this period can be seen
in Carroll’s summary in 1990 of the period from the start of MLAT in 1959 to 1990.

Some researchers have suggested various reasons for the lack of aptitude research in the
last few decades. They have claimed that aptitude measurements appear to be out of date
and are no longer good predictors of L2 success. Skehan (2002) also suggests that
aptitude has only been designed to be a good predictor with specific language teaching
methodology (i.e., audio-lingual) and for being inadequate in a communicative approach.
Furthermore, some scholars such as Krashen (1981) argue that FL aptitude is associated
with a formal learning context and it appears to be useless in predicting the language
success of individuals in a naturalistic environment.

Other criticisms are related to language teacher training. Some researchers believe that
aptitude is a non-dynamic ability. Also, they take the view that it is almost impossible to
overcome the learning difficulties among learners with low aptitude. Many foreign language teachers generally ignore individual differences among learners. Despite the common belief that all learners vary in their abilities, they ignore individual differences because of a lack of language teaching materials designed to address these different needs. As a result, there has been a mismatch between learners’ varying needs and the uniformity of FL pedagogy in terms of content and methodology.

Although some recent studies have abandoned using the term ‘language aptitude’ and its tests (Ellis, 2001; Robinson, 2003), the concept is still widely used in most research because of the strong predictive ability of FL aptitude measurements such as MLAT in relation to foreign language learning success (Dörnyei, 2006). The previous criticisms will be discussed further and more studies will be mentioned for or against these issues.

2.2.5 L2 aptitude research since 2000

The last two decades have witnessed a remarkable resurgence of interest in FL aptitude research. The main shift started at the beginning of the 1990s following a significant publication entitled “Language Aptitude Reconsidered”, edited by Parry and Stansfield (1990).

In the 1990s, there were only a handful of studies on aptitude (Skehan, 1991, 1998; Sasaki, 1993; Harley & Hart, 1997) whereas in the decade after 2000 there was a considerable increase in the number of studies (Rees, 2000; Sawyer & Ranta, 2001; Ranta, 2002; Grigorenko, 2002; Dörnyei & Skehan, 2003; Fujii, 2005; Kiss & Nikolov, 2005; Sparks, Patton, Ganschow, Humbach & Javorsky, 2006; Sparks et al. 2009; Hummel, 2009).
In the 1990s, FL aptitude was investigated as an effective predictor of FL success. The new trend in aptitude research was motivated by three main reasons: 1) The current advances in cognitive psychology, which have provided explanations of different cognitive skills and aptitudes that compose the cognitive ability for learning new languages; 2) The emergence of new teaching methods which have impacted on most teaching practices and classroom activities (Stansfield & Winke, 2008); 3) The possible linkage between FL aptitude and various factors in SLA research such as age, teaching methodology, and learning context, all in relation to L2 success. These issues will be discussed later in various sections.

Recently, Grigorenko et al. (2000) developed CANAL-FL to examine the impact of cognitive abilities such as Cognitive Ability and Novelty in SLA (See Section 2.8 for more details).

More recently, Meara and Lorenzo-Dus developed a computer-based aptitude test (LAT) at the University of Wales, Swansea (Meara and Lorenzo-Dus, 2003). This was based on Carroll’s 1956 MLAT work. Meara and Lorenzo-Dus’s test faced various problems such as the difficulty of coding non-Roman languages and the difficulty of developing a version of the test independent of L1 examinees. To overcome these weaknesses, Meara in 2005, designed the LLAMA language aptitude test. The new test consisted of three subtests (picture stimuli tasks).

These tests were described as explanatory tests and were designed to measure FL aptitude for various foreign languages. Furthermore, not much detail about their validities was mentioned in the Manual of LLAMA Language Aptitude Tests (2005).
2.2.6 Summary

After reviewing the history of L2 FL aptitude testing from 1920 to the present and describing the most well-known foreign language aptitude tests, it can be noted that many previous studies consider MLAT the most usable one. Although it is considered to be out of date, nevertheless it is still the most reliable aptitude test available. It shows a high magnitude of correlations ($r=0.40$ to $0.60$) with L2 success in various learning contexts. Further detail concerning the MLAT will be provided in Section 2.4.

It is time now to consider, in detail, the four main components that Carroll (1981) identified constituting his FL aptitude test, since these four components are found to be common to most FL aptitude tests that show high predictive power, and since this study uses Carroll’s structure.

2.3 FL aptitude components

The components of MLAT and PLAB have been crystallised to be the bases of the majority of modern aptitude tests (Dahlen, 2008; Skehan, 1988; Winke, 2005). According to Carroll (1981), FL aptitude consists of four main independent components and these four independent variables constitute a good indicator of success in foreign-language learning. He describes them as follows:

1. *Phonetic Coding*: the ability “to identify distinct sounds, to form association between these sounds and the symbols representing them, and to retain these associations” (p. 105).

2. *Grammatical sensitivity*: the ability “to recognise the grammatical functions of words (or other linguistic entities) in sentence structure” (p. 105).
3. *Inductive language learning*: the ability “to infer or induce the rules governing a set of language material, given samples of language materials that permit such inference” (p.105).

4. *Rote memory*: “the ability to learn association between sounds and meanings rapidly and efficiently, and to retain these associations” (p.105).

Phonetic coding is associated with the sound processing of input (Skehan 1989; Skehan 2002). This component concerns a learner’s ability to analyse and code sounds for the purpose of retention. It also affects how much input is saved for the next stage of processing (Skehan, 2002).

Carroll (1981) states that the *Inductive Language Learning* ability was not included in MLAT because, in the 1950s, when MLAT was developed, it was so difficult to measure. As a result, Carroll and his colleague preferred to exclude it.

To overcome this weakness, Skehan (1989) merged the last two components (*Grammatical Sensitivity* and *Inductive Language Learning*) to form one component *Language analytic ability* (Skehan, 1989). What encouraged Skehan to combine both components into one component is the fact that the MLAT does not have a subtest to measure *Inductive language learning* ability (Skehan, 1998, p. 191 as cited in Winke, 2005). Regarding the significance of this component in MLAT, Ehrman and Oxford (1995), in their investigation of the relationship between various individual variables among learners of various languages in different instruction contexts, found that the language analytic component seemed the best predictor of success in second-language learning. The *Words in Sentence* subtest of the MLAT indicates the strongest correlation with L2 proficiency. Furthermore, in another supportive study, Sasaki (1993) found that the language analytic ability component of the aptitude test is the best predictor of
learning success among Japanese students rather than the other two components, phonetic coding and memory. This finding is also supported by Harley and Hart’s 1997 study of the impact of age, aptitude, and second-language aptitude in a bilingual context. They found that analytic language ability is the best indicator of success in second-language achievement when learners are intensively exposed to the L2 in adolescence. In the current study, this ability was measured by the grammar sensitivity subtest in the foreign language aptitude test to examine its impact on the performances of Saudi university students learning English as a foreign language.

2.4 Foreign Language aptitude testing

The following sections will shed some light on two of the most well-known scientific tests of FL aptitude in the history of FL aptitude research: The Modern Language Aptitude Test (MLAT) and The Pimsleur Language Aptitude Battery (PLAB). Their characteristics, components, subtests, and more details about MLAT validity, and its versions will be discussed in detail.

2.4.1 Modern Language Aptitude Test (MLAT)

2.4.1.1 MLAT development

After five years of continuous research between 1953 and 1958, the Modern Language Aptitude Test (MLAT), the first scientific aptitude test, was produced by Carroll and Sapon. The development of this test involved many tests and trials on five thousand participants at Harvard University. MLAT was designed by following some practical steps: 1) by selecting two groups of students, one a group with a high level of attributes and the other with a low level of attributes; 2) by asking both groups to undertake a variety of tasks; and 3) by choosing the tasks that differentiated the two groups’
performances and neglecting those tasks with high correlations (Ehrman 1998). This test can be described, therefore, as a combination of unrelated subtests. Altogether it shows a sufficient degree of validity to be used as a whole as a good predictor of foreign language success (Carroll & Sapon, 1959). Detailed descriptions of MLAT development and its validity were reported by Carroll (1962) in *The MLAT Manual* (Carroll & Sapon, 1959). In the current study, the researcher used MLAT, with some adaptations, to measure FL aptitude among the study’s population.

### 2.4.1.2 MLAT subtests

The following section provides a description of MLAT subtests with some examples. This description was originally provided by Carroll and Sapon in 1959. This test is a paper-pencil test that contains five parts. The examinee needs 60-70 minutes to complete the whole test. Part 1 and 2 need recorded materials for the instructions and the phonetic materials in these two parts. A full description of these five parts follows:

1. **Number learning**

   In this subtest, learners are taught a number system, in the fictitious “Kurdish” language, from 1-4 as well as “tens” and “hundreds” in combinations with these numbers. The students are tested by hearing the combinations of the numbers e.g. 321, 441, 13 etc. This task is designed to measure auditory ability as memory (Carroll & Sapon, 1959).

2. **Phonetic Script**

   The learners are provided with nonsense words while they are looking at their phonetic scripts. They are then asked to listen to a word and to choose from four alternatives. This subtest consists of 30 items of four words each. This part requires learners to learn a set of
notations for some of the English sounds. Then he will be asked to listen to these sounds and underline the correct ones e.g. *tik, tiyk, tis, tiys*. This subtest aims to measure phonetic ability.

3. *Spelling Clues*

The learner is given a phonetically spelled word and is asked to choose, from five alternatives, the word with the most similar meaning.

Example:

*prezns*

A. kings        B. explanations      C. dates       D. gifts           E. forecasts

In this example, the correct answer is D because ‘gifts’ and ‘presents’ are synonyms.

4. *Words in Sentences*

This part of MLAT requires students to identify the grammatical function of a specified item in one sentence and in a second sentence to choose an item that shares the same grammatical function. This identification does not require a student to have a high level of meta-linguistic knowledge, but it does require him/her to identify pairs of items with analogous roles. Empirically, *Words in Sentences* subtest in MLAT measures grammatical sensitivity.

Example:

*Sara is cutting the APPLE.*

*My brother John is beating his dog with a big stick.*

\[ \begin{array}{ccc}
A & B & C & D \\
Sara & is & cutting & the & APPLE. \\
My & brother & John & is & beating & his & dog & with & a & big & stick.
\end{array} \]

In this example, C. *his dog* is the correct answer because the APPLE in the key sentence and *his dog* in the second sentence share the same grammatical function. Both of them are noun phrase objects of the main verb.
5. Paired Associates

In this subtest, the learner is provided with 24-word Kurdish-English sets and is given four minutes to memorise 24 Kurdish-English words and their equivalents. Then, the sets are withdrawn and the examinee is asked to choose the Kurdish equivalents for five English words. Carroll and Sapon (1959) used Kurdish language as an example of nonsense language and it does not related to real Kurdish language.

2.4.1.3 Validity of MLAT

The validity of MLAT is a debated issue among researchers in the field of language testing. Some studies provide sound evidence that MLAT has a good validity regardless of the learner’s age, and the teaching method used. Carroll (1981) states the range of the predictive validity coefficients:

“the predictive validity coefficients for foreign language aptitude batteries in representative samples are typically in the range .40 to .60 against suitable criterion measures of success in foreign language attainment, such as final course grades, objective foreign language attainment tests, or instructors’ estimates of foreign learning ability” (p. 96)

He also adds that aptitude test scores are a better predictor of foreign-language success than any other achievement tests.

Besides Carroll’s studies, many other studies have investigated the validity of MLAT. (See Appendix 1 for additional details). Culhane (1970) investigated the validity of MLAT in predicting examination scores at the British University. He found that MLAT alone could be a satisfactory predictor, but that previous learning experience should not be excluded. Ehrman (1998), in her study of 1000 students at the Foreign Service Institute, found that MLAT coefficient validity ranges from .40 to .50; almost the same as the original MLAT validity by Carroll in 1958.
More recently, researchers have investigated the correlation between analytic ability as a component of MLAT and the ultimate foreign-language achievement in communicative settings (Harley & Hart, 2002; Ranta, 2002). Their findings indicated a high validity coefficient between learners’ analytic ability and the achievement scores, making MLAT a good predictor of language-learning success in communicative conditions.

Some researchers and educators have investigated the reasons why aptitude is considered to be a reliable predictor of second-language learning. Skehan (1991, 1998, 2002) notes that FL aptitude tests were designed to measure a set of abilities: linguistics skills and knowledge (the measurements of oral and written skills) and the skills in using decontextualised language (the measurement of general intelligence). Skehan (1986b) states that “language aptitude as reflected in the structural aptitude tests is, therefore, an effective, predictive hybrid of a fundamental ability combined with oral language, with an ability to handle language as a vehicle” (p. 60).

2.4.1.4 Using MLAT with non-native speakers of English

Although the MLAT is designed for native speakers of English, it has been used with non-native speakers of English in various studies (Nagata et al. 1999; Robinson, 1997). For instance, Robinson (1997) used the English MLAT to measure aptitude among Japanese native speakers. He clarified the threat of using English MLAT with non-native speakers of English when he pointed out that “differences in scores on the aptitude test might simply reflect differences in L2 proficiency” (p. 62). Another example was a study conducted by Nagata and his colleagues in the same FL context. They indicated the limitation of their study in using English aptitude tests with Japanese native speakers. They suggested that “[i]deally, [aptitude] tests designed for native speakers of Japanese
should have been used. Unfortunately, they were not available for us” (p. 141). In addition, the English MLAT is considered invalid for non-native speakers of English in certain large government institutions such as the Foreign Service Institute in the USA. The critique was put forward on the basis of ‘face validity’. FL aptitude test scores determined by such a method do not necessarily represent aptitude only. The scores were also likely to be impacted by English-language proficiency. To avoid this kind of threat, the MLAT was adapted into Arabic (L1 of this study’s participants). More methodological procedures related to the constructing of foreign language aptitude test are provided in Chapter 3.5.1.

2.4.1.5 MLAT versions

The MLAT as a main measure of aptitude has subsequently been used widely in different learning contexts in the United States. It is also available for the visually impaired and for non-English speakers. MLAT has been translated into different languages: French (Weshe et al., 1982), Hungarian (Ottó, 2002) and Japanese (Murakami, 1974). Recently, Sasaki(1996) developed a Language Aptitude Battery for Japanese native speakers.

2.4.2 The Pimsleur Language Aptitude Battery (PLAB)

Following the same process as Carroll’s in MLAT, Pimsleur designed another interesting aptitude test in 1966: the Pimsleur Language Aptitude Battery (PLAB). This battery expands the coverage of aptitude tests to include three main components: verbal intelligence, auditory ability and motivation. This test aims to measure auditory ability and motivation, which are not included in MLAT. Pimsleur included motivation in his test because he believed that interest in learning a foreign language can be considered cognitive ability that has a good degree of predictive potential in L2 acquisition.
Like MLAT, PLAB is a paper-pencil test that consists of six sections. The time allowed for its administration is 60 minutes. In Parts 5 and 6, recorded materials are needed for the instructions and phonetic tasks. The test’s six parts are described as follows:

1. **Grade Point Average**

   Students have to record their last achievement scores in various subjects such as history, mathematics, science and English.

2. **Interest in Foreign-Language Learning**

   In this part, students are asked to determine the degree of their interest in learning a foreign language on a 5-point scale.

3. **Vocabulary**

   Students are provided with a list of 24 moderately difficult adjectives. Each adjective is followed by four words. Students are asked to choose the word with the same meaning of the above adjective.

   **Example:** prolonged
   a. prompt c. difficult
   b. decreased d. extended

4. **Language Analysis**

   Students are provided with a list of words and phrases from a foreign language and their English equivalents. They are required to produce new words or phrases in a foreign language. The second section of this part includes fifteen English phrases which were required to translate into the foreign language. Students are required to choose the correct answer from several translations.

   *gade* .................. father, a father
   *shi* ...................... horse, a horse
   *gadeshir le* .............. Father sees a horse.
   *gadeshir la* ............. Father saw a horse.
   *be* ........................ carries
Students are then told to use the above list to figure out the correct translation for the following statement and then to choose the correct translation from the following phrases.

\[ A \text{ horse carried Father} \]

\[ a. \text{ gadeshir be} \quad b. \text{ gadeshirba} \]
\[ c. \text{ shigaderbe} \quad d. \text{ shigaderba} \]

5. **Sound discrimination**

Students are taught three similar sounds in a foreign language by listening to a tape recording. They are then played 30 sentences containing the three sounds and asked to determine the sound in each sentence.

6. **Sound-Symbol Association**

Subjects hear a range of words containing two or three syllables from an unknown language and are then instructed to choose the correct answer from a list of four words.

2.4.3 **Comparing MLAT and PLAB**

These descriptions clearly demonstrate that there are some similarities between MLAT and PLAB; for instance, MLAT’s second part, Phonetic Script, is similar to Part 6 in PLAB’s Sound Symbol Association. There are also some differences between the two batteries. For example, in PLAB there is a focus on auditory ability which is not included in MLAT. Furthermore, the inclusion of previous students’ scores in various subjects and motivation to learn a foreign language is considered to be the main components in PLAB.

Pimsleur (1966) insists that previous learning scores of FL learners and their motivation in learning FL are essential components in designing FL aptitude batteries. In relation to the students’ final scores in previous learning, Pimsleur, in the PLAB Manual, states that
in various studies, the grade point average is a good predictor of success in learning FL and it was an even better predictor than the total score of PLAB itself (Rees, 2000).

This researcher has not been able to locate a study comparing the two tests’ power to predict SL outcomes, but in many studies, MLAT has shown encouraging correlations ranging from .40 to .60 with L2 success (Skehan, 1989).

2.4.4 Summary

After reviewing, in detail, the two most well-known FL aptitude tests (i.e. MLAT and PLAB), the preference is to use MLAT over PLAB for the study of the Saudi context. Choosing MLAT over the other aptitude test is based on the following reasons: 1) MLAT is considered to be the most reliable and valid aptitude test; 2) it has been adapted to non-Roman languages in various learning contexts (Japanese, Polish, and Hungarian).

2.5 Various issues related to foreign language aptitude

2.5.1 The FL aptitude, teaching methods and L2 proficiency link

The relationship between FL aptitude and second-language teaching methods has been one of the most controversial issues in FL aptitude research. Some researchers believe that the FL aptitude predicts L2-learning success regardless of the kind of teaching methodology used. Others dismiss FL aptitude, arguing that the origin of FL aptitude research in the 1950s was associated with the outdated audio-lingual teaching method, and that it is not adequate for the current communicative approach of modern SLA theories or new teaching methods (Ranta, 2002). For instance, Cook (1996) has argued that FL aptitude could be a good predictor of second-language success in a course that is taught by the audio-lingual method, but not when other teaching methods are applied. These claims led to further empirical studies.
On the one hand, various studies have found that FL aptitude has a sufficiently reliable predictive power of likely L2 achievement with various teaching methodologies (Ehrman & Oxford, 1995; Harley & Hart, 1997, 2002; Reves, 1983). In line with Carroll’s belief, in two of the most comprehensive studies, Ehrman and Oxford (1995) and Ehrman (1998) examined the relationship between various learner characteristics such as FL aptitude, learning strategies, learning style, personality, motivation, and anxiety with L2 proficiency in speaking and reading skills, where the instruction was mostly communicative except for the inclusion of some minor audio-lingual components. The subjects were 855 employees who had taken intensive language courses. Their findings showed that FL aptitude test scores using MLAT with FSI (Foreign Service Institute) scores in reading and speaking achieved the highest correlation (.51) compared with other factors such as personality, learning strategies, and language teaching styles. These researchers believe that FL aptitude is still a good predictor of success in L2 even in a communicative teaching environment. For more explanations, Ehrman (1998) provides some possibilities regarding the robust impact of FL aptitude irrespective of the teaching methods used: 1) MLAT is possibly a multi-dimensional test with the ability to measure a set of skills in different teaching methodologies. 2) The internal variations in teaching methods inside classrooms may make a difference. For instance, some language training programs tend to teach analytic skills extensively, which are more specifically measured by MLAT, than other skills are. This can lead to a higher prediction of second-language achievement.

On the other hand, a considerable number of studies oppose the notion of the impact of FL aptitude on L2 achievement in various teaching methodologies. In line with Cook (2001), in one of the most recent studies, Sáfár & Kormos (2008) found that FL aptitude
plays a relatively weak role in predicting success in a one-year communicative teaching program based on focus-on-form instruction among Hungarian high school students. Notably, this study, as most FL aptitude and L2-proficiency studies, could be criticised for simply assessing language proficiency once at the end of the academic year. A single assessment does not show the level of progress among students over the whole program. Furthermore, this study used a quasi-experimental design that failed to show causality. This weakness of the correlation studies is mentioned by the researchers themselves as one of the main limitations of their study: “…finding a significant correlation is no proof for the direction of causality, nor does it reveal its very existence” (p. 132).

Another study was conducted by Ranta (2002) in communicative classrooms and investigated the relationship between FL aptitude subtest (analytic ability) and the ultimate language learning achievement. The results show that the correlation is relatively weak. In another communicative context at the University of Pennsylvania, where the communicative teaching method was the basis of the language program, Goodman et al.,(1990) used the short form of the MLAT to validate the criteria for exemption in the university population. A fairly large number (586) of students in their first year participated in the study. The study aimed to determine the relationship between the aptitude scores of students and their foreign-language course grades in the first and second semester in the freshman year. The researchers found that there was a relatively weak correlation between MLAT and language proficiency and suggested that this weakness might be due to the following: first, there was the restricted range of students grades, since 49% of the students got A+, A, or A-, and 31% got Bs. Second, there was the possible effect of differences in teaching methods. Goodman and his colleagues hypothesised that MLAT could predict language proficiency in the case of traditional teaching methods, which measure grammar, but that it was not sufficient in measuring
communicative skills. The University of Pennsylvania foreign-language programs, however, focused on communicative skills and the use of the language.

It can be seen from the above that the role of FL aptitude measures as a reliable predictor of language success with various teaching methodologies is not clear. More research on this topic needs to be undertaken before the association between FL aptitude and teaching methods can be more clearly understood. The current research is designed to contribute to the field by investigating the relationship between the FL aptitude and the English proficiency in King Khalid University where the task-based teaching method is the dominant method.

2.5.2 FL aptitude and stability

The stability of FL aptitude is another controversial issue among researchers. While some findings support the claim of stability, other findings are ambiguous, inconclusive, or appear to negate the claim. These research designs and arguments concerning the stability of FL aptitude will be identified and reviewed to assess their relevance to this study.

Some researchers believe that FL aptitude is innate and cannot be trained or modified (Carroll, 1981; Kiss & Nikolov, 2005; Skehan, 1989, 1991; Sparks & Ganschow, 2001). For example, Carroll (1993) supports this notion when he states that “to the extent that cognitive abilities are at least relatively stable and resistant to attempts to change them through education or training, and at the same time are possibly predictive of future success, they are often regarded as aptitudes” (p. 16). Similarly, Skehan (1988) maintains that “foreign language aptitude is relatively fixed over a long period of the individual’s life span, and relatively hard to modify in any significant way” (p.86). Empirically, in one of the early studies of the stability of aptitude, Skehan (1989) points out that Politzer and
Weiss’s 1969 attempt to improve the MLAT subtests scores failed. Those learners had finished a short training program related to various cognitive components in language aptitude tests. Following Politzer and Weiss’s study, Harley and Hart (1997) investigated the impact of aptitude at different ages in immersion programs in Canada. The findings indicated that early immersion students with twelve years’ exposure to L2 did not have a higher level of aptitude than did late immersion students with four years exposure to L2. This finding could be criticised for not comparing the FL aptitude for the same group. Instead, Harley and his colleague compared two different groups: an early immersion group and a late immersion group.

By comparison, some researchers believe that FL aptitude is dynamic and it can be affected by previous learning experience (McLaughlin, 1990). Grigorenko et al. (2000) propose that “aptitude should not be viewed as a static personality trait; novices can become experts with experience” (p.173). Eisenstein (1980) also emphasises that “language aptitude is a form of developing expertise rather than as an entity fixed at birth” (p.401). This claim was empirically supported by various studies.

Sáfár and Kormos (2008), in their longitudinal study, compared language aptitude, as measured by HUNLAT (The Standard Hungarian Language Aptitude Test, Ottó 2002), between two groups. The first group included 40 English-Hungarian bilingual students with good L2-learning experience in studying new languages. The second consisted of 21 regular secondary school students without any previous L2-learning experience. The researchers found that the participants’ aptitude scores were significantly improved after exposure to a nine-month intensive English course (576 hours of L2-learning by the end of the course) in both groups, but bilingual students with previous learning experience achieved higher aptitude scores than those students in the regular secondary school
without the same linguistic privilege. The fact that the Sáfár and Kormos’ study compared results of non-comparable groups is significant. Another study was conducted by Skehan (2002); his findings demonstrated that instruction in Latin significantly developed both native-language skills and FL aptitude scores in MLAT among learners with and without learning disabilities after one year of FL instruction. The findings of Skehan’s study need to be handled carefully in any comparisons, since Skehan’s study involved language learning input from a third language (Latin).

Thus, it can be seen from the previous studies that there is no persuasive evidence that aptitude is either stable or dynamic. As Skehan suggests, because of current uncertainty of stability of aptitude, researchers should follow Carroll’s assumption that aptitude is stable over a long period of time. Skehan points out that until the year of 2002, existing evidence supported the idea of aptitude stability, and while this has changed somewhat, more recently, “for now following Carroll, we will assume that aptitude does not change with the seasons!” (p.79). Many later foreign language aptitude studies deal with the concept of aptitude according to Carroll’s assumption, as a stable factor during any language teaching program. Only a few studies have investigated the stability of aptitude and the effect of language learning experience on FL aptitude (Kiss et al., 2005 and Thompson, 2008).

It should be noted that Carroll (1973) himself was somewhat ambivalent on the issue of stability of aptitude. He generally accepted the innateness of aptitude in acquiring L2, but then stated that aptitude may be affected by previous learning experiences of learners (Harley & Hart, 1997). Hence, while FL aptitude appears to be relatively stable over a long period of a learner’s life, previous language experience may affect aptitude levels, especially at an early age (Dörnyei, 2005; 2006).
What motivated the researcher, in this study, to investigate the issue of aptitude stability were the following reasons. First, to the researcher’s knowledge, no study has been conducted to investigate aptitude stability in a Saudi context. In this regard, the results of the current study would, to some extent, enrich FL aptitude research in a novel learning context. Second, most previous studies have dealt with FL aptitude as a stable construct and have not taken into account the possible change of aptitude components during their training programmes. By examining FL aptitude over time, i.e. before and after a seven month period of intensive English language studies, this study would be able to test whether the foreign language aptitude test is stable and not affected by English instruction.

2.5.3 FL aptitude and L1 literacy

After reviewing the relatedness between FL aptitude and previous learning experience, it is necessary to focus on the impact of native language (L1) skills, such as literacy, on FL aptitude. For the last few decades, numerous studies have demonstrated L1-FL aptitude relationships. Most of these studies have investigated possible correlations between L1 skills and FL aptitude (Sparks et al. 1995). Sparks et al. (1998) indicate that individuals with high levels of oral and written proficiency in their L1 achieve high scores in FL aptitude measures, while those individuals who show low levels of L1 skills, or those who lose some L1 skills for any reason, exhibit low scores in FL aptitude tests. An extensive number of studies have been conducted in the United States. Most of these studies examine the relationship between L1 skills, especially phonological ability, and L2-learning, as well as studying the impact of various factors such as cognitive, affective and social factors (Sparks & Ganschow, 2001; Sparks et al. 2008). Their findings show that
early L1 skills, such as spelling and reading skills, play a significant role in L2 achievement in later years.

A number of longitudinal studies have been conducted on the impact of L1 skills on L2-learning and FL aptitude (Sparks et al. 2009 and Skehan, 1989). The Bristol follow-up study is considered to be one of the pioneering studies in aptitude research in 1970s and 1980s (Skehan, 1986a, 1989). The Bristol study (Wells, 1985) investigated the rate of L1 skills development among 125 children. The results of this project indicated that there are noticeable differences in the rate of L1 skills development among the children. Wells noted “scores on measures of oral language at successive age-points yield progressively stronger predictions of later educational attainment” (p. 181). To investigate the origin of aptitude and the possible correlations between L1 skills and L2 achievement and FL aptitude, Skehan used Well's data in his follow up study. He selected 23 secondary school students who had participated in the Bristol project who were at the time of study 13 to 16 years old. Their L2-learning scores in French and German were correlated with previous L1 skills and FL aptitude. The findings showed sound correlations between L1 skills’ developments and French and German proficiency levels and between L1 literacy skills and FL aptitude. Moreover, it can be seen that FL aptitude is a good predictor of FL achievement.

Skehan proposed that the learner’s innate ability to acquire any language is based on L1 development. Correspondingly, Sparks et al. (2006), in the study of the relationship between L1 skills and FL aptitude, also found that a substantial part of the variation in FL aptitude can be explained by L1 literacy.

In some supportive research, Richard et al. (1998) conducted two studies to investigate the relationships between L1 skills, FL aptitude and FL grades at the end of a L2 program
among learners with different levels of proficiency (High, Average and Low). The results of the two studies indicated overall differences in native language literacy skills, FL aptitude and final foreign language grades among all three levels of FL proficiency groups. The findings extracted from both studies clearly support the significance of MLAT as an aptitude measure and as a predictor of oral and written FL proficiency among all the three groups. The aptitude test was found to be not only a good predictor of language learning success in the first year L2-learning program, but also after two years of FL learning. In these studies, MLAT was more highly correlated (.49) with overall language proficiency scores than any other measure of native language skills. Given this, Richard and his colleagues concluded that learners with a good level of L1 literacy skills and a high level of FL aptitude achieve a high level of oral and written FL proficiency as well as higher FL grades at the end of the L2 program.

Recently, in a longitudinal study, Sparks et al. (2009) found that learners with high proficiency in L1 skills show stronger L2 aptitude than the middle and low proficiency L1 learners. These researchers hypothesise that learners’ FL aptitude will account for most of the components of FL competence and formal FL learning success.

To summarise, it can be seen from most of the previous studies that both FL aptitude and L2 achievement appear to be strongly related to early L1 literacy. To put it differently, children who show good progress in their L1 skills achieve high scores in FL aptitude tests (Skehan, 1986a). Educators and researchers should develop sufficient methods to improve teaching L1 skills effectively as a basis of future L2 success (Bialystok & Fröhlich, 1978; Sparks et al., 2009).

The association between Arabic literacy and English proficiency has not been investigated in the study being reported in this thesis. Consequently, it is recommended
that future research investigate the impact of Arabic literacy on English learning in the Saudi context.

2.5.4 FL aptitude and intelligence

One of the most provocative issues is the relationship between FL aptitude and intelligence. Some researchers, for example Dörnyei, deemphasise the existence of language aptitude. Instead, they argue that general cognitive abilities are those that enable students to master a new language (Dörnyei, 2005). According to this notion, language-learning ability can be measured by intelligence tests in the same way as any other general cognitive ability. Intelligence tests were developed in the early 1900s and have been used to measure students’ intellectual ability to study in public schools (Dörnyei, 2005).

Others oppose the notion of having a talent or ‘aptitude’ to learn a new language. They also believe that FL learning ability, which can be measured by FL aptitude measures, differs from general cognitive ability, which can be measured by intelligence tests. This debate raises the need for further empirical studies to investigate the possible relationships between FL aptitude and intelligence. For this purpose, most of these studies correlate L2 aptitude tests scores with intelligence scores. To investigate the relationship between FL aptitude and intelligence empirically, Carroll, in one of his initial studies (1981), found that the correlation between intelligence scores in the IQ test used and FL aptitude scores in MLAT ranged from .34 to .52. It was observed that two learners with the same intelligence scores in the IQ test showed differences in their ease and rate of learning a foreign language (Wesche et al., 1982). In other words, FL aptitude measures, such as MLAT, seem to predict L2 success better than the IQ test predicts L2 success. Carroll proposed that intelligence and FL aptitude are two independent factors in relation to FL
proficiency. He also noted that most intelligence tests were designed to measure general cognitive ability. These abilities differ from those involved in L2 learning.

Later, various supportive studies were conducted in different learning contexts and showed that the correlation between aptitude measures and intelligence measures was not significant (Gardner, 1985; Skehan, 1986b; Stansfield, 1989; Wesche et al., 1982). In Japan, Sasaki (1996) conducted a study to examine the relationship between FL aptitude measure, MLAT, and Primary Mental Abilities (PMA). He found that the correlation between MLAT and PMA was relatively weak. Sasaki noted that MLAT subtests could be used to measure various language learning abilities that differ from the general cognitive ability measured by PMA. Sasaki (1996) also reported a relatively weak correlation between the Wechsler Adult Intelligence Scale and the Language Aptitude Battery for Japanese. It was only .17 in the Japanese context.

The most interesting findings emerging from these aptitude-intelligence studies is that intelligence and FL aptitude appear to have a complex relationship. While they share some features in common, they are not the same (Dörnyei, 2006; Ehrman, 1998).

Skehan (1989) suggests that FL aptitude should not be treated in isolation from general cognitive abilities. He argues that analytical language ability, as a component of FL aptitude, appears to be the best subtest to differentiate between intelligence and aptitude compared with other abilities such as memory and phonetic coding ability. Thus, it can be assumed that there is an overlap between FL aptitude and intelligence.

2.5.5 FL aptitude and learners’ cognitive style

Previous studies have indicated that FL aptitude components are clustered cumulatively to influence foreign-language achievement. There is an assumption that aptitude is not
simply a predictive instrument for language-learning success but that it can be a diagnostic tool to align language learners with appropriate learning methodologies (Carroll, 1962; 1981; Skehan, 1998; Wesche, Edwards, and Wells, 1982). This claim was empirically supported by Harley and Hart (1997) who categorised students into two groups (analytically and memory-based learners) according to their performance in aptitude subtests. In 1997, Harley and Hart investigated the relationship between FL aptitude components and French proficiency among 65 students at different starting learning ages. Based on students’ performances in these subtests, Harley and Hart suggest matching the students to the appropriate teaching methodologies. They found that matched learners expressed great satisfaction with their performance, contrary to their mismatched learners in the other group. These findings demonstrate the positive impact of matching teaching methods to subsets from the aptitude profiles. In this case, educators need to know not only that there are differences in aptitude levels but also how to determine the type of learner (Harley & Hart, 1997). The disadvantages of these categorisations, though, are that the aptitudinal variations among learners are not continuous and there is no clear-cut way to differentiate between analytic learners and memory learners.

The Saudi Arabian educational system officially encourages the analytic way of L2 learning but most teaching practices, inside classrooms, are mainly based on memorisation.

2.5.6 Gender differences and FL aptitude

The issue of gender has not been a much disputed subject within the field of FL aptitude. Despite some studies that found that females were significantly outperforming men in FL
proficiency and in MLAT scores (Kiss & Nikolov, 2005), a number of other studies have shown that gender differences have a weak relationship with FL aptitude scores (Carroll, 2002). Ehrman (1998), in her study of 343 students in FSI found that MLAT total scores and its subtests were not affected by gender differences. In other words, the predictive power of the MLAT is almost the same for both males and females despite the fact that females show higher grades in FL learning and in MLAT scores (Carroll, 2002). In accordance with this finding, Carroll claimed that “no consistent differences in the validity of the MLAT for the two sexes have been observed” (p. 26).

In light of the above, it appears that gender homogeneity is unlikely to limit the external validity of the current study findings. The current study investigated only the impact of aptitude on English proficiency among male learners since because of cultural reasons: male researchers do not have access to female colleges. Therefore, it is recommended for future studies to investigate the relationship between FL aptitude and English proficiency for the female Saudi population.

2.6 Uses of the FL aptitude tests

Aptitude tests used in education

The benefits of FL aptitude have been investigated by many researchers in various learning environments. Since the 1960s, FL aptitude tests and aptitude profiles have been used by language teachers, educators, and decision-makers in educational fields such as universities, schools, colleges, and language teaching centres for a number of purposes.

FL aptitude measures are administered for various purposes, such as selection, placement, streaming, and diagnosis. In foreign language teaching, Cook (2001) has summarised the benefits of aptitude measures in four main areas. First, they can be used to select learners
who learn a new language quickly and successfully. Second, they can categorise learners into different groups on the basis of their aptitudes scores. Third, new teaching methods and new curricula and materials can be developed according to the learners’ types of aptitude. Fourth, students with extremely low aptitude can be eliminated from language programs.

Various studies have been conducted for the same previously mentioned purposes in major institutional settings around the world. For example, in the USA, Carroll (1981) and Reed and Stansfield (2002) point out that MLAT has achieved success in universities and colleges when applied as a diagnostic tool for students who fail to achieve foreign-language requirements. Pimsleur (1966) shed some earlier light on this issue when he proposed, in terms of his own battery of tests, that if aptitude tests were intended to discover areas of strengths and weaknesses in students’ performance, then, on the basis of learners’ aptitude scores, curricula could be adjusted to meet the learners’ needs. Along the same lines, studies by Wesche (1981) and Ehrman (1996, 1998) found that aptitude tests could significantly advantage curriculum designers and educators. For instance, based on students’ scores in aptitude subtests, records of the students’ weaknesses and strengths can be established, and students with low aptitude scores can be advised and guided to use new methods or strategies to overcome their FL-learning problems. In 2008, Kristina Dahlen studied the Turkish vocabulary acquisition among 88 native speakers of English at tertiary level. She investigated the relationship between FL aptitude and Turkish vocabulary acquisition. The findings indicated that students with high phonological awareness scores, measured by MLAT, were able to learn more Turkish words than students with low scores. Furthermore, the components of MLAT, such as phonetic coding, grammar sensitivity and rote memory, can be used to detect students at risk of foreign language learning difficulty (Dahlen, 2008).
In the Saudi context, the results of this study could be used to identify the areas of weakness and strength in students’ aptitude abilities and it could be used also to improve the curriculum. Altogether, this could contribute substantially to improving the current English teaching program in Saudi universities.

Aptitude tests used in government agencies

Many aptitude studies have been conducted in government agencies for selection, placement, and the diagnosis of weaknesses and strengths in learners’ foreign-language acquisition among employees and personnel intending to enrol in intensive foreign-language programs around the world. For instance, in the USA, the Defense Language Aptitude Battery (Peterson & Al-Haik 1976) was designed to be used as a selection device for army personnel for foreign-language programs. The major aim of this test was to minimise training costs and the amount of time required to finish the program (Dörnyei, 2006). Another example is found in Canada, where Wesche (1981) used aptitude tests scores to screen 397 English-speaking Canadian public servants taking part in the French-language training program of the Commission of Canada for their abilities in learning a foreign language. He used a combination of information regarding their FL aptitude and their academic achievement scores. He found that aptitude tests were beneficial in selecting learners with high aptitude to enrol in the intensive program. He also showed that aptitude tests could be used to adapt programs to meet learners’ needs. For instance, in the same study, the aptitude measures were used as both diagnostic and selection tools. In the same way Ehrman (1996, 1998) conducted a study in the Foreign Service Institute (FSI) in the USA among learners from various linguistic backgrounds. Her findings demonstrate that the total score of MLAT can be used as a diagnostic tool to
indicate the areas of weaknesses and strengths in learners’ performances in learning a foreign language.

**FL aptitude and Disability**

MLAT, as the most common FL aptitude test, is considered to be one of the most important diagnostic tools to determine language-learning disabilities in institutional settings.

For instance, in an empirical study at three western universities in the United States, Sparks & Ganschow et al. (1991) examined fifteen students who were identified as language-learning disabled and another fifteen students who were classified as good language learners. The study aimed to predict language-learning difficulties by including both intelligence and aptitude tests (MLAT: short and long forms) as its battery. They found that the MLAT scores could be used as a tool to differentiate between good and disabled language learners.

In another study, at the University of Pennsylvania, Gajar (1987) evaluated the use of MLAT in predicting foreign-language success. He also used MLAT to compare the performances of two groups: one group identified as learning disabled and the other group as mainstream. He found that MLAT was a good predictor of L2 success, particularly in Parts 4 and 5 (Words in Sentences and Paired Associates), both of which significantly correlated with L2 achievement. Gajar found that learning-disabled students faced a weakness in completing the subtest relating to grammatical structures and memory tasks. These two areas, he argued, should be investigated further to overcome the L2-learning difficulties among disabled-learning students. Gajar confirmed the importance of MLAT as a good predictor of success in learning L2 and as a diagnostic tool to determine
learning-disabled students. He also cast doubt on the importance of motivation for L2 learning in the case of those students who do not perform well in MLAT or repeatedly fail in L2 classes.

2.7 Methodological approaches in FL aptitude research

In most of the previous studies, the relationships between FL aptitude and second-language proficiency have been investigated cross-sectionally (Ehrman, 1998; Skehan, 2002; Ehrman, Oxford, 1995). The measures of FL aptitude and L2 proficiency have been administered in the same session or after a very short period of time. Very few studies have followed participants several years later to examine the impact of FL aptitude in predicting L2 success (Sparks et al. 1998, 2006).

Most of the FL aptitude -L2 proficiency long-term studies have been conducted to examine the stability of aptitude or to determine the best predictors of L2 proficiency. Sparks and his colleagues conducted many long-term studies to investigate the effect of L1 skills on later L2 proficiency and L2 aptitude (Sparks et al. 2009). Their studies aimed to provide an answer for the question of why L2 aptitude measure is a good predictor of L2 achievement. They support Carroll’s belief that L2 aptitude is a “residual” of L1 skills. This claim, relating to the link between L2 aptitude and L1 skills, is supported by various studies (Sparks et al., 1998, 2006)

In one of the longitudinal studies, Skehan and Ducroquet (1988) investigated the relationship between L1 skills, L2 aptitude, and L2 achievement. Using the data from the participants in Wells’ 1985 study, Skehan and his colleague administered L2 aptitude and L2 proficiency for the same participants ten years later. The findings showed that there is a strong relationship between L1 skills and L2 aptitude. Even though this is a longitudinal
study, it could be criticised for not directly examining the L2 aptitude and L2 proficiency longitudinally. The study just examined the development of L1 skills over a period of time and followed the same ways of the aforementioned studies in investigating the relationship between L2 aptitude and L2 success cross-sectionally.

In past studies, the stability of FL aptitude was not investigated widely. Most researchers tended to accept Carroll’s claim that aptitude is stable and almost impossible to modify, even though this issue is still controversial as mentioned previously in the FL aptitude stability section (2.5.2). Various studies have investigated this claim experimentally (Harley & Hart, 2002; Kiss & Nikolov, 2005).

Most of the previous studies that investigate the relationship between FL aptitude and the L2 proficiency cross-sectionally and longitudinally have correlated the total scores of FL aptitude measures with the total scores of L2 measures. Only a few studies have considered the correlation between FL aptitude scores and the L2 progress (Haslam, 2010; Erlam, 2005). Even those studies that measured the L2 progress used the absolute value formula, which can be criticised statistically for not measuring the actual progress. Instead the formula only measures the differences between the L2 scores before and after L2 instruction.
The following tables summarise the extent to which the studies noted in this literature review are relevant to the current inquiry.

Table 2.1

*A Summary Table of the Cross-sectional Studies of the Relationships between FL Aptitude and L2 Proficiency at Time 1*

<table>
<thead>
<tr>
<th>Study</th>
<th>Participants</th>
<th>Country and Context of learning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Harley and Hart (1997)</td>
<td>65 eleventh-grade students</td>
<td>Canada (French: SL)</td>
</tr>
<tr>
<td>Sparks and Ganschow (2004)</td>
<td>101 ninth-grade students (fifteen years old)</td>
<td>USA (Spanish: SL; French: SL) and German: FL)</td>
</tr>
<tr>
<td>Hsieh (2004)</td>
<td>200 first-year University students</td>
<td>Taiwan (English: FL)</td>
</tr>
<tr>
<td>Hummel (2009)</td>
<td>77 first-year university students</td>
<td>Canada, Quebec (English: SL)</td>
</tr>
<tr>
<td>Gardner (1997)</td>
<td>University students</td>
<td>Canada (French: SL)</td>
</tr>
</tbody>
</table>
### Table 2.2

**Summary of Cross-sectional Relationships of FL Aptitude and L2 Proficiency at Time 2**

<table>
<thead>
<tr>
<th>Study</th>
<th>Participants</th>
<th>Country and Context of learning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kiss and Nikolov (2005)</td>
<td>419 twelve-year-old Hungarians</td>
<td>Hungary (English: FL)</td>
</tr>
<tr>
<td>Rysiewicz (2008)</td>
<td>Grade 7: Age thirteen</td>
<td>Poland (English: FL)</td>
</tr>
<tr>
<td>Bialystok Fröhlich (1978)</td>
<td>157 Grades 10-12</td>
<td>USA (French: FL)</td>
</tr>
</tbody>
</table>

### Table 2.3

**Summary of the Longitudinal Relationship between FL Aptitude at Time 1 and L2 Proficiency at Time 2**

<table>
<thead>
<tr>
<th>Study</th>
<th>Participants</th>
<th>Country and Context of learning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sparks et al. (2009)</td>
<td>Ninth-grade students</td>
<td>USA (French: FL)</td>
</tr>
<tr>
<td>Ehrman &amp; Oxford, 1995</td>
<td>855 from government departments</td>
<td>USA (different languages)</td>
</tr>
<tr>
<td>Ehrman, 1998</td>
<td>343 (Foreign Service Institute)</td>
<td>USA (different languages)</td>
</tr>
<tr>
<td>Spark et al, 1998)</td>
<td>High school students</td>
<td>USA (Spanish, German and French)</td>
</tr>
<tr>
<td>Thompson (2008)</td>
<td>University students</td>
<td>USA (Portuguese-FL)</td>
</tr>
<tr>
<td>Goodman et al. (1990)</td>
<td>586 first-year university students</td>
<td>USA (different languages)</td>
</tr>
<tr>
<td>Sparks et al (1998)</td>
<td>96 High-school students</td>
<td>USA (Spanish, German, French)</td>
</tr>
<tr>
<td>Sáfár &amp; Kormos (2008)</td>
<td>40 Hungarian secondary-school students</td>
<td>Hungary (English: FL)</td>
</tr>
</tbody>
</table>
2.8 Recent theoretical explanations for FL aptitude and L2 proficiency

2.8.1 Overview

The last few decades have witnessed a number of new trends in FL aptitude research. These new trends go beyond the traditional FL aptitude to the broader area of cognitive abilities in the field of SLA. These new trends in FL aptitude research mainly examine the role of specific cognitive abilities such as working memory and phonological coding on SLA (Dörnyei, 2005). In this section, the best-known works in the last fifteen years are presented: the work of Grigorenko et al. 2000 in developing a new aptitude test (CANAL-FL); the contribution of Sparks et al. (2001) related to the impact of the linguistic coding in L1 literacy skills in predicting L2 success and FL aptitude by using the Linguistic Coding Differences Hypothesis; the role of working memory as a cognitive ability in SLA; Skehan’s work (2002) and Robinson’s research on the aptitude complexes (2002).

2.8.2 CANAL-FT (Cognitive Ability for Novelty in Acquisition of Language-Foreign Test)

During the past ten years further FL aptitude tests have been designed. For instance, a recent aptitude test was developed by Grigorenko, Sternberg, and Ehrman (2000): the CANAL-FT (Cognitive Ability for Novelty in Acquisition of Language-Foreign Test). This test is derived from the cognitive theory in language acquisition. The core assumption of this test is that successful learners of a foreign language tend to use an ability to manage novelty and ambiguity (Ehrman, 1996; Ehrman & Oxford, 1995; Ranta, 2008). This theory, as cited in Grigorenko, et al., (2000), is associated with experimental studies of the triarchic theory of human intelligence. This test is mainly designed to avoid the weaknesses of the former aptitude tests (Abrahamsson & Hyltenstam, 2008). Grigorenko, et al. (2000) have crystallised the differences in this test in three main ways:
1) it is based on cognitive theory in psychology as a theoretical background; 2) it is dynamic, in that it involves some learning during the time of the test; 3) it is a simulation-based test. The CANAL-FT correlates significantly with MLAT \( (r = .75) \). The authors of the test pointed out that CANAL-FT is immature work and it can be considered as a foundation for further research (Grigorenko et al., 2000). Nowadays, the test is still valuable to some extent as a diagnostic tool and to investigate learners’ weaknesses and strengths in learning a new language through their aptitude profiles.

2.8.3 Linguistics Coding Differences Hypothesis (LCDH)

Another significant development in understanding the relationship between FL aptitude and L2 success was the enunciation of the Linguistics Coding Differences Hypothesis (LCDH) by Sparks, with others (1991, 1995). This hypothesis proposes that L2 learning and FL aptitude are rooted in L1 skills such as phonological skills and sound-symbol system skills. This hypothesis stimulated more empirical studies to investigate the nature of these relationships between L1 skills, L2 learning, and FL aptitude.

2.8.4 Working Memory in SLA

Working memory in SLA is one of the more important issues in the future of FL aptitude research (Miyaki & Friedman, 1998). The role of working memory is not as a temporal storage of information but it involves a dynamic combination of processing and storage of information which differs from the traditional rote memory in Carroll’s MLAT (Dörnyei, 2005). These two components — the processing functions and the storage of information — are essential in comprehension and production of L2.

The current study used the rote memory in the traditional aptitude test. The findings of the rote memory in the foreign language aptitude test will be investigated in the chapters that
follow. The findings obtained will allow some implications to be drawn about whether to include the rote memory in the aptitude test in the future or whether to design a tool that measures the use of the working memory instead.

2.8.5 Skehan’s work in theorising aptitude

Early investigations of FL aptitude have had a mainly experimental focus and have not been theoretically linked to SLA processes. Without such a link, with its testable hypotheses, FL aptitude research has been neglected for many years. Skehan launched work in 2001 to focus on providing a theoretical foundation for FL aptitude in relation to the main stream of SLA. His initial work has been instrumental in building a link between aptitude components and the stages of SLA processing (See Table 2.4)

Table 2.4

*SLA Processing Stages and Potential Components (Skehan, 2002: p. 90)*

<table>
<thead>
<tr>
<th>SLA Processing Stages</th>
<th>Aptitude Components</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. noticing</td>
<td>Auditory segmentation</td>
</tr>
<tr>
<td></td>
<td>Attention management</td>
</tr>
<tr>
<td></td>
<td>Working memory</td>
</tr>
<tr>
<td></td>
<td>Phonemic coding</td>
</tr>
<tr>
<td>2. pattern identification</td>
<td>Fast analysis/working memory</td>
</tr>
<tr>
<td></td>
<td>Grammatical sensitivity</td>
</tr>
<tr>
<td>3. extending</td>
<td>Inductive language learning ability</td>
</tr>
<tr>
<td>4. complexifying</td>
<td>Grammatical sensitivity</td>
</tr>
<tr>
<td>5. integrating</td>
<td>Restructuring capacity</td>
</tr>
<tr>
<td>6. becoming accurate, avoiding error</td>
<td>Automatisation</td>
</tr>
<tr>
<td>7. creating a repertoire, achieving salience</td>
<td>Proceduralisation</td>
</tr>
<tr>
<td>8. automatizing rule-based language, achieving fluency</td>
<td>Retrieval processes</td>
</tr>
<tr>
<td>9. lexicalising, dual-coding</td>
<td>Automatisation , procedualisation</td>
</tr>
<tr>
<td></td>
<td>Memory, chunking, retrieval processes</td>
</tr>
</tbody>
</table>
In the first column of the table, the stages (or processes) of SLA were provided in order to relate each stage with putative aptitude components in the second column. Skehan outlined these relationships to explore whether each of the traditional aptitude components (i.e. Carroll’s MLAT) would be fit with each stage (or process) of SLA. Where there was this match, the existing aptitude components would be considered in the development of a revised aptitude construct, but if there is no aptitude component to fit with one of the stages of SLA processing, more aptitude components would need to be developed. This kind of outline helps researchers to diagnose the weaknesses and the strengths of current aptitude tests.

In light of Skehan’s analysis, it is noted that two components of MLAT fit with SLA stages:

1) the first stage of SLA, ‘noticing ability’ which proposes that to acquire any aspect of input, learners should pay some attention to that form. Skehan has noted that the phonemic coding ability of the MLAT test is relevant to this stage. Skehan goes further and indicates that additional research is needed to examine the impact of auditory segmentation, attention management and working memory in the Noticing stage.

2) For Stages 2-5 (Patterning), learners are supposed to detect the pattern in L2, and then be able to make a generalisation and then extension of the patterns. Skehan considers that in regard to Carroll’s aptitude test (MLAT), two components are involved in these stages of SLA: grammar sensitivity will be involved in the initial stages of patterning and the second component, inductive language learning ability, will be involved in more complicated stages of patterning.
However, Skehan also maintains that both components do not have sufficient contribution to completely account for what is happening in these stages. More aptitude subtests should be developed to measure the restructuring ability. Regarding the last stage, ‘lexicalising’, it is expected from the learner that they will have a very high linguistic competence and be able to use the lexical system quickly and efficiently. In this stage some memory component is involved but the rote memory component in MLAT is not adequate to be used to measure the memory ability because it is only associative memory based on encoding. The memory test needed in the future is a test which can identify and measure storage, retrieval and memory organisation.

After reviewing Skehan’s work, it is clear that the traditional aptitude concept is not outmoded; some of it can be fit to the main stream of SLA stages. This will be why the traditional measures have had the power of predictability that they have had. This initial attempt by Skehan of theorising aptitude is a promising area of future research, pointing to ways to develop current aptitude tests to capture the area of weakness in the traditional aptitude batteries instead of neglecting them as out-dated batteries.

2.8.6 Robinson’s research and work in theorising FL aptitude in different learning contexts

One of the recent central issues in IDs research is the impact of L2 context. Robinson was the first researcher who created a contextual link between L2 aptitude and the mainstream of SLA rather than dealing with aptitude as an isolated construct.

After a considerable amount of studies, Robinson (2002b) has provided a theoretical background for the relationships between multiple aptitudes and L2 learning conditions. Here, it is worth mentioning the concept of aptitude complexes which was initially
developed by Snow (1994). The notion of aptitude complexes, with hierarchies between factors, has been investigated by various researchers (Carroll, 1993; Snow, 1994). As a result of learners’ performance in various psychological tests, Snow was able to identify hierarchical models of these cognitive abilities. Robinson extended the theoretical work of Snow (Aptitude complexes), in his pioneering work, he provided empirical support for explanations of the nature of the relationships between various cognitive abilities in different L2 learning conditions (incidental, implicit and explicit learning). He proposed that there are different levels of cognitive resources: first-order level which includes some primary abilities such as pattern recognition, grammar sensitivity and rote memory (two components of MLAT). These primary abilities combine to construct the second level of cognitive abilities such as noticing the gap, memory for contingent speech (MCS), memory for contingent text (MCT), meta-linguistic abilities, advanced semantic operations and metalinguistic rule rehearsal. These second-level, higher cognitive abilities can be joined to form the aptitude complexes that can interact differently in various learning conditions. Robinson nominates four different kinds of aptitude: such as focus on form (via recast), incidental learning (via oral or written contents), and explicit rule learning (see Figure 2.1).
### Aptitude Complexes, e.g.

- **Aptitude for focus (on form)**
  - Aptitude for incidental learning (via oral content)
- **Aptitude for incidental learning (via written content)**
- **Aptitude for explicit rule learning**

<table>
<thead>
<tr>
<th>NTG-MCS</th>
<th>MCS-DSP</th>
<th>DSP-MCT</th>
<th>MCT-MRR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aptitude for focus</td>
<td>Aptitude for incidental learning</td>
<td>Aptitude for incidental learning</td>
<td>Aptitude for explicit rule learning</td>
</tr>
<tr>
<td>(NTG)</td>
<td>(MCS)</td>
<td>(DSP)</td>
<td>(MCT)</td>
</tr>
<tr>
<td>Noticing the gap</td>
<td>Memory for contingent speech</td>
<td>Deep semantic processing</td>
<td>Memory for contingent text</td>
</tr>
<tr>
<td>rehearsal</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Abilities factors, e.g.

- **Perceptual pattern recognition**
- **Phonological speed of WM capacity PWM**
- **Rote Speed**
- **Deep semantic processing**
- **Analogy Inferring**
- **WM word meaning for text**
- **Speed of WMT sensitivity**
- **Metalinguistic rule**

### Abilities test, e.g.

- **Inspective time**
  - (Anderson, 1992)
- **Sound Symbol**
  - (Sasaki, 1996)
- **Listening span test**
  - (Mackey et al., 2002)
- **Words in context**
  - (de Graaff, 1997)
- **Text memory**
  - (Harley & Hart, 1997)
- **Words in sentences/ paired associates**
  - (Carroll & Sapon, 1959)

---

*Figure 2.1 Hierarchical model of cognitive abilities for FL aptitude (derived from Robinson, 2002, p.119)*
The great contribution of Robinson’s work is that it does not only provide a theoretical foundation of FL aptitude but also provides promising practical implications in L2 teaching. Dörnyei (2005, p.60) has highlighted and reaffirmed the benefits of Robinson’s aptitude research by citing Robinson as follows:

“Profiling individual differences in cognitive abilities, and matching these profiles to effective instructional options, such as types of pedagogic tasks, interventionist ‘focus on form’ techniques, and more broadly defined learning conditions, is a major aim of pedagogically oriented language aptitude research.” ([Robinson, 2002] p.113)

Robinson’s work goes beyond Skehan, and into more detail, identifying a number of aptitude factors (‘speed of phonological Working Memory’; ‘Analogies’; ‘speed of working memory for text’) for which tests to measure have not yet been designed. Robinson’s framework, developed from Snow’s idea of complexes, and hierarchical levels, is complex. That complexity will explain, in part, why Carroll’s work has had some success, but not complete success. (Carroll has been measuring some of the ‘top-end’ of SLA). It will also explain why other alternatives to Carroll have had some success, but not complete success, and in some cases are not as effective as Carroll’s battery.

For memory components to be more effective in SLA, it should include more memory-related cognitive abilities such as storing and retrieval (Skehan, 2002).
2.9 Summary

Various theoretical and experimental issues related to FL aptitude were presented in this chapter. There issues include the following: 1) Discussion/definition of the FL aptitude concept; 2) The history of aptitude measurement instruments; 3) The basic components of the measuring instruments, 4) MLAT vs. PLAB; 5) Various issues related to aptitude testing; 6) Uses of aptitude testing results; and 7) Methodological approaches used in previous studies.

This review of previous studies and discussion of the main issues related to FL aptitude as understood by the field of applied linguistics identifies various main issues at the core of this thesis:

Firstly, the findings of previous studies related to the nature of the issue of stability of FL aptitude were controversial. Some studies support the stability of aptitude and others propose that aptitude can be modified. In this regard, the main weakness in the previous studies was the lack comparability; two different groups of participants were used to investigate the stability of aptitude. The current study overcomes this weakness and examines the stability of the foreign language aptitude test before and after the English course for the same participants.

Secondly, the cross-sectional and the longitudinal designs have been the most common methodological designs used to examine the relationships between FL aptitude and L2 proficiency. Almost all of the previous studies have built their findings on the Aptitude-L2 proficiency correlation by using the total scores of the FL aptitude measures and L2 proficiency measures. Few studies have investigated the relationship between FL aptitude and the progress in L2. In those studies concerning the absolute progress of L2, the
progress was measured as the differences between the values of L2 before and after the instruction. This study has been conducted to overcome the weaknesses in the past studies. In this study, like some, but different to most previous studies, both cross-sectional and longitudinal designs are used to investigate the relationship between the total scores of the FL aptitude measures and the total scores of English proficiency measures. Furthermore, the absolute progress is correlated with the foreign language aptitude test. The current study is novel in using the ratio of L2 progress instead of the absolute progress. The ratio of English progress is calculated and the obtained results are correlated with the foreign language aptitude test scores.

All of the obtained results for the cross-sectional design, longitudinal design, the absolute progress design, and the ratio of progress design are calculated and compared with each other in order that we may gain a clearer understanding of the nature of the relationships between FL aptitude and L2 proficiency.

After reviewing the empirical issues related FL aptitude, the recent theoretical trends of FL aptitude and SLA were discussed under the following sub-sections: 1) CANAL-FT (Cognitive Ability for Novelty in Acquisition of Language-Foreign) 2) Linguistics Coding Differences Hypothesis (LCDH). 3) Working Memory in SLA 4) Skehan’s work in theorising aptitude 5) Robinson’s Research and work on theorising FL aptitude in different learning contexts.

While focusing on the empirical, this study will also seek to develop a more inclusive understanding of the nature of FL aptitude. In considering the empirical results, explanations will explore various interpretations related to the Saudi context and the possibility of using the new foreign language aptitude test in predicting English success among Arab- native speakers. Since Carroll’s MLAT is the most effective tool for
measuring aptitude and since Skehan proposes that there is a promising relationship between Carroll’s test components and various and different stages of SLA, this study will try to create a link between the foreign language aptitude test components and the main stages of SLA. In light of Robinson (2002) ‘aptitude complexes’ notion, the current study will consider if there is a different pattern of correlations between different aspects of aptitude and the L2 learning situation in the Saudi context.
Chapter 3

METHODS

3.0 Introduction

The overarching topic of the thesis is the relationship between FL aptitude and English proficiency in Saudi learners of English as a foreign language (EFL) who are first-year university students at King Khalid University, a government tertiary institution in southern Saudi Arabia. In addressing this topic, a number of key issues are to be considered, such as stability of FL aptitude as a trait, the relationship between FL aptitude and FL proficiency cross-sectionally and longitudinally, as well as the relationship between FL aptitude and progress in FL proficiency.

The principal goal of this chapter is to provide a comprehensive description of the issues related to the research protocol deployed in this project. In order to accomplish this, the following sections consider the main dimensions of the research protocol such as the general design of the study, the participant-recruitment procedure, as well as details about the demographics of the sample, research instruments, and data-collection and processing procedures. The chapter concludes with a brief statement of the study’s limitations.

3.1 Research questions and hypotheses

The main aim of this study is to assess whether the foreign language aptitude test instrument predicts English proficiency during a seven months of L2 English learning in a class, using a design that examines the relationship between FL aptitude and English proficiency, cross-sectionally and longitudinally.
In the current study, four questions will be addressed to achieve the main objective of the study:

1. Is the foreign language aptitude testing instrument in the current study a valid test, at a level similar to other equivalent instruments?
2. Do the foreign language aptitude test scores at the beginning of the English course (Time1) differ from the foreign language aptitude test scores at the end of the seven-month English course (Time 2)?
3. Does the foreign language aptitude test predict English-language learning proficiency to a similar standard as other aptitude tests?
4. Do the foreign language aptitude test scores correlate significantly and positively with English proficiency progress after a seven-month intensive English course?

3.2 Study design

3.2.1 Overview

After the review of many FL aptitude studies in the literature chapter in this study, it is clear that the majority of those studies used a cross-sectional design to investigate the relationship between FL aptitude and second-language achievement. In contrast, few longitudinal studies were conducted in the area of FL aptitude-L2 proficiency research. In this section, the properties and benefits/weaknesses of each design will be discussed. Then, further details will be provided as to the reasons for choosing the longitudinal design.
3.2.2 Cross-sectional studies

Cross-sectional design involves testing the correlation between two or more independent variables in the same sample and their relationship with any dependent measures at one occasion. This approach provides a snapshot of information about the relationships between variables in the study at a specific period of time.

The cross-sectional design has various advantages that have motivated many researchers to use it. These advantages can be summarised as follows: 1) Data can be collected from a large number of participants at the same time; 2) Many variables can be included in this design; 3) Data can benefit many researchers; 4) Cross-sectional design can generate hypotheses for future research. The greatest advantage is that it is the most economic design because it involves less time and effort.

Although the cross-sectional design has all of these benefits, it has been criticised for only showing the relationships between variables and not showing the developmental pattern in the sample. Further, it does not measure changes in the values of variables during different periods of time. In addition, the cross-sectional design has been criticised for having only a limited capacity for establishing a cause-and-effect relationship.

3.2.3 Longitudinal studies

Longitudinal design can be defined as follows: 1) Data are collected at two or more points in time; 2) Participants are the same in all periods of time; 3) The data in each period of time can be compared to data in another period of time (Menard, 1991). This design is used in many studies for three main purposes. First, longitudinal investigations map the process of changes in variables. Second, this approach can be used to investigate various factors among the sample and their stability over time (McCall, 1977). Third, longitudinal
design offers findings that can contribute to the development of hypotheses about causality.

Despite the common use of the longitudinal design, it has some disadvantages. For instance, data collection is a costly process that also requires a lot of time and effort. There are possible solutions to this problem such as reducing the sample size or the number of visits to collect the data, but, these solutions are limited because most correlation studies involve large size samples in order to have sufficient statistical power. Using this design also involves other difficulties such as maintaining the same participants over the duration of data collection and the possible risk of loss of data.

Most previous attempts to investigate the relationship between FL aptitude and second-language success have used cross-sectional correlation which only show the correlations among variables in one occasion, resulting in insufficient findings for testing causal explanations. Far fewer studies in which FL aptitude and English proficiency tests are repeated across time have been conducted (Sparks et al., 2009). Longitudinal studies allow the observation of change and trends, and so deliver more information on which to base causal inferences.

The combined cross-sectional and longitudinal design allows for a wider range of comparisons and allows cross-checking to strengthen the evidential support for the conclusions drawn. The comparisons available are provided in Figure 3.1.
3.3 Study setting

At the time of enrolment in the intensive language program, all students had studied English in high school for six years. Each and every Saudi university student will have been exposed to L2 instruction during his or her high school studies. The first exposure to English occurs in grade 7. In Saudi public schools, English is taught as a compulsory subject, in 45-minute sessions, four-times a week, by non-native speakers of English. The high-school English textbooks and tests mainly focus on grammar and vocabulary. Unfortunately, after finishing high school, most Saudi students, even those with high English scores, have insufficient knowledge about communicative English. This can be seen as a common phenomenon among EFL learners, especially in Asia, where students have a lot of information about the structure and lexis, but little knowledge about how to use English as a communicative tool (Fujii, 2005).

The current study took place in the Southern region of Saudi Arabia; more specifically, the study was conducted in the Languages and Translation College in King Khalid
University. Approximately 500 male and female students are enrolled in this college for a four-year Bachelor program in English.

The university policy sets some criteria for new students enrolling in the Languages and Translation College. No English placement tests are administered during the first year; the new students must satisfy strict selection criteria. First, they must have high scores in the General Cognitive Test for Saudi students, a national test designed by the Ministry of Education and Learning in Saudi Arabia for students who want to enroll in Saudi Universities. Second, students need (no lower than 90 per cent) Grade Point Averages (GPAs) in high school.

At the beginning of the first year, students are divided randomly into seven sections. Each section includes 20-30 students. Four intensive courses are provided for the first-year students: Reading Skills, Listening and Speaking, Writing Skills and Grammar. The students study sixteen 50-minute English periods per week. By the end of the program (two semesters), students are expected to have studied English for about 653 hours. It is worth mentioning that English, in Saudi Arabia, is a foreign language and the opportunity to use English outside classrooms is rare. To increase the length of exposure to English, most Saudi universities, if not all, provide an intensive English course in the first year of enrolment.

The Languages and Translation college in King Khalid University aims to achieve various key objectives for its students, especially during the first year, through courses and intensive programs. Firstly, the major goal is to develop the four essential skills of basic English literacy (listening, speaking, writing, and reading) through intensive courses. By measuring these skills, students’ linguistic competence to communicate inside and outside the academic atmosphere is greatly enhanced. Secondly, the intensive English program in
the college is designed with the specific aim of increasing first year students' comprehension of lectures in English. To achieve these aims, a Tapestry Series, edited by Rebecca Oxford in 2000, is taught in the first year. This series focuses on the four skills (listening, speaking, writing, and reading) in three books for four levels: Tapestry Listening and Speaking 1-4, Tapestry Reading 1, and Tapestry Writing 1. They are designed to equip students with sufficient communicative skills to use English inside and outside classrooms.

Most teachers are limited in using these materials without designing further ones to suit the specific student needs or the specific purposes of the class. In the Saudi context, teachers are the center of the learning process. They are responsible for covering the materials, which are pre-selected by the curriculum committee in the college. Different classes are taught by the same non-native English teachers. All teachers have to use the same textbooks and cover the same materials for all the students. In this case, their duty, to a great extent, is to prepare the students for the final unified exams at the end of the academic year. From the students’ perspective, the main goal is to study English in order to pass the final exams. In other words, most Saudi students studied only to pass their exams without the intention of using English for other purposes (i.e. communication). The most obvious weakness in the delivery of this program is neglecting teaching speaking. As a result, most Saudi students face a difficulty in using English to communicate.

3.4 Participants

3.4.1 Sample selection

The sample in this study included all students enrolled in semester 1, 2008, in the Languages and Translation College. The participants in the current study were selected through a non-random sampling method called “Convenience sampling”, which is based
on selection of participants who are available for the study (Mackey and Gass, 2005). It is a commonly used method because it saves time and effort. In the case of this project, the non-random sampling method was the most convenient method for various reasons. The targeted sample was limited to English students in their first year at university and it included only males. The longitudinal nature of the research involved collecting data from the same participants on different occasions during the English course. All of these circumstances made the non-random sampling method the best choice for this project. It needs to be noted, however, that results derived through the non-random sample may not easily yield themselves to generalisations about non-university-student populations of EFL learners in Saudi Arabia. More explanations for the data collection procedures are provided in Section 3.7.

3.4.2 Participant sample

The initial sample consisted of 90 male Saudi university students majoring in English. The participants were in the first semester in the Languages and Translation Department in King Khalid University in Abha (a city in the Southern region of Saudi Arabia), and they were aged between eighteen and 20 years at the time of enrolment. In this study, the researcher was unable to include females because the education policy in Saudi Arabia states that women be segregated from men in all government universities (for more details related to gender impact on L2 success see Chapter Two, Section 5.6). Based on the experience of the researcher as a member of the participants’ society and a teacher in the same department, it can be said that all participants share the same cultural, social, and ethnic backgrounds and speak the same mother tongue.

The participants’ exposure to English was also limited. None of the 90 participants had been involved in any school where English was the medium of instruction. They had all
attended public schools where the medium of instruction was in their native language, Arabic. They had studied English as a subject in secondary school, but only to the standard curriculum level set for minimum schooling requirements. Furthermore, no one had been involved in any trips to English-speaking countries or had been exposed to additional English teaching programs. In other words, all participants had had a similar amount of exposure to English. Relevant literature acknowledges that EFL learners in Saudi Arabia do not get exposure to English outside of the language classroom, which is recognized as a serious obstacle to the attainment of adequate proficiency in that language.

3.4.3 Recruitment of participants

Prior to commencing the study, ethical clearance was sought from The University of Newcastle where the researcher is doing his PhD. After obtaining approval from The University of Newcastle and receiving the signed consent forms from the Dean of Languages and Translation College in King Khalid University, the researcher began data-collection procedures.

In the first week of September 2008, the researcher visited King Khalid University’s Department of Languages and Translation, where he conducted his study. He introduced himself to the students in their classes at the beginning of the English course. He then distributed all the consent forms among the participants and explained, in detail, the information related to the purpose of his study. He also reassured potential participants that all information disclosed by participants, in the tests and in the demographic survey would be kept confidential. To motivate students’ cooperation, the researcher explained the importance of their participation in this study in enhancing English teaching and learning in Saudi universities. They were also informed that there would be no pass or fail
in response to the tests and they had the right to stop completing the tests if and whenever they liked. At the end of this visit, participants were given a week to read the consent forms and to decide whether or not to participate. One week later, the researcher collected the signed consent forms.

3.5 Foreign language aptitude test instrument

3.5.1 Constructing the foreign language aptitude test

This is a leading study in using an Arabic version of an aptitude test. No foreign language aptitude test has previously been designed to measure the language-aptitude abilities for Arabic native speakers. Before conducting any aptitude test in the current study, three options were available for the researcher: 1) To use the original MLAT (English version) which was originally designed for English native speakers; 2) To adapt the MLAT to be used specifically for Arabic students; or 3) to design a new aptitude test for Arabic speakers.

The first option, using MLAT (English version) was rejected because of its potential threat to the validity of the aptitude test. Aptitude tests were developed to measure learners’ aptitude in language abilities, not L2 competence. Carroll’s empirically-based and standard-setting work, in English, was for native English speakers. Chapter 2 of this study has discussed the reasons why MLAT (in English) is not appropriate for non-native speakers of English (see Section 2.4.1.4).

Between the remaining two options: to adapt an existing test or to design a new one, the researcher chose to adapt the MLAT (short form) to be used as a foreign language aptitude test with the same aptitude components as in Carroll’s model. This short form (Parts 3, 4, and 5) was administered in the current study instead of the long form which
includes 5 parts. As shown in the *MLAT Manual*, both forms yield almost the same predictive power for the same target population.

Adaptation of cognitive tests such as aptitude tests from one language to another has been a widespread practice in education, psychology, and applied linguistics research (Rysiewicz, 2008). Hambleton (2006), as cited in Rysiewicz (2008), points out that there is a growing belief in the importance of increasing the number of aptitude tests with different languages in the future. For instance, Ottó (2002) developed the first Hungarian General Aptitude Test. Another example can be found in Sasaki's study in 1993 in developing the Language Aptitude Battery for Japanese (the LABJ). The adaptation of the aptitude test in the current study was motivated by the arguments in favour of such an adaptation provided by Hornowiska and Paluchowski (2004) and Hambleton (2006), as cited in Rysiewicz (2008). First, test adaptation is much easier and less expensive than designing new tests. Second, there is a lack of national experts in designing new aptitude tests in their contexts. Third, it is better to use a well-known tool that has been used in many studies rather than developing new ones. If a well-known tool had not been used, there would have been much less research (almost none) to indicate other findings with which to compare the results, and nothing to substantiate any comparison related to aptitude stability or lack thereof. Furthermore, adaptation tests increase the possibility of comparison between the findings of adapted tests in one country with the findings of well-known tests in another country. The detailed rationale for preferring MLAT over other FL aptitude tests has been discussed in Chapter 2.

Based on the Stansfield and Reed (2003) guidelines for the MLAT adaptation for other languages, and Rysiewics (2008) in his adaptation of the MLAT to Polish L1 speakers, the same procedures were applied to adapt the MLAT to develop the foreign language
aptitude test, which echoes the short form of MLAT (Part 3: Spelling Clues, Part 4: Words in Sentences, Part 5: Paired Associates). Two ways of adaptation were used to adapt the three components of the MLAT (short form): either paraphrasing the construct of the original tasks, or translating the contents of the original task using the new language material and maintaining the same forms of the original task. In this study, Spelling Clues and Words in Sentences tasks were adapted by paraphrasing the content of the original tasks by using Arabic constructs. The last task, Paired Associates was translated into Arabic without changing the forms and the instructions of the original MLAT (see Table 3.1). Furthermore, shortening the original tasks by reducing the number of items (in Spelling Clues and Words in Sentences) was another adaptation strategy used by Rysiewicz (2008) in his study to adapt the MLAT for Polish native-speakers, and followed in the Arabic adaptation used in this study.

Table 3.1

MLAT and Foreign Language Aptitude Test Tasks

<table>
<thead>
<tr>
<th>MLAT</th>
<th>Foreign Language Aptitude Test</th>
<th>Adaptation Method</th>
<th>Target Skill</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spelling Clues</td>
<td>Vocabulary and Writing coding</td>
<td>Constructs translation from the original MLAT to Arabic</td>
<td>Phonetic coding and Arabic vocabulary</td>
</tr>
<tr>
<td>Words in Sentences</td>
<td>Grammar Sensitivity</td>
<td>Paraphrasing the constructs on the original MLAT to Arabic</td>
<td>Grammar sensitivity</td>
</tr>
<tr>
<td>Paired Associates</td>
<td>Memory</td>
<td>Translation</td>
<td>Rote memory</td>
</tr>
</tbody>
</table>
After these types of adaptations, two copies of the adapted foreign language aptitude test were given to two English-Arabic bilingual speakers to be translated back into English. Then, the translated copies were handed to another two English-Arabic bilinguals to translate them back into Arabic. The researcher selected the items that had the same constructs in both languages.

3.5.2 Piloting the Arabic version of the MLAT

To examine the appropriateness of the foreign language aptitude test to the target population, the researcher randomly selected ten male university students, as a representative sample of the target population, to participate in a pilot study. The 45-minute pilot version of the foreign language aptitude test was administered in the second week of September 2008 in King Khalid University. Participants for the test were recruited during in class. For instructions and administrations of the test, the MLAT Manual of 2002 was used. Then the answer sheets were collected.

On the whole, the piloted foreign language aptitude test showed acceptable results. These results were analysed by using Statistical Package for the Social Sciences (SPSS) software. Table 3.2 shows that each individual task of aptitude was significantly correlated with the total score of aptitude tests. These correlations range from .60 to .70. It can be noted also that all aptitude components were significantly correlated with each other except the correlation between Vocabulary and Written coding (VWC) task and Memory task. As a whole, the results indicated that the current aptitude tasks were measuring different components of the FL aptitude construct.
Table 3.2

Correlations of Foreign Language Aptitude Test Components (N=10)

<table>
<thead>
<tr>
<th></th>
<th>VWC</th>
<th>Grammar Sensitivity</th>
<th>Memory</th>
<th>Aptitude</th>
</tr>
</thead>
<tbody>
<tr>
<td>VWC</td>
<td>1.00</td>
<td>.38**</td>
<td>.10</td>
<td>.60**</td>
</tr>
<tr>
<td>Grammar Sensitivity</td>
<td>.38**</td>
<td>1.00</td>
<td>.28*</td>
<td>.69**</td>
</tr>
<tr>
<td>Memory</td>
<td>.10</td>
<td>.28*</td>
<td>1.00</td>
<td>.79**</td>
</tr>
<tr>
<td>Aptitude</td>
<td>.60**</td>
<td>.69**</td>
<td>.79**</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Note. * p < .05 , **p < .01

3.5.3 The revised foreign language aptitude test

After piloting, some minor changes were made to suit the target sample. These changes related to the time allowed for completion of the test and to the type of items. The time was reduced to 30-minutes instead of 45-minutes; because of the time limit for the period it was difficult for the researcher to administer the longer test. For the memory test, the time for memorising the 24-word list was reduced to two minutes instead of three, because in pilot study, this part was too easy for most of the participants. The final components of the foreign language aptitude test are shown in Table 3.3 and see also Appendix 2 for the Arabic version. An English translation of this test is available in Appendix 3.

1. Vocabulary and writing coding errors (diacritics)

This task represents the Arabic equivalent of the Spelling Clues subtest in MLAT. It aims to measure the students’ recognition of synonyms, a particular ambiguity in English because of spelling conventions. (Note that Arabic is phonetically spelt, and diacritics carry meaning information, which, if absent, can be derived from context). In this task, students are presented with 20 Arabic words without their significant diacritics.
words have different meanings when various diacritics are added. The students were asked to examine different phrases and to choose the phrase that best matches the meaning of the provided words without the relevant diacritics. In this part of the test, participants were given ten minutes to complete the task.

Example:

1. Hot (with diacritics, it means a neighbor)
   - حار
   a. A chemical substance
      - عنصر كيميائي
   b. A person who lives nearby.
      - الشخص الذي يسكن بالجوار
   c. A kind of perfumes
      - نوع من العطور
   d. A kind of furniture
      - نوع من الأثاث

2. **Grammar Sensitivity**

The Grammar sensitivity subtest in both tests (Foreign language aptitude test and original MLAT) was designed to measure the learners’ analytic ability. This part is closely modelled on Part 4 of MLAT: *Words in Sentences*, which is the most popular subtest to measure students’ analytic abilities (Carroll and Sapon 1959; Skehan 1989). In this part of the test, students were provided with 20 pairs of Arabic sentences. The first sentence in each item contained an underlined word. From the second sentence, students were asked to choose from five underlined words the word that has the same grammatical function as the underlined word in the first sentence. Ten minutes were allocated to students to finish this task.
For Example:

Sarah cuts THE APPLE with a knife.
My brother, Ahmad, hits the horse with a long stick.

A            B          C          D                          E

3. Memory

The Memory subtest in the adapted foreign language aptitude test was designed to echo the MLAT-Part 5 Paired Associates subtest. It was administered to measure the students’ memory. In this subtest, students had to study a list of 24 pairs of Kurdish-Arabic words for two minutes. As the original MLAT, the ‘Kurdish’ language used in this test was a nonsense language which is unrelated to the real Kurdish language. Then, the students were asked to practise writing the Arabic meanings for the Kurdish words for three minutes. During these practice trials, the students were allowed to check their answers from the list, if they wished. Then, 24 multiple-choice test items were provided. Each item consisted of one Kurdish stimulus and five Arabic answers, only one of which was correct. Five minutes were allowed for students to finish this task.

After reviewing all the components of the foreign language aptitude test, it can be summarised in the following table:
Table 3.3

Summary of the Components of the Foreign Language Aptitude Test

<table>
<thead>
<tr>
<th>Task</th>
<th>Aims of the Task</th>
<th>Type of Task</th>
<th>No. of Items</th>
<th>Scores</th>
<th>Time Allowed</th>
</tr>
</thead>
<tbody>
<tr>
<td>VWC</td>
<td>Measuring the students’ Phonetic Coding ability</td>
<td>Multiple choice</td>
<td>20 items</td>
<td>20 points</td>
<td>10 minutes</td>
</tr>
<tr>
<td>Grammar Sensitivity</td>
<td>Measuring the learners’ analytic ability</td>
<td>Multiple choice</td>
<td>20 items</td>
<td>20 points</td>
<td>10 minutes</td>
</tr>
<tr>
<td>Memory</td>
<td>Measuring rote memory</td>
<td>Multiple choice</td>
<td>24 items</td>
<td>24 points</td>
<td>10 minutes</td>
</tr>
<tr>
<td>Total no. of items</td>
<td></td>
<td></td>
<td>64 items</td>
<td>64 points</td>
<td>30 minutes</td>
</tr>
</tbody>
</table>

3.5.4 Issues associated with the adaptation process

Arabic is distinctly different to English in a number of ways: Arabic script is different; Arabic written text is read from right to left; Arabic spelling is phonetic; Arabic sentence construction is generally object-verb-subject. While undertaking the adaptation of MLAT to Arabic form in the current research, the difference between Arabic and English, in spelling systems, became clearer, and represents an issue that should be considered while adapting the Carroll’s Spelling Clues component of the aptitude test. Arabic’s spelling is phonetic which means that each word is pronounced the way it is spelled, and is unlike English written coding in that regard. In the Arabic writing system, short consonants and
long vowels are represented with letters, but short vowels and the length of consonants are not indicated in writing. Because of this feature in the Arabic writing system, the same root can have different meanings when different diacritics are added. In Carroll’s *Spelling Clues* component, homonyms like ‘presents’ were used. The task of the test was to examine the context of the ambiguous term in order to identify a synonym offered in the multiple choice options. In the current study, the Vocabulary and Written Coding component is an appropriate response to cover what cognitive tasks are involved when recognising ambiguities in written coding of words to be used with phonetic languages such as Arabic. This new adapted component (VWC) was used in this study because in Arabic some spelling errors can occur as a result of ‘Visual letter-confusion errors’ when a learner can be confused by the similar visual shapes of letters (Abu-Rabia and Taha, 2006) and needs to use other clues to decode the meaning intended.

3.6 Other instruments

In this study, the researcher administered a social demographic survey and foreign language aptitude test and an English proficiency test to examine the relationship between FL aptitude and second-language proficiency among Saudi university students in the current sample.

3.6.1 Social demographic sheet

A Social Demographic sheet was provided at the first session in September 2008 which included questions seeking information about age, level of English proficiency, length of learning English, students’ score in English in the final year of high school, their GPAs in high school. A more extensive survey to address some of the socio-economic factors, that Skehan (1989) and Sparks et al., (2006) have raised, was not considered relevant at this stage of the study.
3.6.2 English proficiency test

In this study, the Test of English as Foreign Language (TOEFL) was administered to evaluate students’ English proficiency in three main skills: Listening, Grammar and Reading. The TOEFL was initially developed by the Educational Testing Services in the United States in 1963 by the National Council on the Testing of English as a Foreign Language. It aims to evaluate English proficiency of non-English speakers (Rosenfeld, Oltman et al., 2004). This test has been argued to have the best predictive validity among other English proficiency tests for non-native speakers of English (Hsieh, 2004). Nowadays, every year more than 50 million tests are administered in more than 80 countries at 9000 centers around the world. In this study, the head of the English Program Unit and the course teachers revised the test items and the time allowed for the test. Based on their recommendations, some items were modified to suit the population culture and others were eliminated because of the time limit (see Appendix 2 for the copy of the test, and Appendix 3 for a translation into English.).

The proficiency test consists of 60 multiple-choice items, divided into three sections (20 items in listening, 30 items in grammar and 10 items in the reading part). The following section includes detailed explanations of the proficiency test components:

1. **Listening Skill Test:**

This part of the test consists of 20 items. It aims to measure students’ ability to comprehend spoken English. A tape recorder is used for this subtest. Before starting this subtest, students were orally advised to listen carefully to the test instructions. The researcher informed the participants that once the tape recording started, it would not be stopped until the end of the test. Then, students were asked to listen to short conversations between two speakers in the tape recorder. Each short conversation was followed by a
question. Students were asked to choose the correct answer and write it down in the listening section of the answer sheet. Students were asked to do the same for all items. While listening, the students were not to take notes or to talk. The time allowed was ten minutes.

On the recording, the examinees heard:
(Man)        That exam was just awful.
(Woman)   Oh, it could have been worse.
(Man)        What does the woman mean?
In the test book, they read:

A. The exam was really awful.
B. It was the worse exam she had ever seen.
C. It couldn’t have been more difficult.
D. It wasn’t that hard.

The best answer to the question, “What does the woman mean?” is (D), “It was not that hard.” Therefore, the correct choice is (D).

2. Grammar and Structure Test:

Part 1: Structure

This measurement includes 20 multiple-choice items. The purpose of this subtest is to measure the students’ grammatical and structural abilities to be used in formal written English. The students were provided with 20 incomplete sentences and they were asked to choose the word or phrases that correctly complete the sentence. Students were asked to provide the answers for 20 items and write the answers in the answer sheet. The time allowed for this task was fifteen minutes.

Example:

The president ___________ the election by a landslide.

A. won
B. he won
C. yesterday
D. fortunately
The sentence should read, “The president won the election by a landslide.” Therefore, the answer is (A).

Part 2: Written Expression

This part of the test aims to test students’ structural and grammatical knowledge in written English. Ten sentences were provided. Each sentence had four underlined words or phrases. The students were asked to choose the ungrammatical word or phrase in each sentence. After choosing the appropriate answer, they had to put their answers in the section provided in the answer sheet. The time allowed for the grammar and structure test was fifteen minutes.

Example:

The research for the book Roots taking Alex Haley twelve years.

A                     B                     C                                D

The sentence should read, “The research for the book took Alex Haley twelve years.” Therefore, the examinee should choose (C).

After choosing the appropriate answers, students had to put their answers in spaces provided in the answer sheet. The time allowed for the grammar and structure test was fifteen minutes.

E. Reading Comprehension

The Reading Comprehension subtest consists of ten items. These items aim to measure the students’ abilities to read and understand short English passages. In this part of proficiency test, students were given two short passages. Each passage was followed by 5 questions. The students were asked to read the passages and choose the best answer for each item. After choosing the appropriate answer, students had to fill the spaces provided in the answer sheet. The time given to read the two passages and to answer the ten questions was 20 minutes.
### Table 3.4

**Summary of the Components of the English Proficiency Test**

<table>
<thead>
<tr>
<th>Task</th>
<th>Aims of the Task</th>
<th>Type of Task</th>
<th>No. of Items</th>
<th>Scores</th>
<th>Time Allowed</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Listening Skill Test</strong></td>
<td>Demonstrating students’ English ability to comprehend English talks and conversations</td>
<td>Multiple choice</td>
<td>20 items</td>
<td>20 points</td>
<td>10 minutes</td>
</tr>
<tr>
<td><strong>Grammar &amp; Structure Test</strong></td>
<td>Examining students’ structural and grammatical knowledge in written English</td>
<td>Multiple choice</td>
<td>30 items</td>
<td>30 points</td>
<td>30 minutes</td>
</tr>
<tr>
<td><strong>Reading Comprehension</strong></td>
<td>Measuring the students’ abilities to read and understand short English passages</td>
<td>Multiple choice</td>
<td>10 items</td>
<td>10 points</td>
<td>20 minutes</td>
</tr>
<tr>
<td><strong>Total numbers</strong></td>
<td></td>
<td></td>
<td>60 items</td>
<td>60 points</td>
<td>60 minutes</td>
</tr>
</tbody>
</table>

#### 3.7 Data collection procedures

The data were collected in two phases over a seven-month interval. The first occasion was in September 2008 at the beginning of the English course, and the second time was seven months later, in March 2009 at the end of the same course. All data were collected during class sessions. Three hours were required to administer all tests and surveys in both times (Time1 and Time 2). All testing administrations were supervised by the researcher himself to make sure that all testing procedures were reliable. More details about data-collection procedures follow
3.7.1 Data collection: Time 1

In September 2008, after getting the signed consent forms, the researcher gathered the students who agreed to participate in the study in a classroom separate from other students. Before the survey and tests administration, the researcher considered the standards of the testing conditions related to the physical space. The room was well lit and ventilated. Also, it was spacious enough to allow good spaces between examinees, copying would not be an issue. The sound properties were good: the researcher’s voice was heard clearly and there were no external noises. Then, participants were handed a demographic survey that included information about age, level of English proficiency, years of learning English, English scores in high school and GPA. After filling out the form, the foreign language aptitude test and English proficiency test were administered during a two-hour session. Following the same timing procedures in MLAT and TOEFL, both tests required 90 minutes to be completed (MLAT: 30 minutes and TOEFL: 60 minutes). To avoid the effect of fatigue and tiredness, the students were given a five-minute break between the two tests.

1. Foreign language aptitude test

The foreign language aptitude test was administered for one session (30 minutes), following the same MLAT administration as that recommended by Carroll and Sapon in the 2002 MLAT Manual. The test materials were distributed: each examinee was given a test booklet and an answer sheet. All the directions were explained to the examinees in Arabic. Table 3.1, above, provides more details about test components, aims of tasks, number of items, and timing and scoring. At the end of the test, the answer sheets and the test booklets were collected by the researcher.
2. *English proficiency test*

The English proficiency test was administered during one session (60 minutes). The participants were given test booklets and answer sheets. The time allowed for the test and its instruction was 60 minutes.

3.7.2 *Data collection: Time 2*

Seven months later, in March 2009, the second round of data collection took place for the same target population. Under the researcher’s supervision, and following the exact tests procedures as at Time 1, the same foreign language aptitude test and the same English proficiency test were administered again. A summary of the data-collection procedures and the materials used is provided in Table 3.5.

Table 3.5

*Summary of Data-Collection Procedures*

<table>
<thead>
<tr>
<th></th>
<th>Date</th>
<th>Material</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>First Visit</strong></td>
<td>First week of September 2008</td>
<td>Consent forms</td>
</tr>
<tr>
<td></td>
<td>Second week of September 2008</td>
<td>Pilot study</td>
</tr>
<tr>
<td><strong>TIME 1</strong></td>
<td>Third week of September 2008</td>
<td>Demographic survey</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Foreign language aptitude test</td>
</tr>
<tr>
<td></td>
<td></td>
<td>English proficiency test</td>
</tr>
<tr>
<td><strong>TIME 2</strong></td>
<td>Third week of March 2009</td>
<td>Foreign language aptitude test</td>
</tr>
<tr>
<td></td>
<td></td>
<td>English proficiency test</td>
</tr>
</tbody>
</table>
By the end of the data-collection procedures, the researcher had data from 90 participants. These data included the scores of foreign language aptitude tests and the English proficiency tests at both occasions (Time 1 and Time 2).

3.8 Data preparation

The data were checked to assure that all the test items were readable, clear, and completed. The researcher also made sure that all the collected data including the foreign language aptitude test and the English proficiency test for both occasions (Time 1 and Time 2) were classified correctly in their groups.

3.8.1 Foreign language aptitude test scoring

The foreign language aptitude test consists of 64 multiple-choice items. It was divided into three sections: Vocabulary and Written Coding (20 items), Grammar Sensitivity (20 items) and Memory (24 items). The total score of the whole test was 64 points. Before scoring, the researcher checked all the answer sheets to avoid double answers for the same item. If any item was given more than one answer, that item was omitted from the scoring. The researcher followed the same scoring techniques as those in MLAT Manual of 2002: each correct answer was equivalent to one point (Carroll, 2002). The test was marked by the researcher himself and was then, to ensure that all test sections were marked correctly; it was independently checked by one of the teachers in the department who did not teach the course. In case of mismatching between the researcher’s scores and the teacher’s scores, both of them reviewed the mismatched ones to reach 100 per cent agreement because there was only one correct answer for each item.
3.8.2 English proficiency test (TOEFL) scoring

This test was administered twice to the participants to measure their English proficiency in three main skills: listening, grammar, and reading. The test includes 60 multiple-choice items distributed among three skills as follows: listening (20 items), grammar (30 items) and reading (10 items). One point was given to each correct response and the total score was 60 points (listening 20 points, grammar 30 points, and reading 10 points). In case of two choices for the one item or no response for any item, the answer was considered incorrect. Two copies of participants’ answer sheets were made. One copy was marked by the researcher and the other was given to an independent English teacher. All answer sheets were marked according to the TOEFL scoring. All scores of the foreign language aptitude tests and the TOEFL tests were listed on two separate sheets. After finishing the scoring, the two lists of scores were compared. The mismatched scores were revised by the researcher and the English teacher to ensure all scores were correctly addressed.

3.8.3 Data screening

Before subjecting the raw data to statistical tests, some data-preparation procedures took place. These procedures addressed various issues such as handling of missing data and data classification.

Although 90 students who were enrolled in the first-year program took the aptitude and the proficiency tests, only the data for 56 of these was used in the main analyses. The selected participants were those who had the same length of prior exposure to English and who did FL aptitude tests and English proficiency tests in both sessions (Time 1 and Time 2). In order to control the possible impact of various external factors and to maintain the data of participants who shared the same exposure to English, the partial data of 43
participants (34 at Time 1 and nine at Time 2) were excluded because they were either absent during one of the data-collection sessions, or dropped the course, or were taking the course for the second time. In relation to the data that were excluded from the analysis it should be clarified that the nine participants at T2 were not in addition to the 34 participants at T1. So the overall number of participants excluded was 34, leaving 56 participants whose data were included.

The problem of participants leaving the study before the end of the data collection is one of the most common disadvantages of longitudinal data models (Frees, 2004). Consequently, the data of 56 participants were selected to be analysed in this study. The attrition data analyses are discussed in more detail in the following section.

3.8.3.1 Attrition data analyses

In this section, the SPSS attrition analyses were used to ensure that the data of non-retained students did not differ from the data of retained students in the main study (see Appendix 3). In this respect, the multivariate statistical tool was used to identify the individual impact of each subgroup of participants within the sample study. First, the data of the whole group of participants on both occasions (Time 1 and Time 2) were divided into two main groups (retained and non-retained participants). The statistical analyses were used to investigate whether the results of the retained participants differed from that of the non-retained participants. The multivariate F showed that participants who were excluded from final analyses for a variety of reasons are not statistically different from those who were retained (see non-significant Multivariate F for partstatus variable and univariate Fs for individual components). There is an isolated significant difference on Listening at T1 where the data indicated that the retained students do significantly better, but overall there is no sizeable sign of group differences at T1 and T2.
From more specific analysis, the data of the participants in Time 1 divided into subgroups based on the reasons of their selection: retained, absent, failed, and dropped. The multivariate statistical test was used to examine the impact of each subgroup. The multivariate analyses reveal that there is no multivariate group difference (see Appendix 3: Multivariate Test, F for Partstatus T1 variable). This indicates that overall, the individuals who contributed to the main analyses are not statistically different from those who did not contribute to them for various reasons. This is generally good as it removes concerns over selection bias. In a more finely grained look at results separated into the subcomponents of aptitude and proficiency (tests of between-subjects effects = univariate tests), we see some significant group differences on memory, and aptitude total.

We then look at descriptive and multiple comparisons along these three variables only to discover the following results:

Memory (from highest score to smallest): dropped, absent, retained, and failed
   Dropped is significantly different from both failed and retained, but these two latter are not significantly different from each other.

Aptitude total (from the highest score to smallest): dropped, retained, failed, absent
   Dropped is significantly (or marginally) different from both failed and retained, but the latter two latter are not significantly different from each other.

Also, importantly, the absent students are never statistically different from the retained students; also addressing our concerns of selection bias.

The more detailed attrition analyses showed that the dropped students ($N=9$) outperformed other groups of participants on overall aptitude and memory. This could not be clearly explained in the current study because our data were so limited: more detail about any common characteristics is necessary in order to provide a sufficient explanation. These nine participants did not complete the course for various reasons, such as getting a job or transferring to another college.
To summarise, beside these isolated differences on two of the univariate tests, overall, the attrition analysis suggests no selection bias in the main sample (see multivariate F); hence, there is no reason to suspect that the main results would have changed had all the participants who provided T1 also provided T2.

3.8.3.2 Missing data

Missing data is one of the common problems in data management. Its occurrence can be a result of different reasons, such as participants’ absence during the time of data collection, dropping the course, unwillingness to answer questions, or lack of motivation to cooperate. The researcher encouraged participants to complete all the test items while the tests were being administered in order to avoid the negative impact of missing data in the current study. This procedure helped to decrease the number of missing data, but some missing values still appeared.

Since the nature of the tests’ items’ design was to gather correct or incorrect responses, the researcher dealt with each of the missing values as an ‘incorrect response’. In the cases of missing data as a result of absence during the tests’ administration or carelessness in completing the whole test, the data for the participant was eliminated, rather than being replaced with serial means and in the interests of maximum validity for our longitudinal investigations.

3.8.3.3 Managing outliers

Outliers are any numerical values among one or more variables that are found far away from the rest of the data (Barnett & Lewis, 1994). Outliers are problematic in most studies because they can affect the findings of the study and the generalisation of these findings. In this study, any value greater than two standard deviations (SD) from the group mean
was considered to be outlier. A box plot was used graphically to identify outliers. In addition, the plot box was used to assess the spread of data and the distribution of skewness (Tabachnick & Fidell, 2001).

There are two main ways of dealing with outliers, either through the deletion of extreme values or through their replacement with the most extreme score within the acceptance + or - 2SDs range. The latter was considered more suitable for the current study because the sample was relatively small and no loss of additional data would result from it. In total, 20 outliers were identified by performing the box plot: four outliers at Time 1 and sixteen at Time 2 as shown in Table 3.6. So, finally, the data from 56 participants were computed in the main analyses.

Table 3.6

*Outliers for all Study Variables on both Occasions, Time 1 and Time 2, for (N=56)*

<table>
<thead>
<tr>
<th>TIME 1</th>
<th>No of Outliers</th>
<th>TIME 2</th>
<th>No of Outliers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aptitude T1</td>
<td>0</td>
<td>Aptitude T2</td>
<td>0</td>
</tr>
<tr>
<td>Vocabulary and Written coding</td>
<td>0</td>
<td>Vocabulary and Written coding</td>
<td>2</td>
</tr>
<tr>
<td>Grammar Sensitivity</td>
<td>0</td>
<td>Grammar Sensitivity</td>
<td>0</td>
</tr>
<tr>
<td>Memory</td>
<td>0</td>
<td>Memory</td>
<td>1</td>
</tr>
<tr>
<td>English proficiency T1</td>
<td>1</td>
<td>English proficiency T2</td>
<td>5</td>
</tr>
<tr>
<td>Listening</td>
<td>0</td>
<td>Listening</td>
<td>5</td>
</tr>
<tr>
<td>Grammar</td>
<td>3</td>
<td>Grammar</td>
<td>3</td>
</tr>
<tr>
<td>Reading</td>
<td>0</td>
<td>Reading</td>
<td>0</td>
</tr>
</tbody>
</table>
3.9 Data processing

3.9.1 SPSS package

The data, collected and selected under the above preparation criteria, were processed using SPSS (Statistical Package for the Social Sciences, Version 17, 2008).

3.9.2 Mean difference measures

Paired sample t-tests were used to identify statistically significant differences between two means of the same variable with the same participants on two different occasions. In this study, it was conducted to investigate the stability of aptitude and its components over seven months. In addition, it was used to assess the progress of English after seven months’ exposure to English.

3.9.3 Correlation measures

Correlation tests are used to detect and measure the linear relationships between two variables. In statistics, correlation analysis is used to investigate the relationship between two or more normally distributed variables. In the current study, Pearson’s product moment correlation test was conducted to identify the relationships between the FL aptitude scores with the English proficiency in the target population. Pearson’s correlation coefficient ($r$) was used to measure the predictive power of the FL aptitude for English achievement.

Correlation analyses were conducted to examine the relationship between the FL aptitude scores and the English proficiency scores at Time 1 and Time 2.
3.9.3.1 Regression analyses

Multiple regression analyses can be used to assess the individual contribution of a set of variables in predicting the final scores. In the current study, multiple regression analyses were used to quantify the contribution of each FL aptitude’s subtests (four variables) in predicting English proficiency subtests (four variables) along the seven-month English course.

3.9.3.2 Zero-order correlations

Zero-order correlation analyses were used to investigate the correlation between the foreign language aptitude test components and the English proficiency subtests in the target sample at one occasion. This approach provides a snapshot of information about the relationships between these variables in the study at a specific period of time (Time 1 or Time 2)

3.9.4 Longitudinal studies

3.9.4.1 Role of mean differences measures in longitudinal studies

The aim of the mean differences measures was to investigate the differences between two means of the foreign language aptitude test scores for same participants at both times (Time 1 and Time 2). It was also used to measure the differences of the means in English proficiency scores at Time 1 and Time 2 for the same target participants along the seven-month English course.
3.9.4.2 Role of correlation measures in longitudinal studies

The correlation measures in the current study aimed to measure the magnitude of correlation between the FL aptitude components scores and the English proficiency subtests after the participants’ exposure to the seven-month English course. The correlation between the foreign language aptitude test scores at the beginning of the course (Time 1) with the gained progress in English proficiency subtests after finishing the English course were also investigated.

3.9.5 Calculating English proficiency progress

To answer the key research question related to the longitudinal relationships between FL aptitude and the progress in English proficiency among the target population, three formulas were examined for their relative ability to convey useful information concerning how best to measure English proficiency progress.

The comparison of FL aptitude score with the progress score was then undertaken correlatively. (If the aptitude measure is measuring the capacity to attain foreign language skills (proficiency) within a set time, so that prediction of proficiency is reliable, then a higher aptitude score ought to mean greater rate of progress, so relationship of aptitude and progress would be linear.

3.9.5.1 Formula One (Absolute value)

The absolute value of English proficiency progress, which measures the differences between the two values according to the following calculation:
Absolute progress: \[ \Delta P = \text{Eng. Prof. T2} - \text{Eng. Prof. T1} \]

\[ \Delta P = \text{absolute value of English progress} \]

Eng Prof. T1 = English total score at Time 1
Eng Prof. T2 = English total score at Time 2

For example:

A participant underwent the English proficiency tests at Time 1 and Time 2. He/She got 10 scores in Time 1 and 20 scores in Time 2. According to the above formula, this absolute value of progress was calculated as follows:

\[ \Delta P = \text{Eng Prof. T2} - \text{Eng Prof. T1} \]

Absolute progress value = 20 – 10 = 10

3.9.5.2 Formula Two (Ratio: Absolute value compared with initial proficiency measure)

The ratio of English proficiency progress was calculated according to the following formula:

\[ \text{Ratio of progress} = \frac{(\text{Eng Prof. T2} - \text{Eng Prof. T1}) \times 100}{\text{Eng. Prof. T1}} \]

For example:

A participant did English proficiency tests at Time 1 and Time 2. He/She got 10 scores at Time 1 and 20 scores at Time 2. The ratio of English progress was calculated according to the above formula as follows:

\[ \text{Ratio of Progress} = \frac{(20-10)*100}{10} = 100\% \]

This formula conveyed a ratio of progress in English scores between Time 1 and Time 2, as expected non-linear percentages of progress for all English components in the current study.

The implications of using two different formulas are displayed for some hypothetical perspectives.

For example, supposing two participants had the English proficiency tests two times and the first participant got 10 scores in Time 1 and 20 scores in Time 2. The other participant had 40 in the first test and 50 in the second test. This is the way there scores would be calculated by using the ratio formulas.
Absolute value of progress
1st participant: $20 - 10 = 10$
2nd participant: $50 - 40 = 10$

By using the Ratio progress formula:

$$\frac{(\text{Eng Prof. T2- Eng Prof. T1}) \times 100}{\text{Eng. Prof. T1}}$$

1st Participant: $\frac{(20 - 10)}{10} \times 100 = 100\%$
2nd Participant: $\frac{(50 - 40)}{40} \times 100 = 25\%$

According to the example, the two formulas do not yield the same results. The absolute formula did not show any difference in participants’ progress. Both got a differential score of 10. But the ratio formula showed that the first participant showed 4 times more progress compared with the other participant because the ratio formula factors in participants’ initial performance levels weighting in the lower part of the scale more than in the higher part of the scale.

3.9.5.3 Formula Three (Ratio: Absolute value compared with final proficiency measure)

The ratio of English progress was calculated according to the following formula:

$$\frac{(\text{Eng Prof. T2- Eng Prof. T1}) \times 100}{\text{Eng. Prof. T2}}$$

For example:

A participant had English proficiency test at Time 1 and Time 2. She/He got 10 scores at Time 1 and 20 scores at Time 2. The ratio of English progress was calculated according to the above formula:

Ratio of progress
1st Participant: $\frac{(20-10)}{20} \times 100 = 50\%$
2nd Participant: $\frac{(50-40)}{50} \times 100 = 20\%$
The ratio formula showed that the first participant showed 2.5 times more progress compared with the other participant because the ratio formula factors in participants’ final performance levels, weighting in the lower part of the scale more than in the higher part of the scale, but not weighting it in as much as Formula Two.

In the current study all three comparisons will be undertaken to find if there is any significant correlation. Then, if there is any correlation, the nature of the relationship will be explained.

3.10 Discussion of limitations

After reviewing the design of this study and the relevant statistical processes to examine the relationships between the FL aptitude (its total and its components) and English proficiency (its total and its components), a number of limitations need to be noted. First, the project used a convenience sample that involved males only and did not include females. Second, the sample size is not big enough to generalise the findings of this study to the whole population. Undoubtedly, a larger sample would have increased the generalisation of any strong correlation. Third, although the foreign language aptitude test used in this study was piloted and found to have sufficient degrees of reliability and validity, it has not been administered amongst a large population, nor has it been administered in as many different contexts as the MLAT. Fourth, in the current study, TOEFL was used to measure listening, grammar and reading, but it did not include either writing or speaking because of the limitation of time.
3.11 Summary

This chapter reviewed various issues related to the study setting, full description of participants, the construction and piloting of the foreign language aptitude test. The main issue was related to the methodological design, and for this study two main designs were used. Firstly, a cross-sectional design, which is conducted at both Time 1 and Time 2, was to examine the relationships between the FL aptitude measures and the English proficiency measures at one setting. Secondly, a longitudinal design, which aimed to investigate the relationships between variables before and after the English course, was used. The novel contribution of this study was in using the longitudinal design to examine the relationships between the FL aptitude and English progress during the English course. The results of these investigations will be presented in the next chapter and more evaluations and interpretations of these results will be provided in chapter Five.
Chapter 4

FINDINGS

4.0 Overview

The previous chapters shed some light on FL aptitude testing and the process of developing a tool for Arabic native speakers. This tool was designed to be empirically examined for consistency with other aptitude measuring instruments. Various correlation designs such as cross-sectional design and longitudinal design were conducted to investigate the relationships between the foreign language aptitude test and the English proficiency measures among Saudi university students during the seven-month English course.

This chapter reports the findings of the testing, starting with the preliminary descriptive statistics for the selected sample, and then proceeding to report the material relevant to the four research steps: 1) Foreign language aptitude testing instrument—its empirical validation; 2) FL aptitude predicting English proficiency; 3) Development of English proficiency from seven-month intensive English course; and 4) Examining the relationship between FL aptitude measures and English progress.

Reporting the findings uses a mixture of tabular presentations of the computations carried out with the SPSS package, diagrammatic representations of the steps in the analysis and the dimensions delivered from the measuring instruments, and propositional statements expressing any significant relationships between the variables being examined as conclusions supported by the comparisons of the various computations.
4.1 Preliminary statistical analyses: Descriptive statistics

As indicated in Chapter 3, the selection of the sample of participants was designed to manage extraneous variables that might have corrupted the results. An assessment of the participants’ social demographics revealed that the participants were aged between 18 and 21. They were all native speakers of Arabic and they had learnt English formally in high school for six years; it was a subject in the school curriculum. They had a similar cultural and social background.

After the initial screening of the data for missing values, the scores of the foreign language aptitude tests and English language proficiency tests at Time 1 and Time 2 were statistically analysed using SPSS. For all analyses, the significance level of the analyses (or alpha) was set to .05.

To answer the key research questions and to investigate four basic aspects of the efficacy and relevance of the foreign language aptitude testing instrument in predicting English language proficiency, and in other diagnostic uses, the current study focused on two main variables: first, foreign language aptitude test as an independent variable consisting of three subtests (Vocabulary and Written Coding, Grammar Sensitivity and Memory) and acting as a measure of FL aptitude; second, the Test of English as a foreign Language (TOEFL), as a measure of English proficiency (Educational Testing Services, 2002), also consisting of three subtests (Listening, Grammar, and Reading), but not necessarily mapping directly onto the aptitude test subtests.

The descriptive statistics analyses included checks for the normality of the data (skewness and kurtosis), the mean as central tendency index, maximum, minimum, and standard deviations as measures of dispersion. These are all displayed in Table 4.1.
Table 4.1

*Descriptive Statistics of the Foreign Aptitude Scores and English Proficiency Scores at T1 and T2 (N=56)*

<table>
<thead>
<tr>
<th>Variables</th>
<th>TIME 1</th>
<th></th>
<th>TIME 2</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Min</td>
<td>Max</td>
</tr>
<tr>
<td>Vocabulary</td>
<td>12.33</td>
<td>2.74</td>
<td>7</td>
<td>18</td>
</tr>
<tr>
<td>Written Coding</td>
<td>6.91</td>
<td>2.48</td>
<td>3</td>
<td>12</td>
</tr>
<tr>
<td>Grammar Sensitivity</td>
<td>14.82</td>
<td>4.60</td>
<td>5</td>
<td>24</td>
</tr>
<tr>
<td>Memory</td>
<td>34.01</td>
<td>7.10</td>
<td>20</td>
<td>51</td>
</tr>
<tr>
<td>Aptitude Total</td>
<td>18.41</td>
<td>5.43</td>
<td>6</td>
<td>29</td>
</tr>
<tr>
<td>English Proficiency</td>
<td>8.69</td>
<td>3.06</td>
<td>2</td>
<td>16</td>
</tr>
<tr>
<td>Listening</td>
<td>6.73</td>
<td>2.86</td>
<td>.00</td>
<td>12</td>
</tr>
<tr>
<td>Grammar</td>
<td>3.00</td>
<td>1.47</td>
<td>1</td>
<td>7</td>
</tr>
</tbody>
</table>
The graphical representation of the distribution curves for the variables being examined at T1 and T2, as seen in Figures 4.1, 4.2, 4.3, and 4.4, confirmed that the scores on the two key variables, as total scores, appeared to be normally distributed.

The normality of the data distribution was determined by inspecting skewness and kurtosis. As Table 4.1 shows, the skewness values range between -.59 to .73 and kurtosis ranges between .45 to -.78; that is, they are close to zero. This indicates that the foreign language aptitude test and the English proficiency test were neither too difficult nor too easy for the majority of the target sample and were able to detect differences in performances between participants.
Figure 4.2 Distribution of foreign language aptitude test scores at T2 (N=56)

Figure 4.3 Distribution of English proficiency scores at Time 1 (N=56)
As indicated in Table 4.1, the mean of the total score of foreign language aptitude test was 34.01 (SD = 7.10) for Time 1, which was 53.1% of the total score (64). At T2, the mean was 34.53 (SD = 7.01), which was 53.95% of the total score of the foreign language aptitude test (64 was the total possible). This finding shows stability in foreign language aptitude test scores from Time 1 to Time 2. More formal tests of this stability will be reported later in this chapter.

**Figure 4.4** Distribution of English proficiency scores at Time 2 (N=56)

**Figure 4.5** Visual representation of contributions of subtests to FL aptitude as represented by ranges (Minimum and Maximum)
The mean of English proficiency test was 18.41 (SD = 5.43) at Time 1, which was 30.7% of the English proficiency total score and 21.50 (SD = 7) at Time 2, which was 35.8% of the total score (60 was the total score possible at both T1 and T2). This finding shows that there was marginal progress in English proficiency from Time 1 to Time 2. Again, a more detailed examination of this finding will follow.

As can also be seen in Table 4.1, the standard deviations and the mean differences show good enough variations between the measures at T1 and T2, and with a large enough range of differences in the distributions to allow the conduct of correlational analyses.

4.2 FL aptitude testing instrument – Test of psychometric properties

The question “Is the foreign language aptitude testing instrument in the current study a valid test, to a level similar to other equivalent instruments?” and the sub-question “Do the foreign language aptitude test scores at the beginning of the English course (T1) differ

Figure 4.6 Visual representation of contributions of subtests to English proficiency as represented by ranges (Minimum and Maximum)
from the foreign language aptitude test scores at the end of the seven-month English course (T2)?” are being addressed through the following tests.

### 4.2.1 Mean differences of aptitude measures

![Aptitude stability diagram](image)

*Figure 4.7 Mean differences comparison, diagrammatically

Table 4.2

<table>
<thead>
<tr>
<th>Variables</th>
<th>Means</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>T1</td>
</tr>
<tr>
<td>VWC</td>
<td>12.33</td>
</tr>
<tr>
<td>Grammar Sensitivity</td>
<td>6.91</td>
</tr>
<tr>
<td>Memory</td>
<td>14.82</td>
</tr>
<tr>
<td>Aptitude</td>
<td>34.01</td>
</tr>
</tbody>
</table>

*Note. VWC= Vocabulary and Written coding; t=Paired sample t-test*

A series of t-test pairs, as reported in Table 4.2, demonstrated that there was no statistically significant difference between the mean of the FL aptitude total score at Time 1 and Time 2 (*p > .05*), which indicates, as expected, that aptitude was stable over time. But there was a significant correlation between VWC T1 and VWC T2. This significant correlation shows that there is a change in Vocabulary and Written Coding ability before and after the English course. Generally, this finding shows that the foreign language
aptitude test was not affected by the seven-month English course. The issue of the stability of FL aptitude will be discussed in more detail in Chapter Five.

4.2.2 Correlation between aptitude measures

The question in view here is “Do the foreign language aptitude test scores at the beginning of the English course (Time 1) co-vary with the foreign language aptitude test scores at the end of the seven-month English course (Time 2), for each of the participants in the sample?”

Correlations of Aptitude

Figure 4.8 Correlation of aptitude over time, diagrammatically

Table 4.3

Correlations of aptitude measures between T1 and T2 measurements (N = 56)

<table>
<thead>
<tr>
<th>Aptitude variables</th>
<th>R</th>
<th>Sig. p</th>
</tr>
</thead>
<tbody>
<tr>
<td>VWC T1 and VWCT2</td>
<td>.50</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Grammar Sensitivity T1 and Grammar Sensitivity T2</td>
<td>.51</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Memory T1 and Memory T2</td>
<td>.36</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Aptitude T1 and Aptitude T2</td>
<td>.50</td>
<td>&lt; .001</td>
</tr>
</tbody>
</table>

The foreign language aptitude scores at the beginning and at the end of the course were moderately correlated ($r = .50$, $p < .05$, $N = 56$), adding extra support to the finding of stability as per the means difference test.
As can be seen in Table 4.3, the distribution of the scores of the participants on all aptitude components, Vocabulary and Written Coding (VWC), Grammar Sensitivity (GS) and Memory (MEM), are significantly and moderately correlated with their scores at Time 1 and Time 2. These coefficients range between .36 and .51. These findings indicate the general internal stability of the aptitude measures. The least stable component was memory, but this component was not unstable enough to compromise the stability of the total score.

Together, the findings of mean difference and correlation of distribution indicate that the foreign language aptitude testing instrument is valid and doing what it was intended to do, to a level similar to other equivalent instruments. It also demonstrates that the aptitude measure is stable and does not change with teaching inputs, and also that stability operates for the aptitude test components and for the individuals within the whole distribution.

4.3 FL aptitude predicting English proficiency

In this section of the results chapter, two research questions are addressed. This section examines the cross-sectional relationships between the FL aptitude and English proficiency at T1 and T2. These relationships will be presented separately in two main subsections.
These cross-sectional correlations are represented diagrammatically as follows:

Cross-sectional Correlations

![Diagram of cross-sectional correlations between foreign language aptitude and English proficiency at Time 1 and Time 2.]

Figure 4.9 Cross-sectional correlations between foreign language aptitude and English proficiency (T1 and T2)

4.3.1 Cross-sectional examination of FL aptitude and English proficiency at Time 1 and Time 2

This section aims to address the research questions related to the cross-sectional relationships between foreign language aptitude test and the English proficiency at T1 and T2:

1. Is there any significant cross-sectional relationship between the FL aptitude scores and the English proficiency at the beginning of the English course (Time 1)?
2. Is there any significant cross-sectional relationship between the FL aptitude scores and the English proficiency at the end of the course (Time 2)?
4.3.1.1 Zero-order correlations at Time 1

Zero-order correlations were computed between each aptitude component, as well as aptitude total score, and the English proficiency total score. The results of these tests are shown in Table 4.4.

Table 4.4

Zero-order Correlations between FL Aptitude Components with English Proficiency Subtests and Total Score across the whole Sample at Time 1 (N = 56)

<table>
<thead>
<tr>
<th></th>
<th>Eng. Prof. Total(r)</th>
<th>Listening(r)</th>
<th>Grammar(r)</th>
<th>Reading(r)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aptitude Total</td>
<td>.29*</td>
<td>.30</td>
<td>.24</td>
<td>.02</td>
</tr>
<tr>
<td>VWC</td>
<td>.29*</td>
<td>.31*</td>
<td>.26*</td>
<td>-.13</td>
</tr>
<tr>
<td>Grammar Sensitivity</td>
<td>.21</td>
<td>.21</td>
<td>.17</td>
<td>-.06</td>
</tr>
<tr>
<td>Memory</td>
<td>.16</td>
<td>.16</td>
<td>.11</td>
<td>.18</td>
</tr>
</tbody>
</table>

Notes. Eng. Prof. Total = English proficiency total score at Time 1; VWC = Vocabulary and Written Coding
* p<.05

There was a moderate significant positive correlation between the FL aptitude total score and the English proficiency total score at Time 1 (r = .29, p<.05, N = 56), indicating that the higher the FL aptitude was at T1, the higher the English proficiency was at T1 and suggesting that differences in the aptitude measure amongst a cohort of students with similar background of exposure to FL can be used to predict differences in proficiency after learning/teaching inputs.

The FL aptitude total score showed a significant correlation only with the total score of the English proficiency, but not with any of the three subtests of English proficiency at Time 1.
While there were moderate correlations of the Vocabulary and Written Coding component with the Listening and Grammar subtests, the relationship between the VWC did not correlate with Reading.

In the case of the English proficiency total score and its relation to particular aptitude components, only Vocabulary and Written Coding displayed a significant correlation with the English proficiency total score ($r = .29, p < .05, N = 56$). In the case of the other two FL aptitude components (Grammar Sensitivity and Memory), there were no significant correlations with the total score of English proficiency.

Table 4.4 also provides the inter-correlations among the foreign language aptitude test components and the English proficiency subtests at Time 1. Vocabulary and Written Coding correlated significantly with two components of English proficiency tests: Listening ($r = .31, p < .05, N = 56$) and Grammar ($r = .26, p < .05, N = 56$). Grammar Sensitivity and Memory showed no significant correlation with the English proficiency total score and its components. Based on this result, we should expect those participants who achieve higher scores in Vocabulary and Written Coding to have higher grammar and listening competencies at the early stage of the L2 course.

4.3.1.2 Aptitude Subtests at T1 as simultaneous predictors of proficiency at T1

A first multiple regression analysis was carried out on Time 1 data, using the scores of the English proficiency subtests at Time 1 as criterion variables and the FL aptitude components at Time 1 as predictors. This analysis produced the following findings:

The regression analysis revealed that:
There was a significant relationship between the total score of the FL aptitude subtests and the total score of the English proficiency tests, \( R^2 = .08, \beta = .29, p = .02 \).

The foreign language aptitude test total score at Time 1 was a significant moderate predictor of the English proficiency total score at Time 1 and accounted for 8.7% of the variability in the English proficiency measured at T1.

Table 4.5

*Multiple Regression between Aptitude Components at T1 and English Proficiency Total Scores at T1 (N= 56)*

<table>
<thead>
<tr>
<th>Aptitude variables Time 1</th>
<th>( R^2 )</th>
<th>( \beta )</th>
<th>( p )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vocabulary &amp; Written Coding</td>
<td>.08</td>
<td>.29</td>
<td>.02</td>
</tr>
<tr>
<td>Grammar Sensitivity</td>
<td>.04</td>
<td>.21</td>
<td>.15</td>
</tr>
<tr>
<td>Memory</td>
<td>.02</td>
<td>.16</td>
<td>.23</td>
</tr>
</tbody>
</table>

As presented in Table 4.5, the results of multi-regression analysis, in which all subtests were entered simultaneously, showed that:

- Vocabulary & Written Coding at T1 was a significant predictor of the English proficiency total score at T1 \( (R^2 = .08, \beta = .29, p = .02) \), while controlling for performance on all the other subtests at T1.

- Grammar Sensitivity, as a second component of the foreign language aptitude test, was not a significant predictor of the English proficiency total score \( (R^2 = .04, \beta = .21, p = .15) \) when controlling for performance on all the other subtests at T1.
Also, Memory was not a significant predictor of English proficiency total score T1 ($R^2 = .02$, $\beta = .16$, $p = .23$) when controlling for performance on all the other subtests at T1.

Together, the results of zero-order correlation and multiple regression analyses provide an answer to the research question concerning the relationship between the FL aptitude and English proficiency at T1. These results indicate that there was a moderate significant correlation between the foreign language aptitude test and the English proficiency test at an early stage of L2 instruction. The regression analyses showed that the foreign language aptitude test was a moderate predictor of English proficiency at the beginning of the course.

4.3.1.3 Cross-sectional examination of FL aptitude at Time 2 and English proficiency at Time 2

This second set of cross-sectional analyses of the relationship between aptitude and English proficiency aimed to answer the following question:

Is there any significant cross-sectional relationship between FL aptitude and English proficiency at the end of the same course (Time 2)?

4.3.1.4 Zero-order correlations at Time 2

The same analytical approach and statistical tools used to analyse Time 1 relationships were used for Time 2 variables. These results are presented in Table 4.6.
Table 4.6

Zero–order Correlations between Foreign Aptitude Components with English Proficiency Subtests for the Whole Sample (N=56) in T2

<table>
<thead>
<tr>
<th></th>
<th>Eng. Prof. Total (r)</th>
<th>Listening (r)</th>
<th>Grammar (r)</th>
<th>Reading (r)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aptitude Total</td>
<td>.34**</td>
<td>.36**</td>
<td>.26*</td>
<td>.05</td>
</tr>
<tr>
<td>Vocabulary and Written Coding</td>
<td>.21</td>
<td>.25</td>
<td>.24</td>
<td>-.04</td>
</tr>
<tr>
<td>Grammar Sensitivity</td>
<td>.26*</td>
<td>.24</td>
<td>.20</td>
<td>.10</td>
</tr>
<tr>
<td>Memory</td>
<td>.23</td>
<td>.24</td>
<td>.13</td>
<td>.09</td>
</tr>
</tbody>
</table>

*Note. Eng. Prof Total = English proficiency total score at Time 2

* p = <.05, ** p = <.01

There was also a moderate, positive, significant correlation between FL aptitude total score at Time 2 and English proficiency total score at Time 2 (r = 0.34, p =.005, N=56), indicating that the higher the aptitude scores the higher the English proficiency.

At the end of the course (T2), more significant correlations between the FL aptitude total scores and the English proficiency subtests appeared. Compared to the beginning of the course, the foreign language aptitude tests' total score was significantly correlated with two components of the English proficiency test:

Listening (r = .36, p < .01, N = 56) and Grammar (r = .26 p <.01, N = 56), but not with Reading.

Regarding the relationships between the English proficiency total scores and FL aptitude components, only Grammar Sensitivity showed a significant correlation with the English proficiency total score at Time 2 (r = .26, p < .05, N = 56). Vocabulary and Written Coding and Memory subtests did not correlate significantly with the English proficiency total score at the end of the course (Time 2).
This result indicates that those students who have higher grammar sensitivity after finishing the seven-month English instruction are expected to have higher English proficiency at the end of the course. Table 4.6 also presented the inter-correlations between the FL aptitude components and the English proficiency subtests. None of these inter-correlations were statistically significant, despite a significant correlation between their total scores.

4.3.1.5 *Multiple regression test of FL aptitude T2 as predictor of English proficiency T2*

The second regression analyses were carried out at the end of the seven-month English course (T2) by using the scores of the English proficiency subtests at T2 as the criterion variable and the scores of the FL aptitude components at Time 2 as predictors of English success.

The regression analysis showed a significant relationship between the foreign language aptitude test total score at Time 2 and the English proficiency total score at T2, ($R^2 = .12$, $\beta = .34, p = .001$). Comparing this result with the FL aptitude and English proficiency at Time 1, there is a marginal difference between the regression values ($R^2 = .08$ at T1 and $R^2 = .12$ at T2).
Table 4.7

*Multiple regression between Aptitude components at Time 2 and English proficiency total scores at Time 2 (N= 56)*

<table>
<thead>
<tr>
<th>Aptitude variables Time 2</th>
<th>$R^2$</th>
<th>$\beta$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>VWC 2</td>
<td>.04</td>
<td>.21</td>
<td>.10</td>
</tr>
<tr>
<td>Grammar Sensitivity2</td>
<td>.07</td>
<td>.26</td>
<td>.47</td>
</tr>
<tr>
<td>Memory2</td>
<td>.05</td>
<td>.23</td>
<td>.07</td>
</tr>
</tbody>
</table>

*Notes.* Eng. Prof. Total = English proficiency total score at Time 2; VWC = Vocabulary & Written Coding

As shown in Table 4.7, the multi-regression analyses indicate that there the FL aptitude components at T2 were not good predictors of the total score of the English proficiency test at (T2).

Together, the zero-order correlation and multiple regression analyses, for T2, indicate there was a moderate significant correlation between the foreign language aptitude test and English proficiency at the end of the English course. The regression analyses showed also that the foreign language aptitude test was a significant predictor of success at the end of the course.

4.3.2 Summary of cross-sectional relationships between FL aptitude and English proficiency at Time 1 and Time 2

When comparing T1 and T2 correlations, there was quite a small difference between the size of the correlation coefficients of the total scores of FL aptitude and English proficiency at the beginning of the English course ($r = .29$) and after seven months of English learning ($r = .34$).
Despite this overall similarity of correlations for the total scores of the FL aptitude and English proficiency at T1 and T2, different patterns of correlations appeared among the FL aptitude and the English proficiency subtests.

As shown by comparing Table 4.4 with Table 4.6, the English proficiency total score at T2 related to the FL aptitude Grammar Sensitivity score at T2, whereas the English proficiency total score at T1 related to the FL aptitude Vocabulary and Written Coding score at T1. At T2, the Grammar Sensitivity appeared to be affected by the seven-month English course which focused on teaching grammar. Possible explanations of these cross-time variations will be provided in Chapter Five.

Although the correlation analyses showed significant positive correlations between students’ total FL aptitude scores and their English proficiency scores at both (T1 and T2), the multiple regression analysis revealed that none of the three subtests, separately, stood out as a significant predictor of English proficiency, while controlling for the other FL aptitude subtests.

In summary, the cross-sectional findings in the current study support previous correlational studies which generally found that the strength of correlation ($r$) between aptitude tests and foreign-language success ranges between .20 and .60, and at the lower end (.20 and .40) for adapted versions. More details related to the magnitude of correlations between FL aptitude and language proficiency in previous studies are provided in Appendix 1.

In other words, the cross-sectional findings, together with the findings from Section 4.2, which indicate that aptitude is a stable measure, demonstrate that the foreign language aptitude testing instrument is valid and does what it is intended to do, to a level similar to
other equivalent instruments, and to the extent of predicting English language learning proficiency to a standard similar to other aptitude tests.

4.3.3 Longitudinal examination of FL aptitude and English proficiency

This section investigates the relationship between the FL aptitude and the English proficiency over the seven-month period captured by this study. The following research question was addressed:

Do the foreign language aptitude test scores at the beginning of the seven-month English course (T1) correlate with the English proficiency test scores at the end of the seven-month English course (T2)?

This T1-T2 correlation is represented diagrammatically as follows:

Longitudinal correlation

Seven-month English course

Figure 4.10 Longitudinal correlations of foreign language aptitude measures and English proficiency measures

In this section, the relationship between the foreign language aptitude test scores at the beginning of the seven-month English course (T1) and English proficiency test scores at the end of the seven-month English course (T2) was investigated. The current analyses control for the dependant variables (Aptitude) at Time 1. The results of these correlation analyses are presented in Table 4.8.
4.3.3.1 Test1: Zero-order correlation

Table 4.8

Zero-order Correlations of FL Aptitude Components at T1 with English Proficiency Subtests at T2 for the Whole Sample (N=56)

<table>
<thead>
<tr>
<th></th>
<th>Eng. prof. Total T2 (r)</th>
<th>Listening T2 (r)</th>
<th>Grammar T2 (r)</th>
<th>Reading T2 (r)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aptitude Total T1</td>
<td>.27*</td>
<td>.26*</td>
<td>.23*</td>
<td>.04</td>
</tr>
<tr>
<td>VWC T1</td>
<td>.17</td>
<td>.27*</td>
<td>.16</td>
<td>-.10</td>
</tr>
<tr>
<td>Grammar Sensitivity T1</td>
<td>.14</td>
<td>.11</td>
<td>.16</td>
<td>.13</td>
</tr>
<tr>
<td>Memory T1</td>
<td>.24*</td>
<td>.18</td>
<td>.17</td>
<td>.13</td>
</tr>
</tbody>
</table>

Note. Eng. prof. Total T2 = English proficiency total score at Time 2; VWC = Vocabulary & Written Coding *p = <.05

The results of correlation analyses indicated that FL aptitude scores were significantly although moderately correlated (r = .27, p<.05, N = 56) with the total score of English proficiency at the end of the English course.

Some significant inter-correlations, ranging between .23 and .27, appeared among FL aptitude components at T1 and the subtests of English proficiency test at T2. In particular, the foreign language aptitude test Total score at T1 was significantly correlated with two English proficiency subtests at T2 (Listening r = .26, p<.05 and Grammar r = .23, p<.05) but not with Reading. Regarding aptitude components, Vocabulary and Written Coding at T1 correlated only with Listening at T2 (r = .27, p<.05). Memory at T1 correlated only with the total score of English proficiency at T2 and not with any of its subtests. Contrary to expectations, this study did not find a significant correlation between Grammar
Sensitivity at T1 and English proficiency at T2. More details related to this finding are provided in Chapter Five.

4.3.3.2 Test 2: Multiple regressions of FL aptitude subtests at T1 as simultaneous predictors of English proficiency T2

At the end of the seven-month English course (Time 2), the second regression analyses was applied, the scores of English proficiency subtests at Time 2 were used as the criterion variable and the scores of the FL aptitude components at Time 1 as predictors of English achievement.

The regression analyses revealed that there was a significant relationship between the foreign language aptitude test total score (T1) and the English proficiency total score at T2 ($R^2=.07; \beta=.27, p=.04$). It accounted for 7% of the variance of English proficiency total scores at T2.

Table 4.9

Multiple Regression between Foreign Aptitude Components at Time 1 and English Proficiency Total Scores at Time 2 ($N=56$)

<table>
<thead>
<tr>
<th>Aptitude variables Time1</th>
<th>$R^2$</th>
<th>$\beta$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vocabulary-Written Coding T1</td>
<td>.03</td>
<td>.17</td>
<td>.21</td>
</tr>
<tr>
<td>Grammar Sensitivity T1</td>
<td>.02</td>
<td>.14</td>
<td>.29</td>
</tr>
<tr>
<td>Memory T1</td>
<td>.05</td>
<td>.24</td>
<td>.07</td>
</tr>
</tbody>
</table>

The results of multi-regression analyses of each subtest of foreign language aptitude test at T1 with the total score of English proficiency at (T2) indicate that none of the FL aptitude components was found to be a significant unique predictor of English success at the end of the English course.
4.3.4 Summary of longitudinal examination of FL aptitude and English proficiency

The longitudinal examination of the foreign language aptitude test at the beginning of the course (T1) and the English proficiency scores at the end of the seven-month course showed that these two variables were positively and significantly correlated but the magnitude of correlation was only moderate. The regression analyses revealed that the total score of the FL aptitude can account for 7% of the variance in the English proficiency total score at T2.

Although the correlation analysis showed significant positive correlations between students’ total FL aptitude scores at T1 and their English proficiency scores at T2, the multiple regression analyses revealed that none of the three subtests, separately, stood out as a significant predictor of English proficiency.

The above correlation and regression analyses provided an answer for our research question related to the longitudinal relationship between the foreign language aptitude test and the English proficiency, strengthening the claim for aptitude measures predicting proficiency. The measuring of Aptitude at Time 1 and proficiency at Time 2 represents the typical use and usefulness of undertaking aptitude tests before selecting and screening students for a particular course of studies, or using aptitude measures for streaming for different teaching approaches.

In the next section, our attention turns to a further elaboration on the relationship between the FL aptitude and the English proficiency. These relationships were investigated to examine the relationships between the FL aptitude scores and the progress, rather than the more absolute level, in English proficiency after seven-month course.
4.4 Development of English proficiency from seven-month intensive English course

This section addresses the research question related to the progress of English proficiency after exposure to the intensive English course for seven months: “Do English proficiency test scores at the beginning of the seven-month English course (T1) differ from English proficiency test scores at the end of the course (T2)?”

Mean differences of English proficiency measures

![Diagram showing English Progress from Time 1 to Time 2](image)

Figure 4.11 Mean differences for English proficiency measures

4.4.1 Mean differences of English proficiency measures

The mean differences tests of English proficiency measures before and after the English course revealed that there was a general increase in the level of English proficiency among the participants over time.
Table 4.10

Mean Differences for English Proficiency Measures before and after the Seven-Month Course (N=56)

<table>
<thead>
<tr>
<th>Variables</th>
<th>Means</th>
<th></th>
<th></th>
<th>SD</th>
<th></th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>T1</td>
<td>T2</td>
<td>T2-T1</td>
<td>T1</td>
<td>T2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Listening</td>
<td>8.69</td>
<td>9.21</td>
<td>.52</td>
<td>2.72</td>
<td>2.71</td>
<td>1.42</td>
<td>.16</td>
</tr>
<tr>
<td>Grammar</td>
<td>6.37</td>
<td>8.33</td>
<td>1.6</td>
<td>2.98</td>
<td>2.68</td>
<td>4.03</td>
<td>.001</td>
</tr>
<tr>
<td>Reading</td>
<td>3.00</td>
<td>3.91</td>
<td>.91</td>
<td>2.13</td>
<td>1.81</td>
<td>3.19</td>
<td>.001</td>
</tr>
<tr>
<td>English Proficiency</td>
<td>18.41</td>
<td>21.50</td>
<td>3.09</td>
<td>3.08</td>
<td>5.72</td>
<td>5.22</td>
<td>.001</td>
</tr>
</tbody>
</table>

Note. SD= Standard Deviation; t=Paired sample t-test, MD= Mean Difference

As shown in Table 4.10 the increase in the mean for the total score was 3.09. This result demonstrates that there is a positive and significant ($p < .05$) difference between the mean of English proficiency total score at T1 and T2, thus indicating the expected progress in English proficiency among participants after learning English for seven months.

Reading and Grammar, as subtests of English proficiency, showed significant improvements in mean levels after the seven-month course, but Listening did not. Taking these changes at face value, this suggests that the reading and grammar competencies were improved by the seven-month English course, and, in comparison, that the same course failed to develop the participants’ listening skills.

Grammar showed a comparatively a higher degree of change ($MD = 1.6$) compared with Reading ($MD = .91$). Similarly, one way of describing this is to say that such figures suggest that the seven-month intensive English course was more successful in developing grammar than reading, and unsuccessful in development listening for comprehension.
4.4.2 Correlation analyses for English proficiency

Correlation of English proficiency measures

![Figure 4.12 Correlation of distribution of English proficiency measures](image)

Table 4.11

**Correlations of English Proficiency Measures between T1 and T2 Measurements (N = 56)**

<table>
<thead>
<tr>
<th>English Proficiency variables</th>
<th>r</th>
<th>Sig. (p)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Listening T1 and Listening T2</td>
<td>.56</td>
<td>.001</td>
</tr>
<tr>
<td>Grammar T1 and Grammar T2</td>
<td>.42</td>
<td>.001</td>
</tr>
<tr>
<td>Reading T1 and Reading T2</td>
<td>.17</td>
<td>.21</td>
</tr>
<tr>
<td>English Proficiency T1 and T2</td>
<td>.68</td>
<td>.001</td>
</tr>
</tbody>
</table>

Table 4.11 shows that there was a moderately strong correlation between English proficiency total score across T1 and T2 ($r = .68, p < .05, N = 56$). All English proficiency measures at T1 and T2, except Reading, were significantly and moderately correlated, with coefficients varying between .42 and .56. This indicates a substantial degree of co-variation between time measurements of the same construct.

It can be noted from the mean differences and correlation analyses that there is a significant progress in English proficiency among the target sample after seven-month of English course.
4.5 Examining the relationship between FL aptitude measures and English progress

A positive, significant correlation between the foreign language aptitude test scores at Time 1 and English progress was expected. This expectation was built on past literature. To answer the key research question, “Is there a relationship between the foreign language aptitude test and English progress?” related to the relationship between FL aptitude and the progress in English proficiency in the target population, three slightly different formulas were used to measure English progress. These three formulas reflect different assumptions about whether certain portions of the English proficiency distribution might be more prone to change over time. In the following subsections, new correlations between the FL aptitude scores at Time 1 and the English progress derived from the three different progress measure formulas were reported.

4.5.1 Comparison of aptitude score with Absolute value of English proficiency progress

This measure of English proficiency progress was calculated from the absolute value of English proficiency progress, which measures the differences between the two values according to the following formula. This formula assumes that changes in English proficiency are equally possible in any point along the distribution.

Absolute progress: $\Delta P = (\text{Eng. Prof. T2} - \text{Eng. Prof. T1})$

$\Delta P =$ absolute value of English progress

Eng Prof. T1 = English proficiency total score at Time 1

Eng Prof. T2 = English proficiency total score at Time 2
As shown in Table 4.12, when the correlation between the FL aptitude and the English proficiency progress was investigated using the absolute value formula, there was no significant correlation between the total or the FL aptitude components and the English proficiency subtests. This indicates that participants’ scores on the foreign language aptitude test at T1 were not able to predict participants’ progress, over time, in English proficiency.

### 4.5.2 Relationship of FL aptitude score with the ratio of English progress and initial English proficiency score

The measure of English proficiency progress was also calculated as the ratio of progress in relation to the initial English proficiency measure, according to the following formula. This formula, by giving a greater weight to progress earlier in the distribution, assumes that progress in easier when a student is already quite proficient.

\[
\text{Ratio of progress} = \frac{(\text{Eng Prof. T2} - \text{Eng Prof. T1}) \times 100}{\text{Eng. Prof. T1}}
\]
Table 4.13

Correlations between Foreign Language Aptitude Variables at Time 1 and the Ratio of English Progress along the Seven-Month English Course (N=56)

<table>
<thead>
<tr>
<th>Aptitude Variables</th>
<th>English Total Progress</th>
<th>Listening Progress</th>
<th>Grammar Progress</th>
<th>Reading Progress</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aptitude Total</td>
<td>-.12</td>
<td>-.10</td>
<td>-.08</td>
<td>-.02</td>
</tr>
<tr>
<td>VWC</td>
<td>-.19</td>
<td>-.12</td>
<td>-.11</td>
<td>.03</td>
</tr>
<tr>
<td>Grammar Sensitivity</td>
<td>-.11</td>
<td>-.13</td>
<td>-.13</td>
<td>.01</td>
</tr>
<tr>
<td>Memory</td>
<td>-.01</td>
<td>.00</td>
<td>.00</td>
<td>-.07</td>
</tr>
</tbody>
</table>

*Note. *p*= <.05

Table 4.13 shows that even when using the ratio index, there was no significant correlation between the FL aptitude score at the beginning of the course and the English proficiency progress \( r = -.12, p > .05, N = 56 \) after the seven-month course.

Moreover, there were no significant correlations between the FL aptitude components and the ratio of English progress components.

4.5.3 Relationship of FL aptitude score with the ratio of Absolute English progress and final English proficiency score

The measure of English proficiency progress was calculated as the ratio of progress in relation to the final English proficiency measure, according to the following formula:

\[
\text{Ratio of progress} = \frac{(\text{Eng Prof. T2} - \text{Eng Prof. T1}) \times 100}{\text{Eng. Prof. T2}}
\]
Table 4.14

*Correlations between Foreign Language Aptitude Variables and the Ratio of English Progress along the Seven-Month English Course as per the Third Formula \( (N=56) \)*

<table>
<thead>
<tr>
<th>Aptitude Variables</th>
<th>English Total Progress</th>
<th>Listening Progress</th>
<th>Grammar Progress</th>
<th>Reading Progress</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aptitude Total</td>
<td>-.11</td>
<td>-.19</td>
<td>-.11</td>
<td>.03</td>
</tr>
<tr>
<td>VWC</td>
<td>-.18</td>
<td>-.07</td>
<td>-.17</td>
<td>.01</td>
</tr>
<tr>
<td>Grammar Sensitivity</td>
<td>-.14</td>
<td>-.11</td>
<td>-.06</td>
<td>.08</td>
</tr>
<tr>
<td>Memory</td>
<td>.01</td>
<td>-.07</td>
<td>-.02</td>
<td>-.01</td>
</tr>
</tbody>
</table>

*Note. *\( p = .05 \)*

The results as reported in Table 4.14 shows that there were no significant correlations between the FL aptitude and the ratio of progress in relation to the final English proficiency score. The pattern was the same when focusing on individual components and overall indices.

4.5.4 Summary of longitudinal relationships between the FL aptitude and the English proficiency progress

Despite the significant correlations between the FL aptitude and the English proficiency being found in the cross-sectional and the longitudinal comparisons in the current study, the foreign language aptitude test did not show any significant correlation with English proficiency progress. Considering that the examination of the L2 progress in the current study was to add a new contribution to the field of FL aptitude research, these cross-sectional and longitudinal approaches could be criticised for building their findings on the relationships between the total scores of the FL aptitude tests and L2-proficiency tests without considering the level of progress in L2.
4.6 The foreign language aptitude and TOEFL tests as predictors of achievement

Based on the correlation analyses, the English proficiency score at the beginning of the seven-month course (Time 1) was strongly positively correlated with the English proficiency score at Time 2 ($r = .68, p< 0.001, N = 56$) (see Table 4.11).

To find the best predictor of the English proficiency at Time 2, all Time 1 variables were entered. Turning our attention to looking for the best predictor of English proficiency at the end of the course, multiple regression analyses were used.

Table 4.15

<table>
<thead>
<tr>
<th>All Variables Time1</th>
<th>Eng. Prof. total Time 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$R^2$</td>
</tr>
<tr>
<td>Aptitude Total score</td>
<td>.07</td>
</tr>
<tr>
<td>English Proficiency</td>
<td>.47</td>
</tr>
</tbody>
</table>

An advantage of using multiple regressions, over simple regression, is that it accounts for overlaps between predictors. The findings revealed that English proficiency at T1 accounted for 47% of the variability in English proficiency at T2 ($R^2 = .47$). These findings suggested that English proficiency at the beginning of the course was the best predictor of English success after studying English for seven-months. The foreign language aptitude was the second predictor, it accounted for 7% of the English proficiency at T2 (see Table 4.15).
Table 4.16

*Multiple Regression between all Variables at Time 1 and English Proficiency at Time 2 (N= 56)*

<table>
<thead>
<tr>
<th>All variables Time 1</th>
<th>$R^2$</th>
<th>$\beta$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vocabulary &amp; Written Coding 1</td>
<td>.02</td>
<td>.17</td>
<td>.20</td>
</tr>
<tr>
<td>Grammar Sensitivity 1</td>
<td>.02</td>
<td>.14</td>
<td>.29</td>
</tr>
<tr>
<td>Memory 1</td>
<td>.05</td>
<td>.24</td>
<td>.07</td>
</tr>
<tr>
<td>Listening 1</td>
<td>.43</td>
<td>.66</td>
<td>.001</td>
</tr>
<tr>
<td>Grammar 1</td>
<td>.23</td>
<td>.48</td>
<td>.001</td>
</tr>
<tr>
<td>Reading 1</td>
<td>.10</td>
<td>.31</td>
<td>.001</td>
</tr>
</tbody>
</table>

In relation to the English proficiency subtests at Time 1 and their unique impact on the English proficiency total score at Time 2, the results of multi-regression analyses show that all English proficiency subtests were significant predictors of English proficiency at T2. Both Listening and Grammar at T1 were significant predictors of English proficiency total score at Time 2 ($R^2 = .43$, $p<.01$; $R^2 = .23$, $p<.01$ respectively). Also, Reading at Time 1 was a significant predictor of English proficiency total score at Time 2 ($R^2 = .10$, $p<.01$).

Table 4.15 indicates that prior learning — the English proficiency— was the best predictor of English success in the current study (47%) and the foreign language aptitude was the second (7%) predictor. In total, they account for 54% of the English proficiency at T2 in this study. So, other factors like motivation, age, etc., that have not been considered in this study, account for the remainder. A result covering more than 50% is an important feature of these findings.
4.7 Summary of study findings

In this chapter, the results regarding the four research questions in the current study have been presented. These questions dealt with the nature of the foreign language aptitude test, the foreign language aptitude predicting English proficiency, development of English proficiency from seven-month intensive English course, including any relationship between aptitude and progress. In addition, the data, when processed, allowed for a consideration of the best predictor of English success in the current study.

The results of the study of can be summarised in the following points:

- Performance on the foreign language aptitude test is stable over time supporting the concept of aptitude stability. There was no substantial difference between the two group means of the foreign language aptitude test at both times and the two measurements were highly correlated. Hence, our results show that the foreign language aptitude test was not affected by the intensive seven-month English course and the distribution of the foreign language aptitude scores at the beginning of the course was similar to the distribution of scores of the foreign language aptitude at T2.

- Second, the results of cross-sectional correlations revealed that the foreign language aptitude test scores correlate significantly and positively with the English proficiency test scores at both times (T1 and T2), but the magnitude of correlation was generally weak. The longitudinal correlation between the foreign language aptitude test at the beginning of the course (T1) and the English proficiency test scores at the end of the English course (T2) was significant and positive but it was also weak.
Third, the results regarding progress in English proficiency indicate that there was a measurable and statistically significant progress in English competencies after exposure to the intensive English course for seven months, but the actual size was disappointingly small (of the order of 3 units in 60, i.e. 5%). This result was obtained from the difference between the two means of English proficiency at T1 and T2. This progress reveals that participants’ English skills developed as a result of the English course.

Fourth, another issue is the longitudinal relationship between the foreign language aptitude test scores at the beginning of the English course (T1) and English progress. These results show that there was no significant correlation between the foreign language aptitude measures and the English proficiency progress measures.

Fifth, the computations involved in the study of the four research questions provided information that allowed for the consideration of the best predictor of English proficiency from among all variables in this study (Aptitude = four variables, English proficiency = four variables). These analyses revealed that the best predictor of English proficiency at T2 was the English proficiency at T1, followed by the foreign language aptitude.

These five main issues will be the core of the next chapter, where further discussion and explanations of these results will be provided.
Chapter 5

DISCUSSION

5.0 Overview

This chapter provides an in-depth discussion of the results presented in Chapter 4, focusing on a range of issues as follows: 1) the nature of aptitude measurement in an Arabic-speaking context; 2) the capacity of the foreign language aptitude test to predict English proficiency, by comparison with other similar instruments; 3) the extent of the development of English proficiency over a seven-month intensive English course; 4) the capacity of the foreign language aptitude test to predict English progress. In relation to the latter, the discussion also seeks to establish whether the foreign language aptitude test or the language test (used for the purposes of the research) is a better predictor of language achievement among this population of learners. Some consideration is also given to the data derived through correlational analyses of different components of the aptitude and the language tests.

As noted previously, the design and focus of this study have been primarily empirical. It has been conducted to achieve particular practical ends: having a tool that might help students, teachers and policy-makers make more informed decisions about English language learning/teaching issues, in the light of individual differences in language-learning potential. As with other correlational studies conducted so far, it has less to contribute to current theoretical considerations (Skehan, 2002; Robinson, 2002).

The issues arising from the data derived in the current study have both theoretical and practical dimensions. With regard to the latter the researcher endeavours to find
explanations of the study’s findings based on his already considerable professional experience in working in the Saudi context with Saudi EFL learners. For instance, in relation to the relative lack of achievement over the duration of the course (seven-months), a matter of practical concern to students, teachers and university policy makers alike, one probable explanation includes a range of factors that have little or nothing to do with aptitude, the main focus of this study. Other studies have shown that factors like low learner motivation, inadequate FL input, poor teaching practices, etc., can have an overriding impact on learning outcomes. The low achievement values found in this study, and possibly arising from the operation of such factors, might be responsible for the inconclusive nature of some of the findings of this study. The focus, in the study design, on correlational analyses of the data and the very small difference between the proficiency measures (a surprise) means any correlative comparison is unlikely to deliver meaningful results.

In considering these results for their theoretical implications, this study will focus on various keys concepts: Robinson’s Aptitude Complex Hypothesis (2002) and how it may inform the research agenda for aptitude studies; the idea of ‘noticing’ as a significant element of any learning (as applied to SLA by Schmidt, 1990); Skehan’s enunciation of the relationship between SLA and FL aptitude and his extrapolations about widening the research agenda for FL aptitude studies (2002). For more details of the different theoretical positions and their evaluation, for relevance to this study, see Chapter 2, Section 8.

The understanding of SLA taken into this discussion is firstly that it operates like other learning: noticing is a key first step; focusing on the noticed and applying either memory or analytical problem-solving is the next step. Problem-solving, by developing working
hypotheses and testing those in application, to detect error and to refine the working hypotheses, builds increased competence. Secondly, there are stages of language-learning, and progress from stage to stage relies on a certain level of accomplishment in previous stages. For instance, Skehan’s work would suggest that noticing is needed before patterning can happen, and patterning is a first step to developing a working hypothesis of a rule that might be applicable in other instances, the step that Skehan calls ‘lexicalising’.

The understanding of aptitude developed from the work of comparing these findings with the findings of others, and with current theoretical explanations, is that aptitude is a complex construct that covers a combination of innate cognitive abilities that possibly operate interactively and interdependently in an SLA context. Then, while a number of components of aptitude have been designated empirically and used in the development of MLAT as an aptitude-testing instrument, this way of construction of a testing instrument is open to further refinement, and especially by close attention to the implications of theorising. This study may contribute to informing future aptitude studies.

5.1 The nature of aptitude: a stable or dynamic construct

The first research question addressed the nature of FL aptitude: whether it is stable or subject to change over time. There are two views on the nature of aptitude (see Chapter 2). One view is that aptitude is a stable trait not amenable to training (Carroll, 1981; Skehan, 1989, 1998; Politizer & Weiss, 1969; Sparks et al., 2001). The other view is that aptitude is a dynamic trait, changing with training inputs (Neufeld, 1978; Grigorenko et al., 2000; Ganschow & Sparks, 1995; Sparks et al., 1998; Sparks & Ganshow, 1993; Sparks et al 1992; Sáfár & Kormos, 2008). Research to date has produced equivocal results.
The results of the current study suggest that FL aptitude, in total, is not affected by English instruction. Our data analyses indicated that there was no statistically significant change between the mean differences of participants’ performance in the foreign language aptitude test at the beginning of the English course (T1) and their performance seven months later (T2). The foreign language aptitude total scores on both occasions (T1 and T2) were significantly correlated with each other ($r = .50$). Further, the findings showed that 20 hours a week of learning English over seven months did not create any change in aptitude: it was not sufficient, in neither length nor intensity, to modify FL aptitude. To demonstrate that language aptitude is sensitive to instruction, a longer and/or more intensive exposure to instruction would be required.

The findings of this study, relating to the evidence of stability of aptitude over time, are consistent with Carroll’s conceptualisation of FL aptitude as a measurable talent to learn a new language that is stable and not affected by L2 learning (Skehan, 1991). The current results also add to the support this view has received from previous studies. Put differently, our results do not lend support to a view of aptitude as a dynamic trait amenable to training.

There has been some discussion in the literature about the impact of the length of instruction and/or the intensity of instruction. For instance, in Sawyer et al.’s study (2001), there was no significant change in FL aptitude among Japanese students after studying a short, intensive English course. In Harley and Hart’s study (1997), there was no difference in aptitude levels between early immersion learners (twelve years of L2 exposure) compared to late immersion learners (four years of L2 exposure). While Harley and Hart’s study is of two different groups of students, and so might be subject to the impacts of other variables (and so less valid to compare), Harley and Hart’s principal
objective was to compare length of immersion, and its impact, if any, on aptitude. Their results showed that the group with the longer exposure to L2 instruction did not, on average, have a higher aptitude.

By comparison, other studies, e.g. Ganschow & Sparks (1995) and Sparks & Ganschow (1993), report a change in FL aptitude. In their studies, learners received intensive instruction in Latin over a whole year. In light of such results one could hypothesise that it takes a longer exposure to FL instruction (e.g. at least a year) for there to be any impact on learners’ aptitude. Such an assumption, however, would be in contradiction with Harley and Hart’s research, whose findings show that a very substantial difference in the length of FL instruction had no effect on the learners’ aptitude levels. So far as our study is concerned, 20 hours of English instruction over a seven-month intensive course is not sufficient, in either length or intensity, to modify FL aptitude.

The observation of a very small but statistically significant change in one of the components of the aptitude measure (VWC, the adaptation of the Spelling Clues section of the MLAT), which is discussed later, is the only evidence that might suggest the value of staying open to the idea of dynamism in aptitude.

Consequently, it is recommended that future research investigate the nature of FL aptitude under various conditions such as expanding the length of L2 instruction and exposing learners to more intensive L2 courses.

At the theoretical level, the assumption that aptitude is stable is the base on which its function as a predictor of proficiency relies. There are a number of different
conceptualisations of what ‘aptitude’ denotes or means. One such conception, suggested by Carroll, is that aptitude represents innate cognitive abilities that are used in language use and learning. Carroll chose, from a very comprehensive collection of those components of language learning, those that offered the greatest and distinctive correlations with language learning for his MLAT test. Skehan (2002), working from summarising the recent and accepted findings of SLA research, then re-examined the four MLAT aptitude component principles of Phonemic coding ability, Grammatical sensitivity, Inductive language learning ability, and Associative memory. Relying on an information-processing perspective, Skehan speaks of phases of learning — processing input, central representation, and output — and identifies a number of stages in SLA (though not necessarily in a linear chronological order): noticing, pattern identification, extending, complexifying, integrating, becoming accurate, avoiding error, creating a repertoire, achieving salience, automatising rule-based language, achieving fluency, and lexicalising. Skehan then shows how these stages might be mapped to aptitude, as tested by MLAT, and goes further to consider how MLAT might be developed, if this more extended model of SLA and its understanding of the aptitude-proficiency link is adopted.

By comparison, Robinson (2002) offers an even more intricate model of aptitude. He firstly enunciates an Aptitude Complex/Ability Differentiation Framework. This, in turn, is linked to his Fundamental Similarity Hypothesis of implicit, incidental and explicit adult SLA. Finally comes his Cognition Hypothesis of adult task-based language development, and together they provide a more comprehensive model for thinking about individual differences and learning experience design to deliver improved proficiency.
Clearly, more work needs to be done on developing instruments for measuring aptitude to both test these hypotheses and to develop a diagnostic tool that provides better information for streaming and/or individualised learning programming.

5.2 Foreign language aptitude as a predictor of English proficiency

The second research question in this study sought to examine the extent to which the foreign language aptitude was successful in predicting English proficiency over a seven-month English course. The findings showed that there was a positive, significant correlation between the foreign language aptitude test and the English proficiency. In the current study, the correlation analyses ($r = .27$ to $.34$) and the regression analyses ($R^2 = .07$ to $.12$) showed a moderate level of correlation, indicating that the foreign language aptitude test can be used as a predictor of English proficiency when these two variables are measured cross-sectionally or longitudinally.

Regarding the relationship between FL aptitude and L2 proficiency, research to date, again, has produced inconclusive results. Some studies have found that aptitude is not a significant predictor of L2 proficiency (Headrick, 1984; Hsieh, 2004; Rysiewicz, 2008; Goodman et al., 1990), whereas others have found FL aptitude to be a good ($r = .40$ to $.60$) predictor of L2 proficiency (Carroll, 1981; Ehrman, 1998; Sparks et al., 1995; Garcia, 1996; Ehrman & Oxford, 1995; Sparks et al., 2006; Morton, 2002; Harley & Hart, 2002; Phillips, 1998).

It is now necessary to explore, in greater depth, the relationships between the total scores of the foreign language aptitude test and the English proficiency test. Later, in Section 5.6, the inter-relationships of the foreign language aptitude components and the English
proficiency subtests will be placed in direct comparison with those of the two groups of studies above.

Generally the findings of the current study supported the positive and significant role of the foreign language aptitude test in predicting English proficiency. The correlation between FL aptitude, measured with the foreign language aptitude test, and the English proficiency, measured by TOEFL, was significant, but modest, in predicting L2 proficiency. The coefficient correlations of the total scores of the foreign language aptitude and English proficiency ranged between $r = .27$ and $.34$ (Tables 4.4, 4.6, and 4.8).

The obtained correlations are in the range of Dörnyei and Skehan (2003), who found that the correlations between FL aptitude tests and L2 proficiency ranged between $r = .20$ to $.60$.

The relatively lower correlations findings ($r = .27$ to $.34$) are in accordance with the findings of several studies with the same learning context: where English is a foreign language. For example, Rysiewicz (2008) investigated the correlation between the Polish language aptitude scores and the English proficiency scores in Poland (FL context). The magnitude of correlation was $r = .31$. In Thompson’s (2008) study FL aptitude did not correlate with the performance of 14 university students in a Portuguese proficiency test, although in this case the very small number of participants arguably makes it hard to make any statistically meaningful generalisations. In a Hungarian context, Sáfár and Kormos (2008) found a moderate correlation between FL aptitude test and L2 proficiency ($r = .36$).
The findings of these studies give some support to the view that aptitude-based tests have the capacity to predict L2 success in FL contexts.

The simple regression analysis showed that the total score of the foreign language aptitude test could predict between 7% and 12% of the variance of the English proficiency (Tables 4.5, 4.7 and 4.9). The current finding is in accordance with the findings of various other studies: in the study by Goodman et al. (1990), the French aptitude scores explained 3% of the variance in the English proficiency test; in Rysiewicz’s 2008 study, for Polish speakers, 9.61% of English scores variations could be explained by the variation in the Polish aptitude test (TUNJO) scores.

Our results are in accord with Dörnyei (2005) and Robinson (2005), who have claimed, from such results, that these low percentages raise the question of the adequacy of the traditional aptitude components to measure whatever cognitive abilities are needed for the attainment of a new language in a communicative context. In the absence of such adequacy, Dörnyei and Robinson suggest that ‘Aptitudes complexes’ (as discussed in Chapter 2, Section 8) is a better construct, with different aspects of aptitudes operating at different levels of SL learning. This will be examined more closely in Section 5.6.

There are also studies reporting a higher correlation between aptitude and L2 success than the findings of the current study. The range of correlation in these studies was between .40 and .60 (Carroll, 1981; Ehrman, 1998; Sparks, et al., 1995; Garcia, 1996; Ehrman & Oxford, 1995; Sparks, et al., 2006; Morton, 2002; Harley & Hart, 2002; Phillips, 1998).

The much lower correlation in the current study could be attributed to a variety of factors.

The low correlation values in our study may be due to the very low language achievement of the subjects (the mean TOEFL scores were 30.7% (at T1) and 35.8% (at T2) of the
possible score, and compared to 53–54% for the aptitude measure, see Chapter 4, Section 1). Since there is very little difference between achievement values at T1 and T2, this is quite likely to have disguised any genuine correlation between aptitude and achievement that would have been found if the language differences between T1 and T2 had been greater. It is, however, instructive to note that the correlation of the foreign language aptitude at T2 (with the increased role of VWC in the foreign language aptitude T2 measure) to English proficiency at T2 was the highest correlation ($r = .34, p<.01$) and accounted for 12% of the English proficiency variance.

Interestingly, a review of relevant literature reveals that most FL-context studies employing a non-English aptitude test produced relatively low aptitude-proficiency correlations (with the possible exception of Kiss and Nikolov (2005). Why such a pattern emerges remains unclear although it could possibly have something to do with the morphonological structure of the language involved and/or the design of the non-English aptitude test.

5.3 Development of English proficiency

The third research question investigated whether or not there was a positive, significant change in mean differences of English proficiency after exposure to English instruction for seven months. The findings revealed that the students’ scores in English proficiency at the end of the English course were positively higher (and the difference was statistically significant) than their performances at the beginning of the course. However, as the data in Table 4.10 show, in practical terms, the English progress of our population of learners was quite small: certainly a lot less than what one would expect after an intensive seven-month course.
One possible explanation for the marginal increase in proficiency in English, for our population of learners, may well be due to a combination of learner-external and learner-internal factors such as teaching quality, teaching methods, learner motivation, learner strategies, etc. It was not the purpose of the present research to investigate these factors — there is a voluminous body of literature discussing their importance for L2 achievement.

At a theoretical level, the marginal size of the progress in English proficiency could be explained as a result of a ‘Plateau Stage’. As suggested by Skehan (2002), among other researchers, at some stages of L2 learning, “some learners reach a plateau of development and find it difficult to move beyond particular areas of inter-language” (p.85)

*Figure 5.1* The S-curve for learning (*Adapted from Jovanovic and Lach, 1989*)

The S-curve of learning, offered in Figure 5.1, conveys different learning rates at different times or stages (Jovanovic & Nyarko, 1995). If something like the S-curve operates in this learning situation, then, depending on when T1 and T2 fall along the continuum of variable progress, there may or may not be much increase in learning that can be measured. An assumption taken into this study was that there is steady increase of L2
learning with time, as students applied their varying aptitudes to teaching inputs. Where there is not a steady increase in learning with time, the use of correlation patterns to examine predictability may produce misleading results.

By comparison with Carroll’s students, the participants in our study were not zero-English (T₀) proficiency students; they had studied English for six years in high school. In applying this curve, it is supposed that Carroll’s sample showed a faster and greater rate of progress because his students had no L2 previous experience. In the case of our sample of learners, it cannot be ruled out that they were in the plateau section for some or most of their period of language studies. It is necessary to acknowledge, though, that our study was not designed to examine progressive patterns of learning rates and language development, which is why our data are unrevealing with regard to the hypothesised plateauing of the learning process.

Thus, it is recommended for future study to consider the quality of English teaching and the length of L2 course to create more variations in L2 proficiency scores. A wider range of variations in L2 progress may be needed in order to be able to undertake any valid correlation study. A more detailed longitudinal study, with a larger sample, and a larger number of measures at reasonable intervals (say after every 200 hours of class input) would be needed to determine if the non-linear shape of the S-curve applies to L2 language learning.
5.4 The relationship between the foreign language aptitude measures and English progress

This section of the thesis deals with the relationship between the foreign language aptitude scores and the English proficiency progress. The concept of L2 progress in FL aptitude research has not been given sufficient consideration in many previous studies, and when it has been considered, progress is typically measured by subtracting the mean FL achievement score at T1 from the mean achievement score at T2.

\[ \Delta P = (\text{Eng. Prof. T2} - \text{Eng. Prof. T1}) \]

In the view taken here, as discussed in detail in Chapter 3 section 9.4, this does not constitute an appropriate method of calculating the rate of FL progress, because it only yields a value reflecting learners’ absolute progress. It would seem more appropriate to deal with learners’ relative progress, which is why different formulas to calculate progress have been used in this research:

\[ \text{Ratio 1 of progress} = \frac{(\text{Eng. Prof. T2} - \text{Eng. Prof. T1}) \times 100}{\text{Eng. Prof. T1}} \]

\[ \text{Ratio 2 of progress} = \frac{(\text{Eng. Prof. T2} - \text{Eng. Prof. T1}) \times 100}{\text{Eng. Prof. T2}} \]

The findings (Tables 4.12 — Absolute progress, 4.13—Ratio 1 and 4.14 — Ratio 2) showed that there was no significant correlation between the total aptitude or any of the aptitude components and the English proficiency progress. These findings are in line with Haslam (2010) whose study established no significant correlation between the FL aptitude test and the pronunciation progress among Chinese students in two different learning contexts (ESL and EFL). Notably, the study dealt with absolute, not relative, progress. The lack of correlation between FL aptitude and English progress in Haslam’s study
could be attributed to the shortness of the L2 course (three months). Also it is probably necessary to acknowledge that the phonological component of second/foreign languages is the least sensitive to instruction, especially with post pubescent learners.

By comparison, Erlam’s study (2005) found that MLAT was a good predictor of French progress, which was measured by the absolute formula. The regression analyses showed that FL aptitude could explain 24% of the variance in French writing test. This could be explained as a result of three factors: 1) The French proficiency test measured a wide range of communicative skills; listening comprehension, reading comprehension, written production and oral production. 2) In Erlam’s study the participants were high-school students who had been studying French for two years. 3) A different memory test was used; a working memory component was used instead of Carroll’s rote memory component. In addition, the Sound Discrimination test of the PLAB (Pimsleur, 1966) was administered to measure the students’ phonemic coding ability. These differences between the measures of L2 proficiency tests and the aptitude tests in both studies could be an explanation for the good correlation between aptitude and French proficiency in Erlam’s study.

The absence of the correlation between FL aptitude and English proficiency progress in our study could be explained as a result of various factors. First, one key issue is the overly-homogeneous nature of the study sample. Almost all participants in the present study share the same characteristics such as age, gender, level of education, and cultural backgrounds. This kind of homogeneity was likely to constrain the level of variance in all variables and so compromised the ability to detect relationship between variables. Second, there was a too small variance in English proficiency T2-T1 for meaningful correlation examination. Third, the teaching context involved a large number of smaller classes, all
following the same curriculum, and using the same standard resources. The context included non-native English teachers teaching in what is for them a FL, with unimaginative teaching curriculum and processes. The context also offered little incentive from other study assessment procedures to require functional English literacy in other study disciplines. A systemic avoidance of actual and extended use of English in the academic environment also occurred, arising from an understandable reluctance to teach in English for teachers of other disciplines with limited English language and literacy (see Chapter 1, Section 5). For students, limited motivation and competition for study time means that they make strategic decisions about how much effort to apply to gaining the requisite proficiency applied. Each and all of these contextual factors may well contribute to a context where learning is boring, and study exertions only extend to the compliance level.

A fourth possible explanation, operating at a hypothetical level, is that the learning measured between T1 and T2 involved some of the plateau period represented in the S-curve of learning that might apply to L2 learning, as noted in Section 5.3.

5.5 The foreign language aptitude and TOEFL tests as predictors of achievement

Because this study was longitudinal (in relation to which it differs from most aptitude research), and because it was conducted amongst Arabic students at the tertiary level, with prior English language learning, this study was able to compare the effectiveness of the foreign language aptitude test with an English proficiency test (i.e. the version of the TOEFL used in this study) for their potential to predict English-language achievement.

The interesting finding is that English proficiency, tested at the beginning of the course (T1), turned out to be a better predictor of L2 achievement after studying English for
seven months than the aptitude measure. It was highly correlated with the English language test scores at the end of the course (T2). The foreign language aptitude, by comparison, was a poorer predictor. The multiple regression analyses (Tables 4.15 and 4.16) showed that the English proficiency (T1) could explain 47% of variance of English achievement and the foreign language aptitude could account for 7% to 12% of variance of L2 achievement at the end of the course. This finding is in agreement with the results from a study conducted in Hungary (Nikolov & Ottó, 2006; as cited in Sáfár and Kormos, 2008).

The finding that the TOEFL test is a better predictor than the foreign language aptitude test, in this kind of context, may well prove useful to the more general field of language learning, and especially for Arabic speakers. However, the correlation between proficiency at T1 and T2 could also in part be explained with reference to the low achievement: those learners who had good FL skills at T1 still had those same good FL skills at T2 and applied those skills to deliver the limited change of proficiency between T1 and T2. Until there is greater empirical change in proficiency to measure, it will be difficult to see if there is any real correlation worth examining, and using as a reliable indicator of the instrument’s potential to predict achievement.

Consequently it is recommended that additional studies extend beyond T2 to confirm or disconfirm the usefulness of TOEFL as a predictor for students with some English language proficiency.
5.6 Subtests analyses

In this section more specific detail about the inter-relationships between the foreign language aptitude components and the English proficiency subtests were explored for more contextual and theoretical explanations of the current study results.

As noted in Chapter 2, Section 8, recent developments in theorising about aptitude and SLA, as conveyed in Skehan’s work (2002) and Robinson’s work (2002), suggest that at different stages of learning different aspects of aptitude may be needed. The current findings could be seen to give support to such theorising since there were some indications of difference in ‘behaviour’ of the various components of the aptitude total, over the course of the study, as shown in Tables 5.1 and 5.2 (derived from Tables 4.3, 4.4, 4.5, 4.6, 4.7, 4.8, 4.9 and 4.10). Any significant difference in the behaviour of subtests for aptitude, in measures taken at different times, tends to support the assumption that ‘aptitude’, as measured by MLAT, is a complex concept (i.e. more than one active component, and perhaps with interactivity between components), as per Robinson’s or Skehan’s theorising.

While Skehan’s theory and Robinson’s theories were not the focus of the present research, and the current study was not designed to examine the relationships between aptitude components and SLA stages, the computations conducted as part of this study might help test out the reasonableness of some of their theories of stages or hierarchies. What the variances found suggest is that investigating the relationship between aptitude components and SLA stages and/or expanding aptitude test design to capture different aspects of aptitude as indicated in the theorising, could be a promising area for future research.
Table 5.1

*Mean Differences for Aptitude Total and its Subtest Components and English Proficiency and its Subtest Components (N=56)*

<table>
<thead>
<tr>
<th>Measure (score possible)</th>
<th>Mean Difference T1 and T2</th>
<th>Mean Difference as a percentage of total score possible for the measure</th>
<th>SD</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aptitude total (64)</td>
<td>0.42</td>
<td>0.65%</td>
<td>7.04</td>
<td>.58</td>
</tr>
<tr>
<td>VWC (20)</td>
<td>0.99</td>
<td>4.95%</td>
<td>2.86</td>
<td>.01</td>
</tr>
<tr>
<td>Grammar Sensitivity (20)</td>
<td>0.25</td>
<td>1.25%</td>
<td>2.39</td>
<td>.43</td>
</tr>
<tr>
<td>Memory (24)</td>
<td>-0.75</td>
<td>3.13%</td>
<td>5.51</td>
<td>.31</td>
</tr>
<tr>
<td>English</td>
<td>3.09</td>
<td>5.15%</td>
<td>3.08</td>
<td>.00</td>
</tr>
<tr>
<td>Proficiency (EP) total (60)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Listening (20)</td>
<td>0.52</td>
<td>2.6%</td>
<td>2.72</td>
<td>.16</td>
</tr>
<tr>
<td>Grammar (30)</td>
<td>1.96</td>
<td>6.53%</td>
<td>2.98</td>
<td>.00</td>
</tr>
<tr>
<td>Reading (10)</td>
<td>0.91</td>
<td>9.1%</td>
<td>3.19</td>
<td>.00</td>
</tr>
</tbody>
</table>
Table 5.2

**Correlation Coefficients (r) of Regressions between Aptitude and its Subtests and English Proficiency and its Subtests at Time 1 and Time 2 (N=56)**

<table>
<thead>
<tr>
<th>Aptitude</th>
<th>English Proficiency T1</th>
<th>English Proficiency T2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>EPT1  L1  G1  R1</td>
<td>EPT2  L2  G2  R2</td>
</tr>
<tr>
<td>FLAT1</td>
<td>.29* .30 .24 .02</td>
<td>.27* .26* .23* .04</td>
</tr>
<tr>
<td>VWC1</td>
<td>.29* .31* .26* -.13</td>
<td>.17  .27* .16 -.10</td>
</tr>
<tr>
<td>GS1</td>
<td>.21  .21 .17 -.06</td>
<td>.14  .11 .16 .13</td>
</tr>
<tr>
<td>MEM1</td>
<td>.16  .16 .11 .18</td>
<td>.24* .18 .17 .13</td>
</tr>
<tr>
<td>FLAT2</td>
<td>.28* .36* .27* -.14</td>
<td>.34** .36** .26* .05</td>
</tr>
<tr>
<td>VWC2</td>
<td>.18  .28* .15 -.17</td>
<td>.21  .25 .24 -.04</td>
</tr>
<tr>
<td>GS2</td>
<td>.11  .11 .13 -.09</td>
<td>.26* .24 .20 .10</td>
</tr>
<tr>
<td>MEM2</td>
<td>.24  .30* .24 -.06</td>
<td>.23  .24 .13 .09</td>
</tr>
</tbody>
</table>

**Notes.** FLAT = Foreign language aptitude test; VWC = Vocabulary & Written Coding; GS = Grammar Sensitivity; MEM = Memory; EPT = English proficiency test; L = Listening; G = Grammar; R = Reading

* *p* = <.05; ** *p* = <.01

The different pattern of correlations between components of aptitude and L2 proficiency (Table 5.2) support Robinson’s aptitude complexes theory that hypothesizes that there is a dynamic interplay between aptitude and context. This theory proposes that there are multiple aspects of aptitudes involved in L2 learning under various conditions and mobilised for different tasks of learning.

Robinson has made two remarks that are relevant to this discussion of the nature of the inter-relationship between individual differences such as aptitude and SLA stages: 1) “As Segalowitz noted, while SLA research has produced a substantial body of correlational findings, it has been ‘relatively silent about … mechanisms’ (1997, p. 87)” and 2) “A major challenge for ID/condition interaction research, then, is to try and explain why
patterns of abilities lead to learning outcomes in any one context in terms of proposed SLA processes and mechanisms” (Robinson, 2002, p. 122).

In the current study, a body of ‘correlational findings’ was collected and extended to Time 2, for more informative, longitudinal data. The current study was not designed to examine any of the various mechanisms that could explain ‘why patterns of abilities lead to learning outcomes’, so providing an explanation was out of the scope of the current study.

From this closer look at the behaviour of the subtests in Table 5.1 and Table 5.2, a number of findings warrant closer attention:

- the change in the VWC measure between T1 and T2
- the lack of change in listening proficiency
- the absence of any correlation between reading proficiency and any of the aptitude measures at any time
- the absence of any correlation between Grammar Sensitivity (Aptitude measure) and Grammar proficiency measured by TOEFL
- the lack of correlation between memory and English proficiency.

5.6.1 VWC change and aptitude stability

Comparison between the mean differences relating to foreign language aptitude subtests (Table 4.2, and Table 5.1) revealed that there was no statistically significant change between the first and second assessment occasions for any of the components of foreign language aptitude test, except for the Vocabulary and Written Coding (VWC) component. This particular component appeared to be affected by the English instruction conducted between T1 and T2. However, the change (MD=.99) between the two mean differences was marginal. Figure 4.5 shows that this change, which amounts to one additional correct
answer in a suite of 20 items, as a mean change, did not affect the range of correct answers for this component (compared to changes in the ranges of Grammar Sensitivity and Memory for which the mean change proved to be both smaller, and lacking in statistical significance).

The change could be explained as a result of the difference between Carroll’s *Spelling Clues* component and the VWC used in this study. The *Spelling Clues* component, for the English MLAT, aims to measure learners’ vocabulary ability and to associate sound with spelling. As a result of the difference between Arabic and English language (phonetic/un-phonetic language), this component is used to measure the vocabulary and the association between written items and their spelling. The *Spelling Clues* test is about finding out if the learners are recognising that words with the same sounds may have different meanings, and that the learner needs to look for clues from the context to decide which of the options of meaning apply in the context. Then, in Arabic, the written code also includes diacritics, which convey additional information about meaning. One possible explanation for the change in the VWC component is that over the period T1–T2, as a consequence of exposure to another language, learners have become more sensitive to graphemic features of their primary language, e.g. become more aware of the significance of diacritics.

5.6.2 Lack of correlation between Grammar Sensitivity (aptitude measure) and Grammar proficiency measured by TOEFL

Another finding emerged when the grammar sensitivity was correlated with grammar subset: no significant correlation was found between these two components. This finding does not support Carroll’s proposal (1990) when he pointed out that the Words in Sentences part in the MLAT which measures Grammar sensitivity is the best predictor of
grammar learning. Carroll’s grammar sensitivity was defined as “the individual’s ability to demonstrate his awareness of the syntactical patterning of sentences in a language (Carroll 1973, p. 7)”. Carroll (1990) proposes that the grammar sensitivity component does not directly predict students’ grammatical competence, but it involves a meta-awareness of various grammatical patterns. Our finding also goes against the findings of a study conducted in Japan (Sawyer et al., 2001) that showed that there is a strong correlation between grammar instruction and Words in Sentences (grammar sensitivity subtest in this study).

The lack of correlation between the grammar sensitivity component and TOEFL-Grammar subtest in this study is most unexpected. Broadly speaking, the Grammar Sensitivity component of Carroll’s test measures the learners’ metalinguistic awareness; more specifically it measures their ability to perceive the grammatical functions of different parts of the sentence. The emphasis in the teaching context, on grammar, as described in Section 5.4, suggests that there ought to have been a greater change in grammar proficiency than there was. However, as also noted in Section 5.4, the size of change may not have been sufficient to be able to show any correlation. The correlation between Grammar Sensitivity from the Aptitude measure at T2, and the total English proficiency at T2 is noted. The correlation between the total aptitude measure at T1 and the grammar component of English proficiency at T2 is noted. The correlation between the total aptitude measure at T2 and the grammar component of English proficiency at T1 and T2 is noted. A plausible explanation for the variations in these findings could not be provided. More investigations need to be undertaken in future studies to explore this unexpected result.
5.6.3 Absence of correlations – implications for the measuring of Memory

Memory (Paired associates in MLAT) did not show any significant correlation with the English subtests (Tables 4.4, 4.6, and 4.8). It is interesting to find that rote memory skill, which requires memorisation of lexical items, is not an aptitude component which can be used to predict L2 success in the present study’s context. These findings are consistent with a number of previous studies (Skehan, 1986a; Hummel, 2009). Based on the current findings and the findings obtained from the previous studies, it can be suggested there is no need to include the rote memory components in the development of an FL aptitude test in a similar context where students have some L2 learning experience, since they do not contribute to predicting L2 success. In line with this suggestion, Carroll (1990) himself (the developer of MLAT) was in doubt about rote memory task’s usefulness.

One of the possible explanations relates to the design of the Memory component in the current study (Paired Associates in MLAT). This component is criticised for simply assessing memory decoding. For memory components to be more effective in SLA, it should include more memory-related cognitive abilities such as storing and retrieval (Skehan, 2002).

5.6.4 Proficiency measures noted

The relatively low progress in Listening proficiency from the seven-month intensive English course is probably best understood to be a result of the learning conditions as described in section 5.4 (teaching context, or the plateau effect).
The lack of correlation between any aptitude measure and the reading proficiency measure at both times is probably best understood to be related to the small size of the reading measure.

In future research, developing the aptitude instrument to include the oral components of language, and developing the TOEFL to have better balance of Listening, Grammar and Reading components, and to include an oral component will be needed to provide information to allow the kind of theory testing that will be needed to build a better understanding of the linkage between aptitude and SLA, and to develop an aptitude instrument that has greater predictive capacity.

5.7 Summary

After discussing the main issues of the current study which were posed at the beginning of this chapter, the outcomes of this chapter can be summarised in the following way: first, the foreign language aptitude test has been shown to be a valid and reliable test yielding results that are on par with other non-English aptitude tests. This test could support Arabic native-speakers with a tool to be used freely in the future research for the same target population. Second, the foreign language aptitude test positively and significantly predicts the English proficiency among Saudi student during a seven-month English course. Third, the target sample showed a statistically significant measure of English progress after learning English for seven months but this progress is small. Fourth, the FL aptitude test failed to predict the English progress. Fifth, the English proficiency test at Time 1 was the better predictor of the English proficiency at Time 2 compared to the foreign language aptitude test. Finally, these two variables in the current
study hold some implications in the field of L2 learning; they can be used as diagnostic tools or for the purposes of placement where larger learner intakes are involved.

The findings of the current study have shed some light on various theories related to aptitude and SLA. The inconsistent pattern of correlations between aptitude components and English proficiency on both occasions would appear to support the theoretical positions of Skehan and Robinson concerning the existence of a number of different aspects of aptitude, which play different roles in learning, either at different stages or in different contexts and conditions. While the study was not designed to test neither Skehan’s view nor Robinson’s interlocking hypotheses, our data contain some evidence that ‘noticing’ did have an impact on the VWC measures before and after seven months’ intensive English learning. The lack of an expected correlation between Grammar Sensitivity (Aptitude Measure) and Grammar proficiency tested by TOEFL, in a teaching context where emphasis on grammar was high, remains a puzzle until further research is undertaken. The use of the rote memory (paired associates) section of MLAT, in the context of these students’ learning stage, was also shown be uninformative, supporting Skehan and others’ calls for designing a tool that measures memory, in its different forms, in other ways.

In all, the combination of evaluating the findings, in the context of the current literature of empirical research findings and theoretical considerations, suggests there is much merit in following Skehan and Robinson (among others) in reconsidering how to design aptitude measurement and how to examine the nature of the relationship between what is measured by the aptitude measuring tools and what occurs in SLA.
Chapter 6

CONCLUSION

6.1 Research purpose and questions

This study has been concerned with the development of appropriate tools to support the learning of English as a second language in universities in Saudi Arabia. One such tool, of long standing, has been ‘aptitude testing’, which has been used to identify students most likely to benefit from language teaching, or to stream students of similar ability and to focus teaching techniques to learning needs. In the context of this study an aptitude test for native Arabic L1 learners needed to be developed. Then it needed to be tested to see if it performed at a level equivalent to similar tools developed for English or other L1s (Japanese, Chinese, Hungarian, Polish, etc.). In addition, the current study has investigated various important issues concerning the relationship between the traditional concept of FL aptitude and L2 proficiency, for Arabic L1 learners of EFL.

Four questions guided the study:

1. Is the foreign language aptitude testing instrument in the current study a valid test, at a level similar to other equivalent instruments?
2. Do the foreign language aptitude test scores at the beginning of the English course (Time1) differ from the foreign language aptitude test scores at the end of the seven-month English course (Time 2)?
3. Does the foreign language aptitude test predict English-language learning proficiency to a similar standard as other aptitude tests?
4. Do the foreign language aptitude test scores correlate significantly and positively with English proficiency progress after a seven-month intensive English course?
6.2 Empirical findings

Having constructed and piloted an adapted MLAT test, modified for pencil and paper testing and for L1-Arabic participants, some 56 tertiary students from a Saudi Arabian university participated in the study which used a test design which was intended to overcome some of the limitations of typical aptitude testing designs. Firstly, a cross-sectional design, conducted at both Time 1 and Time 2, was used to examine the relationships between the foreign language aptitude measures and the English proficiency measures. Secondly, a longitudinal design, which aimed to investigate the relationships between variables before and after the English course, was used. The novel contribution of this study was in using the longitudinal design to examine the relationships between the foreign language aptitude and English progress during the English course.

The development of the foreign language aptitude test for Arabic native speakers is one of the main contributions of the current research. The empirical findings of this study indicated that

- The foreign language aptitude test, adapted from the MLAT, worked well. It achieved a similar magnitude of correlation with L2 success (.27−.34) as most previous modified aptitude tests for languages other than English, in the field of FL research.

- The correlations were in general modest.

- The new aptitude test was able to predict English proficiency among Arab native speakers, within limits similar to other adapted instruments.

- The nature of FL aptitude, as measured by the foreign language aptitude test, was stable and not affected by the seven-month English course. This finding is in line
with Carroll’s traditional conceptualization of aptitude. Based on this finding it could be stated that FL aptitude is an innate ability and it is not a skill affected by L2 instruction.

With regards to the role of the foreign language aptitude test in predicting L2 proficiency, the current study, unlike previous studies, addressed all of the cross-sectional and the longitudinal relationships between the FL aptitude and L2 proficiency in the same study. All of these cross-sectional and longitudinal slices of data aimed to build a clearer idea about the role of the foreign language aptitude test in predicting the English proficiency among the Saudi students in the target sample. The findings of the correlation analyses and regression analyses revealed that the foreign language aptitude test had a modest power in predicting English success.

The development of L2 proficiency among the target sample, after studying English for seven months, revealed that there was a statistically significant change in the participants’ English competences after this exposure. Despite the statistically significant change in English proficiency, these findings were disappointing and surprising because the progress was marginal.

The new contribution of our study, which has not been discussed widely in the previous FL aptitude research, was investigating the relationship between the FL aptitude and English progress. Besides using the FL aptitude–L2 progress relationships, this study was novel in using the ratio of L2 progress instead using the absolute index of L2 progress. Despite the significant correlations between the foreign language aptitude tests and the English proficiency in the cross-sectional and longitudinal designs, the foreign language aptitude test failed to correlate with the English progress.
At this stage, this lack of correlation has been attributed to the smallness of the size of increase in proficiency, which was a surprising result, as mentioned above, making correlative comparisons difficult.

During the process of examining these four aspects, two further matters came under consideration. Firstly, a comparison of the foreign language aptitude test and the TOEFL, amongst a group of students with prior English learning experience, was possible, in order to determine which test offered the better prediction of English proficiency in such circumstances. Secondly, some consideration of what was happening for the subtests of both the foreign language aptitude test and the TOEFL was possible.

The moderate power of the foreign language aptitude test, in predicting the English proficiency, raised the question of what might be a better alternative predictor of English proficiency. The findings of the regression analyses for all of the variables in the current study showed that English proficiency, as measured by the modified TOEFL test, at the beginning of the course was the best predictor of English proficiency at the end of the seven-month English intensive course. Despite the high correlation between English proficiency at T1 and T2, this finding is limited because of the marginal change in English proficiency after studying English for seven months.

Consequently, it is highly recommended for future studies, to measure the English proficiency by TOEFL in more extended times (T2, T3 etc.) to ascertain the usefulness of using TOEFL as a valid predictor of L2 success.

6.3 Theoretical considerations

As noted, the information from the findings for the subtests of both aptitude and proficiency were available to be examined, and for material that might inform our
understanding of the role of aptitude in SLA. More recent literature has begun to theorise about the connection between what is being measured when measuring aptitude and what might be involved in SLA. The findings in this study showed signs of supporting the theorising of Skehan and of Robinson, who, like others, have raised questions about the efficacy of the MLAT construction as it presently stands. Further, both Skehan and Robinson have made suggestions about how the MLAT might be developed, and in a way that helps link aptitude measures to learning processes and contexts for SLA.

While the MLAT construction (which informed the adaptation to this Arabic tool) is still the most effective aptitude test, empirically, the consideration of what is going on in SLA suggests that this tool might be improved if there is a clearer match of the different aspects of the tool to the nature and context of learning. The way the correlations between different aspects of the aptitude test varied with different components of the TOEFL in this study, tends to support the contention that more work on developing aptitude testing, in line with the theorising, and in a way that will allow closer testing of the current theories, will be productive for the field.

6.5 Implications for FL aptitude and SLA research

There are new directions for the future FL aptitude research. One of the new directions is to find a link between various aptitude components and the stages of SLA. The foreign language aptitude components in the present study revealed a possible connection with the stages of SLA. Other studies are important to develop new aptitude components to fit with more stages of SLA as suggested by Skehan (2002). Secondly, the role of context and the dynamic interplay interaction between the complex aptitudes and the L2 situations
and types of learning should be considered for future research as suggested by Robinson (2002). A new area of research should investigate the impact of various cognitive abilities in different L2-learning conditions. In line with Sparks and his colleagues, the impact of L1 on both FL aptitude and the ultimate success in L2 is a promising area of research. In addition, the role of working memory as suggested by Miyake and Freedman (1998), and other measures of memory processes may need to be used in future investigations of ID variables.

As noted in Chapter 2, Section 5.3, the association between Arabic literacy and English proficiency has not been investigated in this, or other studies. Consequently, it is recommended that future research investigate the impact of Arabic literacy on English learning in the Saudi context.

6.6 Practical implications of the current study

Implications for FL pedagogy

The present study has some practical implications for FL pedagogy in the Saudi context where the present study was undertaken and for any other similar educational context. The most common characteristics for this context are large classes, unmotivated students, paper-based evaluation, deemphasised oral activities, teacher-centered classes, and a tendency to teach grammar out of context. These practices in the FL instruction led to a noticeable weakness in the level of English proficiency among the first-year students at King Khalid University, despite the statistically significant change in English proficiency. The findings of this study suggest that any endeavours to generate quick and effective change in the English program in the department will be risky. Unfortunately, the size of
progress with English language learning, which has been so disappointing, prompts a call for change, from policy-makers and teachers.

More research should be conducted to evaluate the English program in the college. What are needed are changes that are appropriate to the cultural situation and operate on the best evidence-based research for this level of language learning. Offering testing, by using TOEFL as external formative feedback on current English language performance, might assist students in managing personal learning objectives and taking additional interest in, and responsibility for, the language learning they need for their ongoing academic and professional life.

*Implications and Possible applications of the foreign language aptitude test*

The foreign language aptitude test is stable and operates at a similar level of predictive power as other aptitude instruments adapted from MLAT, and while not the best predictor of the English proficiency for this group of students, should still be able to be used with students for considering their L2 learning capacity, and possibly for streaming into different course structures/classes for more or less intensive language teaching, before undertaking any L2 studies.

*Using TOEFL as a diagnostic tool*

This study indicated that the TOEFL test (English proficiency at Time 1) can be conducted amongst students who have been exposed to English language learning, and is a better predictor of further English language learning at the end of the course than the foreign language aptitude test at Time 1. Consequently, for institutions offering further English language learning for ongoing academic or vocational preparation, conducting the
TOEFL test at the commencement of the additional studies might assist teachers and students in identifying individualised learning plans to focus further learning effort.

6.4 Limitations of the current study and recommendations for future research in the same area

As noted earlier the current study has various significant contributions to the field of FL aptitude research, namely: combining the two correlational designs—the cross-sectional and the longitudinal designs—in the same study; developing a FL aptitude test for Arabic native speakers; considering L2 progress in predicting L2 proficiency. Typically for this type of research, the project design was subject to a variety of limitations, often of practical nature, which need to be recognised in order to put the research findings in perspective.

Generalisability

The issue of generalisation is a common limitation in most studies, including in the current study. The target sample in this study was limited to the first year male students in the English department at King Khalid University. It is therefore necessary to be cautious in interpreting the results’ applicability to other populations of learners and other learning contexts.

Because of cultural sensitivity, it was difficult to access female participants since in Saudi Arabia women are segregated from men in all government universities. Investigation of the role of the foreign language aptitude test in predicting English proficiency among female university students will be needed to know if the foreign language aptitude test works to the same extent, when used amongst female students, so that the predictive power of this measure can be generalised. The indications from other FL aptitude studies
in various contexts would suggest that gender will not make a significant difference, but it is highly recommended to investigate the impact of this factor in the Saudi context in future studies.

To increase the generalisability of these findings, it is recommended to conduct the same study with various students (males and females) from different university departments in Saudi Arabia and with those students in the higher levels of English in the college.

*Adaptation of MLAT and TOEFL tests in this situation*

Another limitation of this study that should be taken into account in any future research is the validity of the measurements of the foreign language aptitude test and the English proficiency test, TOEFL, because these measures were modified and adapted from the original ones. These modifications were done to suit the culture and the language of the target population. As described in Chapter 3, the two instruments used in our study were designed to reflect the constructs of the original ones, but it cannot completely be ruled out that the amendments may have affected their validity and reliability. In the case of the foreign language aptitude test, it is important to note that the adapted foreign language aptitude test form (MLAT—short form) in the current study did not address the oral abilities that are measured by these two components in MLAT: Part 1: Number learning and Part 2: Phonetic script.

It is recommended to include these two oral measures to the foreign language aptitude test to shed more light on the possible further impact of this ability on the relationships between FL aptitude and L2 proficiency.

Using TOEFL as a standardised English proficiency measure in the current study meant that only three competences were measured: listening, grammar and reading.
For future research, more competences such as speaking and writing should be investigated to measure the English proficiency.

*Using correlational design*

As with most individual differences (IDs) studies, the present study is a correlational study based on a body of ‘correlational findings’ (Robinson, 2002) collected and extended to Time 2. While this study could examine and measure the level of correlation between FL aptitude and L2 proficiency, it could not confirm a causal relationship between these two variables. Also, it was similarly limited in providing an explanation of the various mechanisms that could explain why patterns of abilities lead to learning outcomes (Robinson, 2002).

Further studies with different designs are needed to investigate the nature of the relationship between these variables and to determine the possible causal link between FL aptitude and L2 proficiency.

**6.7 Final remarks**

This thesis has some contribution in the field of FL aptitude research. I hope to launch new research in this context, based on the findings of the current study. These studies aim to move the traditional aptitude test from its purely empirical nature to be based more closely on a theoretical foundation related to the main stream of SLA. To achieve this goal, this future research will focus on two main issues: practicality and research refinement.
Practicality

First, the current study will be extended to another couple of longitudinal times to indicate if there is a possible impact of the hypothetical ‘plateau’ stage idea/theory on English development. Also, the extension of this study may provide valid information for considering the stability of aptitude and to allow current students to have more chances to increase their English competences through exposure to more L2 instruction. The possible increase in English progress may create a stronger correlation between the foreign language aptitude test and the English proficiency.

Research refinement

Second, in light of the findings in the present study, and especially by attending to the theorising of Skehan, Dörnyei and Robinson, some changes in the foreign language aptitude test and the English proficiency test will need to be considered in any future research. These changes will include the following: developing the Memory component in the foreign language aptitude test to replace it with a working memory component; and including the oral subtests to both the foreign language aptitude test and TOEFL.

In the foreign language aptitude test, the two oral aptitude components in MLAT (Part 1: Number learning and Part 2: Phonetic script) will be included to shed some light on the possible further relationships between FL aptitude and L2 proficiency. In addition, oral subtests such as speaking test will be added the English proficiency test (TOEFL) to measure more communicative competences.

Further, the design of the current study will be developed to consider not only the correlational relationships but also the direction and the causality. I intend to investigate
the causal relationship between aptitude and L2 proficiency and to exclude the possible impact of other external factors such as motivation, age etc.

These changes represent the first stages of refining future research, in order to gather more meaningful data.

Further refinement will be possible by more closely attending to the directions offered by the theoretical explanations enunciated by Skehan, Dörnyei and Robinson, and any research developed in the field in more recent times, that looks to use research designs that allow us to test current explanatory hypotheses. Any particular suggestions for developing aptitude testing by using different and more useful memory units that test out these theories will be one priority. Any suggestions of relevant oral units—for aptitude measurement and for TOEFL measurement—associated with testing these theories would be another priority. It is also expected that different research designs will be needed to test hypotheses developing out of theoretical explanations. Some of these theoretical and practical considerations can be explored while doing the current practical extension.

With a more developed foreign language aptitude test, and an expanded TOEFL, the next step will be to undertake another round of primary data collection with a new design. Using improved instruments, and starting at T₀, and by working with the same ‘size’ units of language learning input (similar amount of content and intensity) through T₁−T₄, data will be collected to help us answer questions about what is going on in second-language acquisition, and how the different cognitive abilities measured by aptitude instruments, operate to deliver successful learning in different contexts, and for different stages of language learning. In this way, English language learning in Saudi Arabian universities can be developed to produce more effective outcomes.
BIBLIOGRAPHY


Keitges, S. (1986). *The relationships between foreign language aptitude, attitudes and motivations and personality traits and the attained English language proficiency of Japanese university students*. The University of Texas at Austin Austin.


## Appendix 1: A summary of all the previous foreign language aptitude studies

<table>
<thead>
<tr>
<th>Study</th>
<th>Data Sources</th>
<th>Country</th>
<th>Data Collection</th>
<th>Magnitude of correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bialystok (1978)</td>
<td>157 grade 10 - 12</td>
<td>USA (French)</td>
<td>MLAT Motivation/attitude</td>
<td>R^2 = .28</td>
</tr>
<tr>
<td>Headrick (1984)</td>
<td>69 University of Missouri-Columbia Elementary German I students</td>
<td>USA -German</td>
<td>Aptitude test attitude questionnaire 4 course examinations and a final examination</td>
<td>No Significant Correlation</td>
</tr>
<tr>
<td>Gardner et al. (1984)</td>
<td>170 university students</td>
<td>Canada (French)</td>
<td>MLAT French proficiency Motivation and attitudes measurement</td>
<td>r = .40</td>
</tr>
<tr>
<td>Keitges (1986)</td>
<td>168 private university (124 F, 44 M)</td>
<td>Japan (English)</td>
<td>-Aptitude test for Japanese-Michigan Test</td>
<td>r = .18-.27</td>
</tr>
<tr>
<td>Gardner (1989)</td>
<td>89 six-week immersion French program</td>
<td>Canada (French)</td>
<td>MLAT short form French proficiency test</td>
<td>No significant correlation</td>
</tr>
<tr>
<td>Ehrman (1994)</td>
<td>343 Students FSI (Foreign Service Institute). Age 37, male, beginners, intensive L training. and female, BA and MA, Language 1.8,</td>
<td>USA, (Multi-languages), MLAT for native speakers of English. Here it is used with non-native (validity).</td>
<td>r = .60</td>
<td></td>
</tr>
<tr>
<td>Ehrman &amp; Oxford (1995)</td>
<td>855 persons, 55% M, 45 % F. government Departments</td>
<td>USA</td>
<td>MLAT (long form) Proficiency test at the end of the course (speaking, reading, listening)</td>
<td>r = .55</td>
</tr>
<tr>
<td>Study</td>
<td>Data Sources</td>
<td>Country</td>
<td>Data Collection</td>
<td>Magnitude of correlation</td>
</tr>
<tr>
<td>-------------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>---------------------</td>
<td>---------------------------------------------------------------------------------</td>
<td>--------------------------</td>
</tr>
<tr>
<td>Sparks (1995)</td>
<td>154 9th and 10 grade girls (private single sex high school) -100 9th grade in public school</td>
<td>Spanish French</td>
<td>MALT English grade Academic achievement The grade in 8th grade English 5 native language measure</td>
<td>r=.55</td>
</tr>
<tr>
<td>Harley &amp; Hart (1997)</td>
<td>65 11grade students,</td>
<td>Canada (French)</td>
<td>Memory : MLAT 5 Words Pairs Memory Test Analytical Ability :MLAT Words in Sentences and PLAB 5 language analytic ability L2 proficiency tests: Voc, listening, writing, oral</td>
<td>No significant correlation between FLA (Analytic ability) and SLP among children and no significant correlation between the memory ability and the adult learners.</td>
</tr>
<tr>
<td>Nagata (1999)</td>
<td>177 students English as a FL ( 30 f, 147 M) 18-22 y.</td>
<td>Japan (English)</td>
<td>Pre-test 75 English words Aptitude tests (Kurdish Memory test ( Skehan,1682,MLAT part 5paried association, part 4 Words in sentences, PLAB parts 5, 6 sounds)</td>
<td>r=.40 -60</td>
</tr>
<tr>
<td>Harley and Hart (2002)</td>
<td>33 (24 F, 7 M) secondary school students</td>
<td>In Canada , English (French) 3 month French program</td>
<td>Language aptitude tests(Memory and analytic tests) French proficiency test</td>
<td>r=.60</td>
</tr>
<tr>
<td>Hsieh (2004)</td>
<td>200 university student</td>
<td>Taiwan (English)</td>
<td>TOEFL Motivation questionnaire MLAT</td>
<td>r=-.17</td>
</tr>
<tr>
<td>Study</td>
<td>Data Sources</td>
<td>Country</td>
<td>Data Collection</td>
<td>Magnitude of correlation</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>------------------------------------------------------</td>
<td>------------------</td>
<td>---------------------------------------------------------------------------------</td>
<td>--------------------------</td>
</tr>
<tr>
<td>Erlam (2005) FL Aptitude# 3</td>
<td>60 high school age Direct object pronouns</td>
<td>New Zealand (French)</td>
<td>Language Aptitude Skeh (1998) L2 proficiency( listening, reading, writing, speaking)</td>
<td>r=.52</td>
</tr>
<tr>
<td>instructions( deductive,</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>inductive, structured input)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kiss et al (2005)</td>
<td>Hungary, Hungarian</td>
<td>419 12 years Hungarian student from 10 schools.</td>
<td>Revised aptitude test -Background questionnaire -English proficiency test( listening ,reading ,writing) -Motivation scale (Nikolov,2003)</td>
<td>r=.634</td>
</tr>
<tr>
<td>Developing, piloting,</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>validating , Aptitude test for Hungarian student</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ehrman (2006) Aptitude,</td>
<td>855(55% male-45% female) from different government agencies. Age 39. MA and PhD.</td>
<td>USA (34 languages)</td>
<td>Questionnaire MLAT Motivation- Affective Survey(Ehrman &amp; Oxford, 1991) Language proficiency (Reading, Speaking)</td>
<td>r=.60</td>
</tr>
<tr>
<td>learning strategies, L style,</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>personality, motivation,</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>anxiety # language proficiency</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rysiewicz (2008) Aptitude</td>
<td>Piloting study: 200 Secondary school students age (18). 2nd study 250 aged 18-22</td>
<td>Poland (Polish)</td>
<td>TUNJO :Polish (MLAT)</td>
<td>r=.31</td>
</tr>
<tr>
<td>Polish (MLAT)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amy S. Thompson (2008) Aptitude and course scores and previous language experience</td>
<td>14 university with different linguistic backgrounds, studying Portuguese at first semester</td>
<td>USA (Portuguese)</td>
<td>MLAT Part 4 (words in sentences) -course scores -questionnaire</td>
<td>No correlation</td>
</tr>
<tr>
<td>Sparks (2009)</td>
<td>77(11 M , 66F) French adults</td>
<td>French (L2, English)</td>
<td>MLAT short form (French version) Michigan Test (Vocabulary, reading and Grammar)</td>
<td>r=.25-.29</td>
</tr>
<tr>
<td>Aptitude, phonological</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>memory, and L2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix 2: Foreign language aptitude test

في السؤال التالي سوف تجد (الاتجاه) مجموعة من الكلمات في كل فقرة والكلمة التي في بداية الفقرة غير صحيحة إملايا، وذلك عن طريق كتابة الكلمة بدون نقطة، والمطلوب منك في هذا السؤال التعرف على الكلمة الصحيحة، والبحث عن المرادفة لها من بين الأربعة الخيارات الموجودة تحت كل كلمة. وعند التعرف على الكلمة الصحيحة المطلوب منك اختيار الإجابة الصحيحة في ورقه الإجابة.

الرجاء منك عدم كتابة الإجابة في ورقة الأسئلة:

مثال:

حجر
أ اسم فاكهة صيفية
ب تحديد حرية الحركة
ت تعلم ببطء
ث مدينة عربية

الإجابة الصحيحة (حجز) والعبارة المرادفة لها الاختيار (ب) تحديد حرية الحركة.

الآن أيداعيًا للإجابة على الأسئلة التالية:

1. حبار
أ عنصر كيميائي
ب الشخص الذي يسكن بالجوار
ت أثاث منزلي
ث نوع من المطبخ

2. حبار
أ حيوان بحري
ب مدينة سعودية
ت تنوع في المثل
ث نباتات عطرية

4. الحياة
أ الموت
ب الشاب الخلفي
ت المنذون
ث مكان للتخزين

6. نهر
أ بارد جدا
ب يستغل
ت الأمواج
ث يحرك بقوة

7. حفيف
أ مرض نفسي
ب عكس ضعيف
ت صوت الحياة
ث خصائص فيزيائية
8. ثلج
أ. اسم مدنه
ب. نوع من العطور
ت. الاستقامة
ث. الكيفي النيبي

7. هندسة
أ. نوع من الأكالات
ب. حيوان بحري
ت. من أدوات الزينة
ث. توقف عن الحرب

10. الشعر
أ. آيات منزلي
ب. تضاريس جغرافيه
ت. مرفقات أشياء
ث. حيوان بحري

8. ثلج
أ. اسم مدينه
ب. نوع من العطور
ت. الاستقامة
ث. الكيفي النيبي

9. ذهب
أ. معنى مدين
ب. ركود أقصادي
ت. يستثير
ث. أدب إسلامي

12. ضرير
أ. حالة مصر
ب. كوارع من المعان
ت. آثار
ث. مدافن أعمى

11. زلال
أ. شديد الصفاء
ب. إهتزاز الظراية
ت. ظاهرة صحيحة
ث. مدينه أبوبية

14. حمرة
أ. اسم مذكر
ب. كميري عنصر
ت. عكس الضوء
ث. اسم بحيرة

13. اختبار
أ. الاستمرار
ب. العمل الجمعي
ت. تنظيم الحدود
ث. تنفيق النشوة

16. أحده
أ. اسم مئثث
ب. فعل ماضي معنى يحصل على
ت. عدد من الأعداد
ث. يجمل

15. الحرب
أ. التعب
ب. نوع من الفاكهة
ت. مدينه البهن
ث. يكلل سياسي

18. الحسر
أ. اسم سورة قرآنية
ب. الترف
ت. اسم فاكهة
ث. الحوت

17. الحبل
أ. بحيرة
ب. جمع معنى خداع
ت. مستنقع
ث. اسم مذكر يعني الجبيل البركاني

20. محمد
أ. الاتصال اللاسلكي
ب. نبخر
ت. البركان
ث. طريقة لحفظ الأغنية

19. عبادة
أ. مدينه خليجية
ب. قسم من أفهام المستشفى
ت. نوع من البهارات
ث. اسم بمعنى الثقة
السؤال الثاني
الكلمات في جمل

هذا الجزء من الاختبار يقيس مدى فهمك للوظائف النحوية للكلمات والعبارات في الجمل.

فضلاً انظر ( أنظري) للمثال التالي:

عاشة تقطع النفاخ بالسكون
أخي أحمد يضرب الحصان بعصا طويلة.

في الجملة الأولى تلاحظ ( تلاحظين) أن كلمة/نفخ هي التي وقع عليها الفعل وهو القطع، وفي الجملة الثانية كلمة/الحصان أيضًا هي التي وقع عليها الفعل وهو الضرب. إذن تستطيع أن تقول أن الإجابة (د) هي الإجابة الصحيحة.

لا ان كلا الكلمتين في محل نصب مفعول به.

المثال الثاني:

المال هو كل ما يطمغ إليه.

قبل سنوات قليلة أغلب الفلاحات كانت بدون.

أب ج د هـ

الإجابة الصحيحة هي/الفلاحات لأنها تؤدي نفس الوظيفة اللغوية لكلمة/مال في الجملة الأولى وبناء على ذلك الإجابة.

(د) هي الإجابة الصحيحة.

الآن ابدأ الإجابة على الأسئلة التالية:

1. أحمد سقط على الأرض في فارس تعثر أيضا.

أب ج د هـ

الآن يمكنك الانتظار خارجا أو العودة يوم الجمعة صباحا إذا أردت.

2. كان البعض غانبا عن الاجتماع.

أب ج د هـ

بالرغم من العروض الكثير التي قدمت ، واحد نظيف من الممكن أن يوافق عليه.

3. حانم بعث بدره دراجته الهوائية.

أب ج د هـ

أذا كان عملهم بريقي للمسئو المطلوب فإلي اضمن لهم الحصول على مكافأة آخر الأسبوع.

4. أصحى صار منتفخ من الالتهاب.

أب ج د هـ

الجدار العالي أصبح مخيفا عن الظلال بسبب الأشجار الشائكة.
5. قاد سيارته من الرياض إلى المدينة.

6. سوف اشترى سيارة عندما يكون لدي المال الكافي.

7. بعد مغادرتك ليلة البارحة. أغلب الطلاب مكتَ من الخياطة الحفل.

8. بعض الموظفين قدموا أوراقهم للوظيفة الجديدة.

9. أثق اللوحة على الحائط الخاص.

10. ما هو اللون الذي يفضله أكثر؟

11. نحن نخطط لدراسة الهندسة هذه السنة.

12. المقابلة الأسبوعية عادة تقام يوم الجمعة مساءً؛ بشكل مبكر وتعتبر منشطًا ثابتًا لبرنامج الكشافة.

13. كان أحمد هنا.

14. رجع البعض من السفر.

"صعود خمسة بعدها، دخل الرجل المنزل وهو يتنفس بسرعة كبيرة وعلى أثر التعبد.

"بعد صعود خمسة طوابق. دخل الرجل المنزل وهو يتنفس بسرعة كبيرة وعلى أثر التعبد.

"بعض الموظفين قدموا أوراقهم للوظيفة الجديدة.

"وجدت العديد من المرشحين ليس لديه خدمة أكثر من ثلاث سنوات.

"عط اللوحة على الحائط الخاص.

"ما هو اللون الذي يفضله أكثر؟

"هذا اللون يناسبني أكثر من ذلك، ولا يشكل أي فرق بالنسبة لي.

"تحت تناحي الفرصة لإنقاص سويا سوف نقوم ببعض التجهيزات لإعداد الشواء.

"المقابلة الأسبوعية عادة تقام يوم الجمعة مساءً؛ بشكل مبكر وتعتبر منشطًا ثابتًا لبرنامج الكشافة.

"يعتبر راشد أميركيفيد رفض التناهي الذي قدم له من مناصريه.

"كان أحمد هنا.

"نظرًا للحاجة المستمرة للمنتج، يجب على اللجنة المطالبة به فورًا.

"في وسط البحرية سوف تجد جزيرة صغيرة مغطاة بشجرة واحدة.

"jabedheh
15. رأى أحمد بعض الأسماك تسخج ببطء قريبا منه.
بينما هو يمشي بالقرب من الحديقة. وقف منهضا من الذين وصلوا قبله.

16. الذرة تنمو أطول في فصل الصيف.
زرعت سارة بعض من الأزهار الصفراء في حديقتها الصغيرة.
العاصفة تزداد سواءا كثما كانت الرياح أقوى.

17. هذا قدم بيت في القرية.
إنه أبعد من الفندق الذي رأيناه سابقا ولكن أفضل مثال آثار بني في زمان الأجداد السابقين.

18. ذهب صديقى للمنزل.
خفت المنزل وبالقرب من إلغاء تتمزرة عكبة.

19. هل تلك قاعتك؟
هذا يبدو مناسبًا لك بالرغم من أن تلك المعاطف أقل سعرا من تلك التي في ذلك المعرض.

20. ذهب أحمد لإجراء مكالمة تليفونية.
 تحتاج لـجردلين لـحمل ذلك الصندوق لمسيرية لـإنه تقبل جدا.
السؤال الثالث: التوافق الثنائي

تعليمات: المطلوب منك تذكر المفردات العربية - الكردية في الأسفل. انتظر حتى يطلب منك المشرف على الاختبار البدء. وعندما سوف تعطي دقيقين لدرس المفردات بالأسفل. وعند نهاية الدقيقين سوف يقوم المشرف على الاختبار بإعطائك التعليمات للبدء بحل الأسئلة.

الكردية - العربية

<table>
<thead>
<tr>
<th>الكردية</th>
<th>العربية</th>
</tr>
</thead>
<tbody>
<tr>
<td>hij</td>
<td>kete</td>
</tr>
<tr>
<td>naq</td>
<td>chie</td>
</tr>
<tr>
<td>sidqu</td>
<td>yong</td>
</tr>
<tr>
<td>nente</td>
<td>hui</td>
</tr>
<tr>
<td>a</td>
<td>xozo</td>
</tr>
<tr>
<td>ngoz</td>
<td>mep</td>
</tr>
<tr>
<td>tsep</td>
<td>lah</td>
</tr>
<tr>
<td>johong</td>
<td>wener</td>
</tr>
<tr>
<td>mupa</td>
<td>mi</td>
</tr>
<tr>
<td>nung</td>
<td>jate</td>
</tr>
<tr>
<td>chomco</td>
<td>e</td>
</tr>
<tr>
<td>oor</td>
<td>hon</td>
</tr>
</tbody>
</table>
A. على
B. بارد
T. في
T. يسقط
J. التي

ja.14
A. عم
B. يوم
T. على
T. شمس
J. لا

ngoz.16
A. يدخل
B. سيدة
T. التي
T. ظلام
J. في

nun.18
A. يسأل
B. نسر
T. يوم
T. التي
J. صداع

hij.20
A. الخريف
B. يوم
T. سهل
T. يرسم
J. شمس
J. ح.
أ. طائرة ورقيقة
ب. ذنب
ت. جمل
ث. التي
ج. يلمس

أ. يسل
ب. يرسم
ت. يوم
ث. صينية
ج. يدخل

أ. يمس
ب. ذنب
ت. بعض
ث. يوم
ج. الخريف
Each item below has a group of words. The word at the top of the group is not spelled in
the usual way. Instead, it is spelled approximately as it is pronounced. Your task is to
recognize the disguised word from the spelling. In order to show that you recognize the
disguised word, look for one of the five words beneath it that corresponds most nearly
in meaning with the disguised word. When you find this word or phrase, make a mark in
the appropriate space on your answer sheet.

Here are some sample items:

**S. blocking**

A. a summer fruit

B. the process of controlling movement

C. learning slowly

D. Arabian city

**Blocking** is almost has the same meaning of *the process of controlling movement*, so B
has been marked as the best choice.

START THE QUESTIONS:

1. **hot (a neighbor)**

   a. a chemical substance
   b. a person who lives nearby.
   c. a home furniture
   d. a kind of perfume
2. Cuttlefish (sea animal)
   a. a sea animal
   b. Saudi city
   c. Tiredness and boring
   d. Scented plant

3. Perplexed
   a. Physical disease
   b. Coastal city
   c. gravity
   d. tenser

4. life (guilty people)
   a. death
   b. similar appearance
   c. sinners
   d. store room

5. light (sound of a snake)
   a. psychological disease
   b. opposite of weak
   c. snake's sound
   d. physical property

6. river (to shake)
   a. freezing
   b. to exploit
   c. waves
   d. to move strongly

7. truce (the ceasing of war)
   a. a type of food
   b. sea animal
   c. a tool for decoration
   d. stopping war

8. ice (name of city)
   a. name of a city
   b. a type of perfumes
   c. straightness
   d. environmental adaptation
9. went (gold)
a. an expensive metal
b. economic downturn
c. to consulate
d. Islamic behavior

10. hair (price)
a. home furniture
b. geographic heights
c. the synonym of price
d. sea animal

11. (name of a person) clear
a. highly pure
b. crust's shaking
c. healthy phenomenon
d. an Asian city

12. noise (a blind person)
a. psychological problem
b. a type of metal
c. furniture
d. synonym of blind

13. test (passing)
a. continuously
b. group work
c. accessing borders
d. edification young people

14. redness (a name of male)
a. chemical element
b. the opposite of light
c. name of a lake
d. a male's name

15. war (political party)
a. tiredness
b. type of fruit
c. an expensive metal
d. political group

16. someone (to find)
a. name of female
b. past verb means to find
c. a number
d. to soothe
17. rope (tricks)
a. a lake
b. a plural which means tricks
c. swamp
d. a male's name which means volcanic mountain

18. bridge (a chapter's name in the Quran)
a. chapter in the Quran
b. richness
c. name of a fruit
d. whale

19. worship (a clinic)
a. a gulf city
b. a section in a hospital
c. a kind of seasoning.
d. A noun which means confidence.

20. Mohammad (chilled)
a. Wireless communication
b. scent
c. volcano
d. frozen
WORDS IN SENTENCES

Part 2

This is a test of your ability to understand the function of words and phrases in sentences.

Look at the following example:

**Ayesha is cutting the APPLE with knife.**

w.

**My brother Ahamd is beating the horse with a long stick.**

A B C D E

In the key sentence, **APPLE** is the name of the thing which is being cut; in the second sentence, **horse** is the thing which is being beaten. Therefore, D is the correct answer.

Example 2:

**MONEY is his only object.**

Not so many **years ago, most farming was done by hand.**

A B C D E

The right answer is **farming**: it performs the same function in the second sentence as **MONEY** does in the key sentence. Therefore you should have marked space D.

START THE QUESTIONS

1. Ahmad 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

2. **SEVERAL** were absent from the meeting.

   In spite of the **many** proposals **which** were made, **only one** could be adopted.

   A B C D E
3. Hatim sold **BANDAR** his bicycle.

If their **work** is up to **standard**, I will guarantee **them** a bonus at the **end** of the **week**.

A B C D E

4. My finger became **SWOLLEN** from the infection.

The child grew **strong** from the **healing** sunshine.

A B

The **high** wall was nearly **hidden** from view by the **foliage**.

C D E

5. He drove **FROM** Riyadh to Al.Mdinah.

To be **safe**, he decided to **buy** spare parts for **any** emergency.

A B C D E

6. I will buy a car **WHEN** I get the money.

After you left **last** night, most of the students remained until the **end**.

A B C D E

7. She read the Quran **EXTREME**LY well.

**Promptly** on the dot of five, he came up the stairs, **quite** **flushed** with excitement and breathing **very** **heavily**.

A B C

D E
8. A NUMBER of people applied for the position.
I find many candidates who cannot offer more than two years' experience.

9. He nailed the board TIGHT against the house.
He always did the job well.

10. Which color do YOU like best?
This one suits me better than the other.

11. We plan to study engineering this year
On the chance that he would see us, we took steps to put up a beacon.

12. The weekly meeting, usually held on Friday night, is a fixed ACTIVITY of the Scout program.
Washington was the first president of the United States; he refused the crown that some of

his admirers wanted him to have.
13. AHMAD was here.
There is an urgent need for this product. So, the committee should order it immediately.

A                   B       C     D               E

14. THEY came back from their travel.
In the middle of the lake will be found a small island crowned with a single tree.

A                 B                  C  D                 E

15. Ahamd saw some fish SWIMMING slowly by.
As he was walking down the lane, he found himself wondering who had been there before he arrived.

A               B               C       D

16. The corn grew TALL during the summer.
Sarah raised yellow tulips in her small garden.

A             B           C

The storm proved worse as the wind became stronger.

D               E
17. That is the OLDEST house.
It is farther from your hotel than the one we saw before, but it is the best example of earlier dwellings constructed by our former inhabitants.

18. My FRIEND went home.
Behind the house but near the forest stood a barn.

19. Is THAT your hat?
This looks better on you even though those suits are better bargains than the ones on this

20. Ahmad has gone TO make a telephone call.
Two people are needed to carry this box to the car because it is too heavy for one.
PAIRED ASSOCIATES

Part 3

Instructions: Your task is to MEMORIZE the Kurdish-English vocabulary below. Wait for the signal, then you will be given two minutes to study the vocabulary printed below. At the end of the two minutes the examiner will give you the signal to start answering the questions.

Vocabulary (Memorize for 2 minutes)

<table>
<thead>
<tr>
<th>Kurdish - English</th>
<th>Kurdish - English</th>
</tr>
</thead>
<tbody>
<tr>
<td>hij - draw</td>
<td>kete - camel</td>
</tr>
<tr>
<td>naq - that</td>
<td>chie - few</td>
</tr>
<tr>
<td>sidqu - news</td>
<td>yong - hawk</td>
</tr>
<tr>
<td>nente - lady</td>
<td>hui - fall</td>
</tr>
<tr>
<td>ja - day</td>
<td>xozo - easy</td>
</tr>
<tr>
<td>ngoz - dark</td>
<td>mep - on</td>
</tr>
<tr>
<td>tsep - enter</td>
<td>lah - woH</td>
</tr>
<tr>
<td>lohong - ask</td>
<td>wener - book</td>
</tr>
<tr>
<td>mupa - anger</td>
<td>mi - touch</td>
</tr>
<tr>
<td>nung - frog</td>
<td>jate - sun</td>
</tr>
<tr>
<td>chomco - body</td>
<td>e - bowl</td>
</tr>
<tr>
<td>roo - art</td>
<td>hon - cold</td>
</tr>
</tbody>
</table>

Instructions: In the next part of the test, you will be given the Kurdish words and 5 choices in Arabic.
PAIRED ASSOCIATES (CONTINUED)

1. mep
   A. in
   B. on
   C. that
   D. enter
   E. art

2. e
   A. ball
   B. at
   C. body
   D. cold
   E. bowl

3. lah
   A. wolf
   B. camel
   C. dark
   D. last
   E. on

4. tesp
   A. ask
   B. anger
   C. enter
   D. touch
   E. draw

5. jate
   A. frog
   B. body
   C. hawk
   D. sun
   E. book

6. xozo
   A. day
   B. easy
   C. news
   D. touch
   E. bowl
<table>
<thead>
<tr>
<th></th>
<th>chomco</th>
<th></th>
<th>nente</th>
<th></th>
<th>mupa</th>
<th></th>
<th>yong</th>
<th></th>
<th>chie</th>
<th></th>
<th>hon</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>A. art</td>
<td></td>
<td>A. enter</td>
<td></td>
<td>A. map</td>
<td></td>
<td>A. easy</td>
<td></td>
<td>A. cold</td>
<td></td>
<td>A. on</td>
<td></td>
</tr>
<tr>
<td></td>
<td>B. body</td>
<td></td>
<td>B. anger</td>
<td></td>
<td>B. draw</td>
<td></td>
<td>B. young</td>
<td></td>
<td>B. camel</td>
<td></td>
<td>B. cold</td>
<td></td>
</tr>
<tr>
<td></td>
<td>C. cold</td>
<td></td>
<td>C. lady</td>
<td></td>
<td>C. cold</td>
<td></td>
<td>C. few</td>
<td></td>
<td>C. bowl</td>
<td></td>
<td>C. in</td>
<td></td>
</tr>
<tr>
<td></td>
<td>D. news</td>
<td></td>
<td>D. few</td>
<td></td>
<td>D. wolf</td>
<td></td>
<td>D. touch</td>
<td></td>
<td>D. few</td>
<td></td>
<td>D. fall</td>
<td></td>
</tr>
<tr>
<td></td>
<td>E. dark</td>
<td></td>
<td>E. fall</td>
<td></td>
<td>E. anger</td>
<td></td>
<td>E. hawk</td>
<td></td>
<td>E. Frog</td>
<td></td>
<td>E. that</td>
<td></td>
</tr>
</tbody>
</table>
13. naq
   A. not
   B. day
   C. that
   D. art
   E. ask

14. ja
   A. yes
   B. day
   C. on
   D. sun
   E. no

15. roo
   A. art
   B. draw
   C. ask
   D. run
   E. camel

16. ngoz
   A. enter
   B. lady
   C. that
   D. dark
   E. on

17. wener
   A. never
   B. book
   C. anger
   D. few
   E. touch

18. nung
   A. ask
   B. hawk
   C. day
   D. that
   E. frog

19. lohong
   A. frog
   B. wolf
   C. body
   D. ask
   E. cold
20. hij
   A. fall
   B. day
   C. easy
   D. draw
   E. sun

21. kete
   A. kite
   B. wolf
   C. camel
   D. that
   E. touch

22. sidqu
   A. easy
   B. cold
   C. news
   D. dark
   E. book

23. hui
   A. ask
   B. wolf
   C. few
   D. day
   E. fall

24. mi
   A. touch
   B. draw
   C. day
   D. bowl
   E. enter
Appendix 4: The English proficiency test (adapted TOEFL)

SECTION 1
LISTENING COMPREHENSION
Time: 10 minutes

In this section of the test, you will have an opportunity to demonstrate your ability to understand conversations and tasks in English. Answer all the questions on the basis of what is stated or implied by the speakers you hear.

Do not take notes or write in your test book at any time. Do not turn the pages until you are told to do so.

Directions

You will hear short conversations between two people. After each conversation, you will hear a question about the conversation. The conversations and the questions will not be repeated. After you hear a question, read the four possible answers in your test book and choose the best answer. Then in your answer sheet, find the number of the question and fill in the space that corresponds to the letter of the answer you have chosen.

Listen to an example.

On the recording you hear:

(Man)        That exam was just awful.
(Woman)   Oh, it could have been worse.
(Man)        What does the woman mean?

In your test book, you read:

A. The exam was really awful
B. It was the worse exam she had ever seen.
C. It couldn’t have been more difficult
D. It wasn’t that hard.

You learn from the conversation that the man thought the exam was very difficult and that the woman disagreed with the man. The best answer to the question, “What does the woman mean?” is (D), “It was not that hard.” Therefore, the correct choice is (D)
1. (A) Carla does not live very far away.
   (B) What does Carla said was unjust.
   (C) He does not fear what anyone says.
   (D) Carla is fairly rude to the others.

2. (A) She thinks it is an improvement.
   (B) The fire tress in it, are better.
   (C) It resembles the last one.
   (D) It is the best the man has ever done.

3. (A) He graduated last in the class.
   (B) He is the last person in his family to graduate.
   (C) He doesn’t believe he can improve gradually.
   (D) He has finally finished his studies.

4. (A) He’s surprised there were five dresses.
   (B) It was an expectedly inexpensive dress.
   (C) He would like to know what color was it.
   (D) The dress want not that cheap.

5. (A) Leave the car somewhere else.
   (B) Ignore the parking tickets.
   (C) Add more money to the meter.
   (D) Pay the parking tickets attendant.

6. (A) He does not like to hold too many books at one time.
   (B) There is no bookstore in his neighbour-hood.
   (C) It is not possible to obtain the book yet.
   (D) He needs to talk to someone at the bookstore.

7. (A) it was incomplete.
   (B) It finished on time.
   (C) It was about honor.
   (D) It was too long.

8. (A) She needs to use the man’s notes.
   (B) Yesterday’s physics class was quiet boring.
   (C) She took some very good notes in physics class.
   (D) She would like to lend the man her notes.
9. (A) It’s her birthday today.
   (B) She is looking for a birthday gift.
   (C) She wants to go shopping with her dad.
   (D) She wants a new wallet for herself.

10. (A) He prefers cold water.
    (B) His toes are too big.
    (C) The pool left quiet refreshing.
    (D) He didn’t go for a swim.

11. (A) She just left her sister’s house.
    (B) Her sister is not at home.
    (C) She’s not exactly sure where her sweater is.
    (D) She does not know here her sister lives.

12. (A) She does not have time to complete additional reports.
    (B) She cannot finish the reports that she is already working on.
    (C) She is scared of having responsibility for the reports.
    (D) It is not time for accounting reports to be completed.

13. (A) He’s had enough exercise.
    (B) He’s going to give himself a reward for the hard work.
    (C) He is going to stay on for quite some time.
    (D) He would like to give the woman an exercise machine as a gift.

14. (A) He cannot see the huge waves.
    (B) The waves are not coming in.
    (C) He would like the woman to repeat what she said.
    (D) He agrees with the woman.

15. (A) The exam was postponed.
    (B) The man should have studies harder.
    (C) Night is the best time to study for exams.
    (D) She is completely prepared for the exam.

16. (A) Students who want to change schedules should form a line.
    (B) It is only possible to make four changes in the schedule.
    (C) It is necessary to change to submit the form quickly.
    (D) Problems occur when people don’t wait their turn.
17.  
(A) In a mine.  
(B) In a jewellery store.  
(C) In a clothing store.  
(D) In a bank.  

18.  
(A) A visit to the woman’s family  
(B) The telephone bill  
(C) The cost of a new telephone.  
(D) How far way the woman’s family lives.  

19.  
(A) She has not met her new boss yet.  
(B) She has a good opinion of her boss.  
(C) Her boss has asked her about her impressions of the company.  
(D) Her boss has been putting a lot of pressure on her.  

20.  
(A) The recital starts in three hours.  
(B) He intends to recite three different poems.  
(C) He received a citation on the third of the month.  
(D) He thinks the performance begins at three.
This section is designed to measure your ability to recognize language that appropriate for standard written English. There are two types of questions in this section, with special directions for each type.

**Structure**
Directions: Question 1-10 are incomplete sentences. Beneath each sentence you will see four words or phrases, marked (A), (B), (C), (D). Choose the one word or phrase that best completes the sentence. Then, on your answer sheet, find the number of the question and fill in the space that corresponds to the letter of the answer you have chosen. Fill in the space so that the letter inside the oval cannot be seen.

Example:
The president __________ the election by a landslide.

(A) won
(B) He won
(C) yesterday
(D) Fortunately

The sentence should read, “The president **won** the election by a landslide.” Therefore, you should choose (A).

1. ……………………., the outermost layer of the skin, is about as thick as a sheet of paper over most of the skin.

   A. it is the epidermis
   B. in the epidermis
   C. the epidermis
   D. the epidermis is

2. Sam Spade in *The Maltese Falcon* and Rick Blaine in *Casablanca* ……….. of Humphrey Bogart’s more famous roles.

   A. They are two
   B. Two of them are
   C. Two of them
   D. Are two

215
3. The compound microscope has not one…………… two lenses.

A. And also
B. But
C. And there are
D. But there are

4. During the Precambrian period, the Earth’s crust formed, and life …….. in the seas.

A. First appeared
B. First to appear
C. Is first appearing
D. Appearing

5. The hard palate forms a partition …………… and nasal passages.

A. The mouth
B. Between the mouth
C. Is between the mouth
D. It is between the mouth

6. Conditions required for seed germination include abundant water, an adequate supply of oxygen, and …………..

A. The temperatures must be appropriate
B. Having appropriate temperatures
C. Appropriate temperatures
D. Appropriately temperate

7. When fluid accumulates against the eardrum, a second more insidious type of …………..

A. *otitis media* may develop
B. developing *otitis media*
C. the development of *otitis media*
D. to develop *otitis media*

8. Some general theories of motivation ……… of central motives, from with other motives develop.

A. Identify a limited number
B. Identification of a limited amount
C. Identify a limited amount
D. Identifying a limited number.
9. Before the Statue of Liberty arrived in the United States, newspaper invited the public to help determine where …….. placed after its arrival.

   A. Should the status be
   B. The statue being
   C. It should be in the status
   D. The status should be

10. Hydroelectric power can be produced by ………. And using tidal flow to run turbines.

   A. Water basins are dammed
   B. Damming water basins
   C. To dam water basins
   D. Dams in water basins.

Written Expression

Directions, in questions 11-20, each sentence has four underlined words or phrases. The four underlined parts of the sentence are marker (A),(B),(C) and (D). Identify the one underlined word or phrase that must be changed in order for the sentence to be correct. Then in your answer sheet, find the number of the question and fill in the space that corresponds to the letter of the answer you have chosen.

Example:

The research or the book Roots taking Alex Haley twelve years.

A     B     C     D

The sentence should read, “The research for the book took Alex Haley twelve years.” Therefore you should choose (C).

11. Mosquitoes will accepts the malaria parasite at only one stage of the parasite complex life cycle.

A     B     C     D

12. The counterpart of negative electrons is the positive proton.

A     B     C     D
13. The ankle joint occur where the lower ends of the tibia and fibula slot neatly around the talus.
A                      B                                                          C
D

14. The neocortex is, in evolutionary terms, most recent layer of the brain.
A                             B        C                  D

15. Supersonic flight is flight that is faster the speed of sound.
A                       B        C                  D

16. A zoom lens produces an inverted real image, either on the film in a camera and on the light-sensitive tube of a television camera.
A                        B                  C
D

17. It is a common observation that liquids will soak through some materials but not through other.
A                                                     B                                              C
D

18. Unlikely gas sport balloons, hot air balloons do not have nets.
A                 B                                                         C     D

A        B                                   C
D

20. In space, with no gravity for muscles to work against, the body becomes weakly.
A                                                  B             C
D
SECTION 3
READING COMPREHENSION
Time: 20 minutes

This section is designed to measure your ability to read and understand short passages similar in topic and style to those that students are likely to encounter in universities and colleges.

Directions: in this section, you will read several passages. Each one is followed by a number of questions about it. You are to choose the one best answer, (A), (B), (C), or (D) to each question. Then, on you answer sheet, find the number of the question and fill in the space that corresponds to the letter for the answer you have chosen.

Answer all the questions about the information in a passage on the basis of what is stated and implied in that passage.

Now begin work in the passages.

Question 1-5

The hippopotamus is the third largest land animal, smaller only than the elephant and the rhinoceros. Its name comes from two Greek words which mean “river horse” the long name of this animal is often shortened to the easier to handle term” hippo”

The hippo has a natural affinity for the water. It does not float on top of the water; instead it can easily walk along the bottom of a body of water. The hippo commonly remains underwater for three to five minutes and has boon known to stay under for up to half an hour before coming up for air.

In spite of its name, the hippo has relatively little in common with the horse and instead has a number of interesting similarities in common with the whale. When a hippo comes up after stay at the bottom of a lake or river, it releases air through a blowhole, just like a whale. In addition, the hippo resembles the whale in that they both have thick layers of blubber for protection and they are almost completely hairless.

1. The topic of this passage is
   A. the largest animal land
   B. the derivations of animal name
   C. the characteristics of the hippo
   D. the relation between the hippo and the whale

2. The word “float” in line 4 is closest on meaning to
   A. sink
   B. drift
   C. eat
   D. flap
3. According to the passage, what is the maximum time that hippos have been known to stay underwater?

A. three minutes  
B. five minutes  
C. thirty minutes  
D. ninety minutes

4. The passages state that one way in which a hippo is similar to a whale is that

A. the both live on the bottom of rivers
B. they both have blowholes
C. they are both names after horses
D. they both breathe underwater

5. The word “bubbler” in line 11 is closest in meaning to

A. fat  
B. metal  
C. water  
D. skin

**QUESTION 6-10**

People are often surprised to learn how long some varieties of trees can live. If asked to estimate the age of the oldest living trees on Earth, they often come up with guesses in the neighbourhood of two or perhaps three hundred years. The real answer is considerably larger than that, more than five thousand years.

The tree that wins the prize for its considerable maturity is the bristlecone pine of California. This venerable pine predates wonders of the ancient world such as pyramids of Egypt, the Hanging Gardens of Babylon, and the Colossus of Rhodes. It can not nearly as tall as the giant redwood that is found in California, and, in fact it is actually not very tall compared with many other trees, often little more than five meters in height. This relatively short height may be one of the factors that aid the bristlecone pine in living to a ripe old age – high wind and inclement weather cannot easily reach the shorter trees and cause damage. An additional factor that contributes to the long life of the bristlecone pine is that this type of tree has a high percentage of resin, which prevents rot from developing in the tree trunk and branches.
6. The best title of the passage would be

A. The Size of the Bristlecone Pine
B. Three-Hundred-Year-Old Forests
C. The Wonders of the Ancient World
D. An Amazingly Enduring Tree

7. The word “estimate” in line 2 is closest in meaning to

A. measure
B. approximate
C. evaluate
D. view

8. According to the passage, approximately how old are the oldest trees on the Earth?

A. Two hundred years old
B. Three hundred years old
C. Five hundred years old
D. Fine thousand years old

9. The author mentioned the Egyptian pyramid is as an example of something that is

A. far away
B. believed to be strong
C. extremely tall
D. known to be old

10. The word “inclement” in line 10 could best be replaced by

A. sunny
B. bad
C. unusual
D. strong

END OF THE TEST
Appendix 5: Attrition data analyse

<table>
<thead>
<tr>
<th>Between-Subjects Factors</th>
<th>Value</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participants who failed, dropped the course or absent</td>
<td>0</td>
<td>failed 19</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>absent 6</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>retained 56</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>dropped 9</td>
</tr>
</tbody>
</table>

### Descriptive Statistics

<table>
<thead>
<tr>
<th></th>
<th>Participants who failed, dropped the course or absent</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Spelling Clues</strong></td>
<td>failed</td>
<td>12.5263</td>
<td>3.30603</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>absent</td>
<td>11.0000</td>
<td>2.96648</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>retained</td>
<td>12.2857</td>
<td>2.86492</td>
<td>56</td>
</tr>
<tr>
<td></td>
<td>dropped</td>
<td>13.7778</td>
<td>2.63523</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>12.4000</td>
<td>2.95585</td>
<td>90</td>
</tr>
<tr>
<td><strong>Grammar Sensitivity</strong></td>
<td>failed</td>
<td>7.1053</td>
<td>1.72867</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>absent</td>
<td>6.3333</td>
<td>2.25093</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>retained</td>
<td>6.9107</td>
<td>2.48835</td>
<td>56</td>
</tr>
<tr>
<td></td>
<td>dropped</td>
<td>7.5556</td>
<td>3.67801</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>6.9778</td>
<td>2.44939</td>
<td>90</td>
</tr>
<tr>
<td><strong>Memory</strong></td>
<td>failed</td>
<td>14.0526</td>
<td>5.14753</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>absent</td>
<td>15.1667</td>
<td>3.18852</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>retained</td>
<td>14.8214</td>
<td>4.60082</td>
<td>56</td>
</tr>
<tr>
<td></td>
<td>dropped</td>
<td>19.2222</td>
<td>2.81859</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>15.1222</td>
<td>4.66117</td>
<td>90</td>
</tr>
<tr>
<td></td>
<td>Participants who failed, dropped the course or absent</td>
<td>Mean</td>
<td>Std. Deviation</td>
<td>N</td>
</tr>
<tr>
<td>--------------------------</td>
<td>------------------------------------------------------</td>
<td>--------</td>
<td>----------------</td>
<td>----</td>
</tr>
<tr>
<td><strong>Aptitude Total</strong></td>
<td>failed</td>
<td>33.6842</td>
<td>5.99122</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>absent</td>
<td>32.5000</td>
<td>6.80441</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>retained</td>
<td>34.0179</td>
<td>7.10055</td>
<td>56</td>
</tr>
<tr>
<td></td>
<td>dropped</td>
<td>40.5556</td>
<td>6.22718</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>34.5000</td>
<td>6.98353</td>
<td>90</td>
</tr>
<tr>
<td><strong>Listening</strong></td>
<td>failed</td>
<td>6.7895</td>
<td>1.98827</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>absent</td>
<td>7.3333</td>
<td>1.50555</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>retained</td>
<td>8.6964</td>
<td>3.06843</td>
<td>56</td>
</tr>
<tr>
<td></td>
<td>dropped</td>
<td>8.0000</td>
<td>3.31662</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>8.1333</td>
<td>2.89207</td>
<td>90</td>
</tr>
<tr>
<td><strong>Grammar</strong></td>
<td>failed</td>
<td>5.8947</td>
<td>2.49209</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>absent</td>
<td>6.6667</td>
<td>2.80476</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>retained</td>
<td>6.8036</td>
<td>3.01764</td>
<td>56</td>
</tr>
<tr>
<td></td>
<td>dropped</td>
<td>5.6667</td>
<td>2.78388</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>6.4889</td>
<td>2.86883</td>
<td>90</td>
</tr>
<tr>
<td><strong>Reading</strong></td>
<td>failed</td>
<td>3.4211</td>
<td>1.21636</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>absent</td>
<td>3.0000</td>
<td>0.63246</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>retained</td>
<td>3.0000</td>
<td>1.47710</td>
<td>56</td>
</tr>
<tr>
<td></td>
<td>dropped</td>
<td>3.1111</td>
<td>1.26930</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>3.1000</td>
<td>1.35787</td>
<td>90</td>
</tr>
<tr>
<td><strong>English Prof. Total</strong></td>
<td>failed</td>
<td>16.1053</td>
<td>3.54173</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>absent</td>
<td>17.0000</td>
<td>3.79473</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>retained</td>
<td>18.5000</td>
<td>5.65042</td>
<td>56</td>
</tr>
<tr>
<td></td>
<td>dropped</td>
<td>16.7778</td>
<td>2.77389</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>17.7222</td>
<td>4.98318</td>
<td>90</td>
</tr>
</tbody>
</table>
### Multivariate Tests

<table>
<thead>
<tr>
<th>Effect</th>
<th>Pillai's Trace</th>
<th>Wilks' Lambda</th>
<th>Hotelling's Trace</th>
<th>Roy's Largest Root</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>.000</td>
<td>.953</td>
<td>.000</td>
<td>.953</td>
</tr>
<tr>
<td>Partstatus 1</td>
<td>.148</td>
<td>.090</td>
<td>.152</td>
<td>.091</td>
</tr>
<tr>
<td></td>
<td>.156</td>
<td>.093</td>
<td>.048</td>
<td>.139</td>
</tr>
</tbody>
</table>

### Levene's Test of Equality of Error Variances

<table>
<thead>
<tr>
<th></th>
<th>F</th>
<th>df1</th>
<th>df2</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spelling Clues</td>
<td>.493</td>
<td>3</td>
<td>86</td>
<td>.688</td>
</tr>
<tr>
<td>Grammar Sensitivity</td>
<td>2.940</td>
<td>3</td>
<td>86</td>
<td>.038</td>
</tr>
<tr>
<td>Memory</td>
<td>2.216</td>
<td>3</td>
<td>86</td>
<td>.092</td>
</tr>
<tr>
<td>Aptitude Total</td>
<td>.868</td>
<td>3</td>
<td>86</td>
<td>.461</td>
</tr>
<tr>
<td>Listening</td>
<td>3.576</td>
<td>3</td>
<td>86</td>
<td>.017</td>
</tr>
<tr>
<td>Grammar</td>
<td>.645</td>
<td>3</td>
<td>86</td>
<td>.588</td>
</tr>
<tr>
<td>Reading</td>
<td>1.226</td>
<td>3</td>
<td>86</td>
<td>.305</td>
</tr>
<tr>
<td>English Prof. Total</td>
<td>2.607</td>
<td>3</td>
<td>86</td>
<td>.057</td>
</tr>
</tbody>
</table>
## Tests of Between-Subjects Effects

<table>
<thead>
<tr>
<th>Source</th>
<th>Dependent Variable</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
<th>Partial Eta Square d</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corrected Model</td>
<td>Spelling Clues</td>
<td>29.879a</td>
<td>3</td>
<td>9.960</td>
<td>1.146</td>
<td>.335</td>
<td>.038</td>
</tr>
<tr>
<td></td>
<td>Grammar Sensitivity</td>
<td>6.057b</td>
<td>3</td>
<td>2.019</td>
<td>.329</td>
<td>.804</td>
<td>.011</td>
</tr>
<tr>
<td></td>
<td>Memory</td>
<td>178.105c</td>
<td>3</td>
<td>59.368</td>
<td>2.908</td>
<td>.039</td>
<td>.092</td>
</tr>
<tr>
<td></td>
<td>Aptitude Total</td>
<td>379.690d</td>
<td>3</td>
<td>126.563</td>
<td>2.748</td>
<td>.048</td>
<td>.087</td>
</tr>
<tr>
<td></td>
<td>Listening</td>
<td>56.069a</td>
<td>3</td>
<td>18.690</td>
<td>2.335</td>
<td>.079</td>
<td>.075</td>
</tr>
<tr>
<td></td>
<td>Grammar</td>
<td>18.527f</td>
<td>3</td>
<td>6.176</td>
<td>.744</td>
<td>.529</td>
<td>.025</td>
</tr>
<tr>
<td></td>
<td>Reading</td>
<td>2.580g</td>
<td>3</td>
<td>.860</td>
<td>.458</td>
<td>.712</td>
<td>.016</td>
</tr>
<tr>
<td></td>
<td>English Prof. Total</td>
<td>94.711h</td>
<td>3</td>
<td>31.570</td>
<td>1.283</td>
<td>.285</td>
<td>.043</td>
</tr>
<tr>
<td>Intercept</td>
<td>Spelling Clues</td>
<td>7061.113</td>
<td>1</td>
<td>7061.113</td>
<td>812.142</td>
<td>.000</td>
<td>.904</td>
</tr>
<tr>
<td></td>
<td>Grammar Sensitivity</td>
<td>2235.878</td>
<td>1</td>
<td>2235.878</td>
<td>364.247</td>
<td>.000</td>
<td>.809</td>
</tr>
<tr>
<td></td>
<td>Memory</td>
<td>11491.776</td>
<td>1</td>
<td>11491.776</td>
<td>562.953</td>
<td>.000</td>
<td>.867</td>
</tr>
<tr>
<td></td>
<td>Aptitude Total</td>
<td>56889.504</td>
<td>1</td>
<td>56889.504</td>
<td>1235.227</td>
<td>.000</td>
<td>.935</td>
</tr>
<tr>
<td></td>
<td>Listening</td>
<td>2727.294</td>
<td>1</td>
<td>2727.294</td>
<td>340.748</td>
<td>.000</td>
<td>.798</td>
</tr>
<tr>
<td></td>
<td>Grammar</td>
<td>1799.148</td>
<td>1</td>
<td>1799.148</td>
<td>216.716</td>
<td>.000</td>
<td>.716</td>
</tr>
<tr>
<td></td>
<td>Reading</td>
<td>450.962</td>
<td>1</td>
<td>450.962</td>
<td>240.111</td>
<td>.000</td>
<td>.736</td>
</tr>
<tr>
<td></td>
<td>English Prof. Total</td>
<td>13427.190</td>
<td>1</td>
<td>13427.190</td>
<td>545.887</td>
<td>.000</td>
<td>.864</td>
</tr>
<tr>
<td>PartstatusT1</td>
<td>Spelling Clues</td>
<td>29.879</td>
<td>3</td>
<td>9.960</td>
<td>1.146</td>
<td>.335</td>
<td>.038</td>
</tr>
<tr>
<td></td>
<td>Grammar Sensitivity</td>
<td>6.057</td>
<td>3</td>
<td>2.019</td>
<td>.329</td>
<td>.804</td>
<td>.011</td>
</tr>
<tr>
<td></td>
<td>Memory</td>
<td>178.105</td>
<td>3</td>
<td>59.368</td>
<td>2.908</td>
<td>.039</td>
<td>.092</td>
</tr>
<tr>
<td></td>
<td>Aptitude Total</td>
<td>379.690</td>
<td>3</td>
<td>126.563</td>
<td>2.748</td>
<td>.048</td>
<td>.087</td>
</tr>
<tr>
<td></td>
<td>Listening</td>
<td>56.069</td>
<td>3</td>
<td>18.690</td>
<td>2.335</td>
<td>.079</td>
<td>.075</td>
</tr>
<tr>
<td></td>
<td>Grammar</td>
<td>18.527</td>
<td>3</td>
<td>6.176</td>
<td>.744</td>
<td>.529</td>
<td>.025</td>
</tr>
<tr>
<td></td>
<td>Reading</td>
<td>2.580</td>
<td>3</td>
<td>.860</td>
<td>.458</td>
<td>.712</td>
<td>.016</td>
</tr>
<tr>
<td></td>
<td>English Prof. Total</td>
<td>94.711</td>
<td>3</td>
<td>31.570</td>
<td>1.283</td>
<td>.285</td>
<td>.043</td>
</tr>
<tr>
<td>Error</td>
<td>Spelling Clues</td>
<td>747.721</td>
<td>86</td>
<td>8.694</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Grammar</td>
<td>527.899</td>
<td>86</td>
<td>6.138</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sensitivity</td>
<td>Memory</td>
<td>Listening</td>
<td>Grammar</td>
<td>Reading</td>
<td>English Prof. Total</td>
<td></td>
</tr>
<tr>
<td>------------------</td>
<td>-------------</td>
<td>--------</td>
<td>-----------</td>
<td>---------</td>
<td>---------</td>
<td>--------------------</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1755.551</td>
<td>86</td>
<td>688.331</td>
<td>713.962</td>
<td>161.520</td>
<td>2115.345</td>
<td></td>
</tr>
<tr>
<td>Aptitude Total</td>
<td>3960.810</td>
<td>86</td>
<td>86</td>
<td>86</td>
<td>86</td>
<td>111463.000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>6698.000</td>
<td>68</td>
<td>744.400</td>
<td>732.489</td>
<td>164.100</td>
<td>2115.345</td>
<td></td>
</tr>
<tr>
<td>English Prof. Total</td>
<td>2210.056</td>
<td>86</td>
<td>86</td>
<td>86</td>
<td>86</td>
<td>4340.500</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Sensitivity</th>
<th>Memory</th>
<th>Listening</th>
<th>Grammar</th>
<th>Reading</th>
<th>English Prof. Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>4916.000</td>
<td>90</td>
<td>4522.000</td>
<td>4522.000</td>
<td>1029.000</td>
<td>30477.000</td>
</tr>
<tr>
<td>Grammar</td>
<td>14616.000</td>
<td>90</td>
<td>6698.000</td>
<td>6698.000</td>
<td>1029.000</td>
<td>30477.000</td>
</tr>
<tr>
<td>Total</td>
<td>22515.000</td>
<td>90</td>
<td>111463.000</td>
<td>111463.000</td>
<td>1029.000</td>
<td>30477.000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Sensitivity</th>
<th>Memory</th>
<th>Listening</th>
<th>Grammar</th>
<th>Reading</th>
<th>English Prof. Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>777.600</td>
<td>89</td>
<td>744.400</td>
<td>744.400</td>
<td>164.100</td>
<td>2210.056</td>
</tr>
<tr>
<td>Grammar</td>
<td>533.956</td>
<td>89</td>
<td>732.489</td>
<td>732.489</td>
<td>164.100</td>
<td>2210.056</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>1933.656</td>
<td>89</td>
<td>744.400</td>
<td>744.400</td>
<td>164.100</td>
<td>2210.056</td>
</tr>
<tr>
<td>Total</td>
<td>22515.000</td>
<td>90</td>
<td>111463.000</td>
<td>111463.000</td>
<td>1029.000</td>
<td>30477.000</td>
</tr>
</tbody>
</table>