Chapter 13
Foreign Language Aptitude Components and Different Levels of Foreign Language Proficiency Among Chinese English Majors

Lanrong Li

Abstract This study aims to explore the relationship between language aptitude components and different levels of English proficiency among Chinese English majors. Sixty-four second-year English majors from a university in Beijing participated in the study. An aptitude test composed of three subtests of Pimsleur Language Aptitude Battery and two self-developed subtests was administered to the participants. The students' scores on two national English proficiency tests (Test for English majors – Band 4 and Band 8, abbreviated as TEM-4 and TEM-8) were used as measures of their English proficiency. Correlational analysis, multiple regression analysis and *t*-tests were conducted. Results showed that different aptitude components had different relationship with different levels of language proficiency. Regression analysis showed that two aptitude components (sound discrimination and memory for text) were significant predictors of both of the students' TEM-4 and TEM-8 scores, while only inductive language learning ability could significantly predict their TEM-8 scores. Further analysis showed that students with higher and lower TEM-4 and TEM-8 scores also differed significantly in different language aptitude components. The results lend support to Skehan's hypothesis (A cognitive approach to language learning, Oxford University Press, Oxford, 1998) that different language aptitude components play different roles in second language acquisition when the learner is at different levels of proficiency.

Keywords Foreign language aptitude components • Memory for text • Language analysis • Sound discrimination • Foreign language proficiency

L. Li (✉)
English Department, Beijing Normal University, Beijing, China

Q. Zhang and H. Yang (eds.), *Pacific Rim Objective Measurement Symposium (PROMS) 2012 Conference Proceeding*, DOI 10.1007/978-3-642-37592-7_13,

13.1 Introduction

Foreign Language (FL) aptitude is widely considered as one of the central individual variables in second language (L2) learning (Skehan 1989). Abundant research has found high correlations between foreign language aptitude and various forms of L2 attainment, and the coefficients were usually around 0.4–0.6 (see reviews by J. B. Carroll 1981; Sawyer and Ranta 2001; Dornyei and Skehan 2003). It is believed that foreign language aptitude may account for the difficulties of unsuccessful language learners and the advantages of successful language learners. However, according to J. B. Carroll (1962, 1981, 1990), FL aptitude is not a unitary concept, but is composed of several components, which are phonetic coding ability, grammatical sensitivity, inductive language learning ability, and associative memory.

The componential view of aptitude suggests that learners may draw upon different aptitude components at different developmental stages of L2 proficiency. There is also some evidence (Wesche 1981; Skehan 1986) showing that learners have their strengths and weaknesses in their FL aptitude profiles, and different aptitude components may account for the failure and success of FL learning (e.g. Pimsleur et al. 1963; Ioup et al. 1994; Dekeyser 2000; Rysiewicz 2008). Based on a review of studies on successful and unsuccessful FL learners from the perspective of language aptitude, Skehan (1998) hypothesized that aptitude components play different roles at different proficiency levels and proposed a diagram showing explicitly the relationship between them. His hypotheses make it possible to adopt different instructional methods to facilitate L2 learning at different developmental stages and identify those L2 learners with learning difficulties or stronger potentials.

However, not many studies have attempted to test Skehan's hypotheses. The small number of studies which touched on this issue also generated inconsistent findings (Ma and Wang 2011; Winke 2005; Hummel 2009), especially on the role of grammatical sensitivity and memory components. Due to lack of sufficient empirical evidence, we know even less about the role of inductive language learning ability in L2 development. Thus, the aim of this present study is to test Skehan's hypotheses on the relationship between aptitude components and different levels of L2 proficiency. This article presents a study in which FL proficiency at different developmental stages and FL aptitude components formed the two main variables. In the first section, background about FL aptitude as well as its measurement tools is provided, followed by a literature review on relevant studies on aptitude components and FL proficiency at different levels. The last two sections describe the study and report the major findings.

13.2 Foreign Language Aptitude and Its Measurement Tools

Foreign language aptitude refers to "the individual's initial state of readiness and capacity for learning a foreign language, and probable degree of facility in doing so" given the presence of motivation and opportunity (J. B. Carroll 1981, p. 86). It is

Table 13.1 Structure of the MLAT and the PLAB

Aptitude test	Subtest	Components measured
MLAT	I Number learning	Memory & phonetic coding ability
	II Phonetic script	Phonetic coding ability
	III Spelling clues	Phonetic coding ability & vocabulary
	IV Words in sentences	Grammatical sensitivity
	V Paired associates	Associative memory
PLAB	I Grades in major subjects	
	II Interest	Motivation
	III Vocabulary	Knowledge of vocabulary
	IV Language analysis	Inductive language learning
	V Sound discrimination	Auditory ability
	IV Sound-symbol association	

regarded as a cognitively based learner characteristic that controls the rate of progress the learner will make in foreign language learning. Based on his early research on language aptitude, J. B. Carroll (1981) proposed the following four aptitude components:

1. phonetic coding ability – sound-symbol association ability. An ability to identify distinct sounds, to form associations between those sounds and the symbols representing them, and to retain these associations.
2. grammatical sensitivity – the ability to recognize the grammatical functions of words (or other linguistic entities) in sentence structures;
3. rote learning ability for foreign language materials – the ability to learn associations between sounds and meanings rapidly an deficiently, and to retain these associations; and
4. inductive language learning ability – the ability to infer or induce the rules governing a set of language materials, given samples of language materials that permit such inferences. It is the ability to extract syntactic and morphological patterns from a given corpus of language material and to extrapolate from such pattern to create new sentences (Carroll 1981, p. 105).

The MLAT (Modern Language Aptitude Test), developed by J. B. Carroll and Sapon (1959/2002), was intended to test these aptitude components except inductive language learning ability, which was only weakly and indirectly measured (J. B. Carroll 1981). However, this aptitude component was well represented by the third subtest of PLAB (*Pimsleur Language Aptitude*, Pimsleur et al. 2004). However, PLAB does not have a subtest for testing memory ability, and different from the MLAT, it also takes motivation and average grade points of other subjects as part of aptitude. Besides, the MLAT was designed for adult literate native-speakers of English, while PLAB was designed for native-English speakers aged between 13 and 19. Their structures are presented in Table 13.1

As can be found, there is no one-to-one correspondence between subtests of the MLAT and PLAB and the aptitude components they are intended to measure. Another thing worth noting is that the last two subtests of PLAB measure auditory

ability according to the PLAB manual (Pimsleur et al. 2004). In J. B. Carroll's (1981) model, these two subtests actually tap into phonetic coding ability. However, PLAT-V Sound Discrimination may involve more of the ability to associate sound patterns with their meaning than the ability to distinguish between different sounds (J. B. Carroll 1981; Skehan 1989), as J. B. Carroll (1962) also found that the test of sound discrimination was not predictive of L2 success. In this sense, PLAT-V Sound Discrimination may actually measures two subcomponents: the ability to identify sounds and memory and it may be the latter that gives the validity of this subtest (J. B. Carroll 1981).

Different as the two aptitude tests are, they are complementary rather than contradictory. Since their publication, the two language aptitude tests have been widely used among researchers and have "proved to be strong predictors of language learning success" (Kiss and Nilolov 2005, p. 107). Despite their high predictive validity, some researchers (J. B. Carroll 1990; Robinson 2005) also point out that traditional language aptitude tests may not be able to predict L2 success at the advanced level. Other components like working memory have been proposed as potential aptitude components that predict the success at the advanced level of L2 proficiency (e.g. Miyake and Friedman 1998; Sawyer and Ranta 2001; Robinson 2002, 2005; Dornyei 2005; Erlam 2005). Although the present study did not take working memory as a variable, it is still necessary to provide sufficient evidence to show the limitations of memory ability measured by traditional aptitude test.

13.3 FL Aptitude Components and L2 Learning at Different Levels of Proficiency

By reviewing relevant research on successful and unsuccessful language learners, Skehan (1998) puts forth a diagram showing the relationship between relative importance of the different language aptitude components and different levels of FL language proficiency (as shown in Fig. 13.1).

As can be seen, at the early developmental stage of L2 proficiency, phonemic coding ability is of great importance. However, subsequently, it reaches a plateau, indicating that it only contributes to foreign language learning at the initial stage and once the learner achieved a certain level of foreign language proficiency, it appears not so important. Thus, it is less able to distinguish among learners with relatively advanced foreign language proficiency than the other two components. Language analytic ability has a linear relationship with L2 success, suggesting that it is important at each stage of L2 development. This is basically the same with memory except at exceptionally advanced levels. However, as Skehan (1998, p. 218) admits that "more research is needed, of course, to make this diagram more than a convenient schematic representation", which is also the motivation for the present study. Actually, there is already some partial evidence that lends support to the hypotheses.

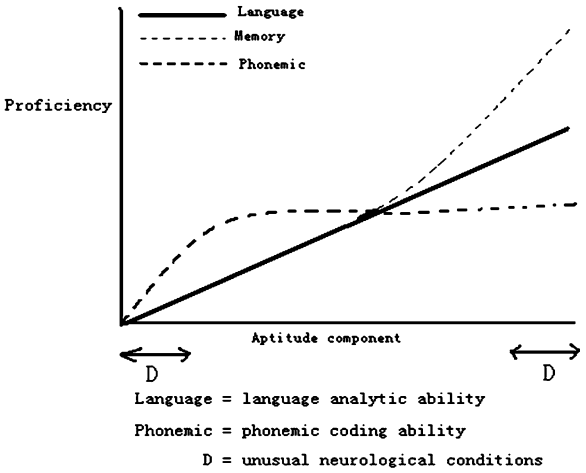

Fig. 13.1 The relationship between aptitude components and proficiency level (Adapted from Skehan 1998, p. 217)

Phonetic coding ability was the component that attracted much attention from Pimsleur et al. (1963) found that auditory ability was important in distinguishing underachieving students from those who showed no difficulty and poor auditory ability often accounted for intelligent or motivated students who seemed unable to learn a foreign language. Sparks and his collaborators (Sparks and Ganschow 1991, 1993, 2001; Sparks et al. 1992; Ganschow et al. 1998) have done a great deal of research on this component. They explored the relationship between phonetic coding ability and measures of mild dyslexia. They noticed that those adolescent or adult novice language learners with low levels in phonological and syntactic coding of their native language tended to experience difficulty in FL learning.

In terms of inductive language learning ability and grammatical sensitivity, Skehan (1998) suggests that the two meld into language analytic ability. Skehan (1986) investigated the relationship between language success and the components of aptitude (language analytic ability and memory ability). By using cluster analysis, he identified patterns in aptitude score profiles associated with success in Arabic and found that some successful language learners relied on their strong language analytic abilities while other successful learners drew more upon their good memory ability. Dekeyser (2000) found that those few adult immigrants who scored within the same range of grammaticality judgment test as child arrivals all got very high scores on MLAT-IV Words in Sentences. Therefore, it can be concluded that language analytic ability may play an important role among adult FL learners and may be related to higher language proficiency.

Memory has received the most interest and attention from researchers and has also undergone considerable change in concept. J. B. Carroll (1981) emphasizes the associative memory between sound and symbols, but Skehan (1982, as cited in Skehan 1998) failed to find any relationship between associative memory and

language learning success. Thus, Skehan (1989, 1998) posits that memory for unfamiliar materials and the ability to impose organizational structures on new materials could better predict the success of language learning. However, Harrington and Sawyer (1992) and Robinson (Robinson 2002) found the relationship between working memory, measured by Deneman and Carpenter's (1980) reading span test, and language learning performance. Skehan (2002) interprets the findings by arguing that in his early research, only a simple span test was used, "which did not require the strong executive working memory, but only rehearsal", while "reading span test is designed to provoke significant computation within working memory" (pp. 75–76). Thus, Skehan concludes that the key issue might be the need to "operate upon the material that is held in immediate memory" (2002, p. 76). Harley and Hart (1997) also failed to find a relationship between associative memory and French proficiency of adolescent English learners but they found that a test of memory for text was correlated significantly with L2 success of early immersion learners (who began a French immersion program from Grade 1).

With regard to the relationship between language aptitude components and different levels of L2 proficiency, there are relatively fewer studies. Winke (2005), who studied foreign language aptitude (measured by MLAT) and working memory of adult English learners of Chinese, found that associative memory, grammatical sensitivity, and phonological memory were associated with learning at the beginning level of Chinese while only phonological memory was associated with learning at the advanced level. Hummel (2009) studied the relationship between aptitude components, working memory and L2 proficiency of adult learners of French at a relatively advanced level and found that grammatical sensitivity and phonological memory could significantly predict the learning of this group. When this group was divided into different groups based on a median split, however, only phonological memory predicted L2 learning of lower-proficiency subgroup, while none of the aforementioned variables predicted L2 learning of higher-proficiency subgroup. Ma and Wang (2011) also studied the relationship between aptitude, working memory and L2 reading among English majors. They found that though all MLAT subtests (except MLAT-V Paired Associates) were correlated with measures of English reading, only MLAT-I Number Learning could predict English reading of the subgroup with lower reading scores while MLAT-IV Words in Sentences and working memory could significantly predict English reading of the subgroup with higher reading scores (the group was also classified by a median split).

In sum, Skehan's hypotheses did not stimulate much empirical research. Existing studies suggest that phonological coding ability seem to play an important role at the initial stage of L2 learning. However, with regard to the role of memory and language analytic ability, relevant studies have not produced consistent findings. Furthermore, studies exploring the relationship between aptitude components and different levels of L2 proficiency all used the MLAT. We know much less about the role of inductive language learning ability and memory and phonetic coding ability measured by subtests of PLAB. The present study aims to

address this issue. Specifically, the study aims to answer the following three questions:

1. What is the relationship between foreign language aptitude and its components and foreign language proficiency among Chinese English majors?
2. Which language aptitude components could predict English majors' proficiency at different stages of development?
3. What are the differences in language aptitude components between higher-proficiency and lower-proficiency groups at different stage of L2 development?

13.4 Methodology

13.4.1 Participants

Sixty-four second-year English majors from a key university in Beijing participated in this study, and 62 finished the aptitude tests. Among them, 9 were males, and 53 were females. Their ages ranged from 18 to 23 with a mean of 20.5 years old. Coming from different parts of the country, 47 students were Hans and the other 15 students were national minorities. On average, they had spent about 7 years learning English when they took the aptitude test. All the participants were from two parallel intact classes enrolled in the same course of Extensive English Reading taught by two different teachers.

13.4.2 Instruments

Owing to practicality, a 45-min aptitude test was designed. Due to lack of appropriate and available tools for measuring Chinese learners' foreign language aptitude, two aptitude subtests were self-developed and the other half adopted the last three subtests of PLAB.

There were five parts in the overall test. Each part was intended to measure different aptitude components. The first two parts were used to measure memory ability, including memory for text and associative memory (see Appendix), which were based on one previous study (Harley and Hart 1997) and modeled on MLAT-VPaired Associates respectively. Inductive language learning ability was measured by PLAB-IV-Language Analysis. Phonetic coding ability was measured by PLAB-V-Sound Discrimination and VI-Sound-Symbol Association. The materials included some slides, a set of booklets, answer sheets, and a questionnaire. The structure of the whole language aptitude test and the points assigned to each part are shown in Table 13.2. There were 143 items in total and each item was worth one point. Detailed information on each subtest of the test is presented in the following section.

The test of memory for text was designed according to Skehan's findings that this aspect of memory ability – the ability "to analyze text, to extract its

Table 13.2 Structure of the language aptitude test

Part	Name	Main focus	N	Task types
1	Memory for text	Memory for text	50	Subjective
2	Paired words	Associative memory	24	Multiple choice
3	Language analysis	Inductive language learning	15	Multiple choice
4	Sound discrimination	Memory	30	Multiple choice
5	Sound-symbol association	Phonetic coding ability	24	Multiple choice
Total	FL aptitude test	/	143	/

propositional content, and remember such content" (Skehan 1989, p. 31)—was significantly correlated with L2 learning success. The test of memory for text in this study was designed on the basis of the work of Harley and Hart (1997).

A narrative Chinese story with about 300 words was selected and prerecorded on the tape. Immediately after listening once to the recording of the story, the students were asked to write down as much of the story as they could recall on the answer sheets within 3 min. The story does not require any special background knowledge to understand. In terms of scoring, the text was deemed to contain 50 information bits based on prepositional analysis. It is generally agreed that a sentence "can be represented by a proposition consisting two or more concepts and some form of relation between them" (D. Carroll 2008, p. 154). For example, sentence "John hit Jack" can be represented as a proposition "hit (John, Jack)". Thus, the following sentences, "Jack was hit by John", "it was John that hit Jack" all can be represented by the same one proposition despite their superficial dissimilarities. The full score for this part was 50 points. A student's score consisted of the number of information bits in each proposition included in their written versions. Another rater was invited to participate in the scoring of this part to decrease the subjectivity in scoring that could have affected the results. The inter-rater reliability coefficient reached .92.

Associative memory test was designed by the author and was modeled on MLAT-4 Paired Associates. First, students were presented a slide showing a list of 24 Icelandic words along with their English equivalents. Then, the students were asked to memorize the English meanings of the 24 words in 2 min, next they needed to choose the corresponding English equivalents for each of them from five choices without looking at the word list in 3 min. Look at the following sample:

fara – go
fara A. road B. ill C. sun D. go E. fly

If the student could choose the choice D by recalling, they would be awarded one point.

With respect to inductive language ability, PLAB-IV Language Analysis was adopted. First, participants read a list of words from a foreign language and the English equivalents of these words. Look at the following sample.

jiban. . . . . .boy, a boy
jojo. . . . .dog, a dog
jiban njojo za. . . . . .A boy likes a dog.

By referring to the above list, they were asked to figure out how the following statement should be expressed in this foreign language.

A dog likes a boy.

This English sentence is followed by four choices and they were asked to choose the correct one.

PLAB-V requires the examinee to differentiate spoken words in an unfamiliar language and learn the meanings of the new words. For the first 15 items, the students were taught 2 words and then must indicate which of two words printed in the test booklet was spoken on the tape. For the subsequent 15 items, the students must choose among all three words and indicate which word was contained in each sentence. PLAB-VI consists of nonsense words based on English syllable structure. The voice on the tape pronounces one of the four words (like trapled, tarpled, tarpdel, trapdel) in each response set, and the students simply indicate which word was spoken.

In terms of the students' English proficiency, their scores on two national English proficiency tests –Test for English Majors Band 4 and Band 8 (abbreviated as TEM-4 and TEM-8) were used as criteria and were collected in 2008 and 2010. TEM-4 contains measures of English listening, vocabulary and grammar, reading and writing while TEM-8 measures English listening, reading, writing and translation. Neither of the two tests contains a subtest measuring speaking. Students are entitled to take the oral test only after they achieve a certain level of written proficiency. TEM-8 is supposed to test students with advanced English levels whereas TEM-4 tests students with intermediate levels. The English Group of the Teaching Guiding Committee for College Foreign Language Majors under the Ministry of Education designed and administered the two tests to second-year and fourth-year English majors respectively in the mainland of China each spring. The two tests have a history of about 20 years and students' scores of the tests serve as important evidence of their English proficiency.

Lastly, a questionnaire was designed to obtain the background information of learners' English learning including sex, age, ethnic background, length of English learning, and their evaluation of the difficulty level of the test.

13.4.3 Procedures

In March 2008, a trial of the test was conducted to test the appropriateness and reliability of the aptitude test. Based on the results of the trial test, the original test was revised. In May, 2008, the revised aptitude test was administered to the participants.

The main study was completed in the last week of May. The whole test took about 45 min and was conducted in one period of normal class. Sixty-four students participated in the study, of whom 62 completed the test. One of the teachers was

invited to invigilate the test with the author to ensure that all the testing procedures be properly followed. Before the test, the purpose and procedures of the test were explained to students in Chinese. Before each of the first three subtests of the aptitude test, the instruction part and a practice session were presented to students on slides. The instruction of the last three aptitude subtests was pre-recorded on the tape.

Students' scores on TEM-4 in 2008 and TEM-8 in 2010 were obtained from the academic office of the department in the school.

13.4.4 Data Analysis

All the data were entered into to one spreadsheet and analyzed by SPSS 16.0. In order to assess the quality of the language aptitude test, descriptive statistics were obtained, including the calculation of means, ranges, and standard deviations of each subscale. Furthermore, for the aptitude subtest of Memory for Text, inter-rater reliability was calculated. For the rest of the aptitude test, item analysis was performed including the difficulty, discrimination index and reliability coefficients by using Cronbach's alpha. To answer the first research question, Pearson product–moment correlational analyses was performed. All tests of correlational significance were two-tailed. The significance level was set at $p < .05$ for this study. To answer the second research question, stepwise multiple regression analysis was conducted. To answer the last question, the participants were first divided into two groups according to the median splits of their TEM-4 and TEM-8 scores. Independent-sample T-test analysis, stepwise multiple regression analyses were employed to compare the differences in language aptitude components between high- and low-proficiency groups.

13.5 Results and Discussion

13.5.1 Students' Performance on English Proficiency Tests and FL Aptitude Tests

Before answering the first question, it is necessary to have a look at the students' performance on language proficiency tests and the aptitude tests. In collecting the data on the participants' English proficiency, one student's TEM-4 score and four students' TEM-8 scores were not available. Table 13.3 presents descriptive statistics about the students' TEM-4 and TEM-8 scores and FL aptitude scores. From the results we can see that the first subtest Memory for Text had a relatively smaller standard deviation, indicating that this subtest was not able to differentiate between the students so well as other subtests.

From Tables 13.3 and 13.4, it can be found that Part 1 Memory for text was the most difficult part of the test. The reason might be that the students were not given

Table 13.3 Descriptive statistics of aptitude test and English proficiency measures

	Min.	Max.	M.	SD
Memory for text	9.50	28.00	17.53	4.55
Paired words	5.00	24.00	18.34	4.27
Language analysis	5.00	15.00	12.35	2.46
Sound discrimination	14.00	30.00	24.16	4.57
Sound-symbol association	2.00	24.00	18.35	3.54
Language aptitude total	63.50	107.50	90.74	10.34
TEM-4	44.00	88.00	69.11	9.71
TEM-8	44.00	82.00	64.23	8.02

Table 13.4 Performance on language aptitude test

	Name	Mean test score (%)	Cronbach's alpha	Mean biserial
Part 1	Memory for text	33.48	/	/
Part 2	Paired words	76.41	.81	0.48
Part 3	Language analysis	82.37	.71	0.44
Part 4	Sound discrimination	80.54	.81	0.39
Part 5	Sound-symbol association	76.48	.75	0.40
Total	Language aptitude	69.86	Part 2–5: .87	Part 2–5: 0.43

enough time to finish this task. The easiest part was Part 3 which measures inductive language learning ability. A possible explanation is that since PLAB is designed for teenagers and the English language involved in this part is rather simple, the inductive language learning ability of the second-year Chinese English majors might have been more developed that the subtest could measure. However, on the whole, we can see that the aptitude test was of appropriate difficulty and reached higher internal consistency and could discriminate students well.

13.5.2 Correlations among Measures of Language Aptitude, Its Components, and L2 Proficiency

Table 13.5 shows the interrelationship between measures of FL aptitude components. First, we can see that although Paired Words and Memory for Text were intended to measure different aspects of memory, their insignificant relationship suggests that they measure distinct cognitive abilities. Paired Words was correlated with Language Analysis significantly at .398 ($p = .001$). This result lends support to J. B. Carroll's (1981) speculation that Language Analysis involves some extent of rote memory ability in that while students were doing the Language Analysis, they also need to establish the association between the newly-learned stimuli and the meanings they represent in a very short time in order to apply the rules underlying the given materials. This result is similar to Harley and Hart (1997)'s findings, in which they found that MLAT-IV Word Pairs and PLAB-IV Language Analysis subtest correlated significantly among both the early and late

Table 13.5 Correlations among measures of language aptitude and its components

	Memory for text	Paired words	Language analysis	Sound discrimination	Sound-symbol association
Paired words	−.077				
Language analysis	.234	.398**			
Sound discrimination	−.155	.032	.267*		
Sound-symbol association	−.123	.111	.200	.444**	
Total	.353**	.526**	.691**	.602**	.577**

Note: *$p < .05$; **$p < .01$

Table 13.6 Correlations between measures of language aptitude and English proficiency

	MfT	PW	LA	SD	SSA	Total	TEM-8
TEM-4	.285*	−.100	.308*	.476**	.339**	.484**	.826**
TEM-8	.299*	.065	.423**	.457**	.251*	.546**	1

Note: *MfT* memory for text, *PW* paired words, *LA* language analysis, *SD* sound discrimination, *SSA* sound-symbol association
*$p < .05$; **$p < .01$

emersion students (at .38 and .54 respectively). Though the participants were different in the two studies, the findings were very similar.

The weak correlation between Sound Discrimination and Language Analysis may be interpreted as that both may involve some extent of rote memory ability. The strong correlation between Sound-Symbol Association and Sound Discrimination was expected as both may have involved the ability to distinguish between and code different sounds.

Table 13.6 shows the relationship between aptitude measures and English proficiency tests. First, the aptitude composite scores were found to be correlated significantly with both TEM-4 and TEM-8 scores. This result is consistent with many of earlier findings (e.g. Gardner and MacIntyre 1992; Ehrman and Oxford 1995; Dai 2006). In terms of the aptitude components, it can be found that all the other aptitude components were correlated with the two English proficiency tests except associative memory measured by Paired Words. This result corroborates some previous findings (e.g. Harley and Hart 1997; Winke 2005), suggesting that associative memory may not play a major role in L2 learning. Even J. B. Carroll (1990) himself admits that he was not confident about the validity of this subtest as its validity seems to vary wildly with different samples.

13.5.3 Differences in Aptitude Components Between Learners at Different Developmental Stages

Table 13.7 shows the results of stepwise regression analysis. As is shown, both Sound Discrimination and Memory for Text were significant predictors of both

Table 13.7 Results of stepwise regression analysis between language aptitude component and TEM-4 and TEM-8

Variable	Step	Predictor	R^2	Adj. R^2	F value[a]	Sig. of F
TEM-4	1	Sound discrimination	.227	.214	17.605	.000
	2	Memory for text	.359	.337	16.492	.000
TEM-8	1	Sound discrimination	.209	.195	15.819	.000
	2	Memory for text	.349	.327	15.786	.000
	3	Language analysis	.394	.363	12.565	.000

Note: Probability for inclusion = .05; probability for exclusion = .01
[a]For equation

TEM-4 and TEM-8. Sound Discrimination, as discussed earlier, may involve more memory ability than the ability to discriminate between different sounds (J. B. Carroll 1981). Sound Discrimination and Memory for Text could jointly explain about 36 and 35 % of the total variance of TEM-4 and TEM-8 respectively. This result suggests the important role of memory ability across different stages of L2 development, thus lending support to Skehan's (1998) hypothesis with respect to memory. However, Language Analysis which was intended to measure inductive language learning ability could only predict the students' TEM-8 scores but not their TEM-4 scores, indicating that this aptitude component becomes more important at the advanced level. Dekeyser (2000) found that those English learners who immigrated to America as adults and achieved native-like proficiency all had outstanding performance on MLAT-IVWords in Sentence. Harley and Hart (1997) also found that inductive language learning ability was associated L2 proficiency of late French immersion learner and speculating that late L2 learners may draw more on their language analytic ability than memory ability. This result seems to give partial support to Skehan's (1998) hypothesis on language analytic ability. One of the interpretations might be that inductive language learning ability might be more associated with productive skills than receptive skills and the former usually develop later than latter ones.

13.5.4 Differences in Aptitude Components Between Subgroups with Higher and Lower TEM-4 and TEM-8 Scores

Tables 13.8 and 13.9 show the differences between the subgroups with higher and lower TEM-4 and TEM-8 scores. First, it can be found that the subgroup with higher TEM-4 scores different from those with lower TEM-4 scores differed significantly at two aptitude component measures – Sound Discrimination and Sound-Symbol Association, suggesting that phonetic coding ability was still important for the learner when they were at the intermediate level of L2 proficiency. However, when the students were in the fourth year of their English study, sound-symbol association was not able to differentiate the two groups any more. Sound discrimination and language analysis became the two factors which could significantly differentiate between the

Table 13.8 Comparison between the two subgroups with higher and lower TEM-4 scores

	Higher TEM-4 (31)		Lower TEM-4 (31)		*t*-test (2-tailed)	
Aptitude	M	SD	M	SD	*t*	*p*
Memory for text	18.06	4.18	17.00	4.90	.920	.361
Paired words	17.90	5.08	18.77	3.29	−.801	.427
Language analysis	12.81	2.48	11.90	2.39	1.461	.149
Sound discrimination	26.00	4.06	22.32	4.35	3.440	.001
Sound-symbol association	19.58	2.17	17.13	4.20	2.886	.006
Total	94.35	9.33	87.13	10.16	2.915	.005

Table 13.9 Comparison between the two subgroups with higher and lower TEM-8 scores

	Higher TEM-8 (31)		Lower TEM-8 (31)		*t*-test (2-tailed)	
Aptitude	M	SD	M	SD	*t*	*p*
Memory for text	18.42	4.38	16.65	4.61	1.553	.126
Paired words	18.45	4.69	18.23	3.88	.207	.837
Language analysis	13.06	2.34	11.65	2.40	2.358	.022
Sound discrimination	26.06	4.06	22.26	4.30	3.586	.001
Sound-symbol association	19.06	3.71	17.65	3.26	1.599	.115
Total	95.06	9.43	86.42	9.48	3.600	.001

two subgroups with higher and lower TEM-8 scores. This result again shows the importance of memory ability across different stages of L2 development. Meanwhile, it also suggests that the ability to associate sounds and their written symbols were not so important once the learner achieves a certain level of L2 proficiency while inductive language learning ability only comes into play at a higher proficiency level.

In order to further find out which aptitude component could significantly predict L2 proficiency of learners at different proficiency levels, another stepwise multiple regression analysis was performed. It can be seen from Table 13.10 that Memory for Text and Sound Discrimination were strongly associated with the subgroup with lower TEM-4 scores, showing the importance of memory at this stage. Perhaps at this stage, the students needed to memorize a larger number of words, phrases, sentences or texts, which might heavily depended upon their memory ability. Previous studies (Skehan 1986; Harley and Hart 1997) suggest that younger learners tend to rely more on their memory ability, while older learners tend to rely more on their language analytic ability. This result seems to suggest that the relationship between aptitude components and L2 learning is not only associated with age but also stages of L2 learning. For the subgroup with higher TEM-4 scores, Sound-Symbol Association was the only significant predictor of their English proficiency. This result is a little bit difficult to interpret and one possible interpretation is that this ability is more related with written than with oral skills and the former might have been developed later than the later.

For the group with lower TEM-8 scores, Language Analysis was the significant predictor of their English proficiency, indicating inductive language learning ability tends to be more important at the lower advanced level. However, for the group

Table 13.10 Stepwise regression analyses: predicting English proficiency at different levels of proficiency from the dimensions of language aptitude

Group	Step	Predictor	R^2	Adj. R^2	F value[a]	Sig. of F value
TEM-4 lower	1	Memory for text	.141	.112	4.776	.037
	2	Sound discrimination	.309	.260	6.267	.006
TEM-4 higher	1	Sound-symbol association	.126	.096	4.192	.050
TEM-8 lower	1	Language analysis	.234	.208	8.880	.006
TEM-8 higher	1	No entry				

Note: Probability for inclusion = .05; probability for exclusion = .01

with higher TEM-8 scores, none of the aptitude components could significantly predict the students' English proficiency. This result supports some previous findings (Winke 2005; Hummel 2009) and could be interpreted as follows. First, it might be due to the limitations of the aptitude test, indicating that some other possible language aptitude components like working memory might be at play at the advanced stage of L2 development. Another possibility is that some other individual variables (e.g. motivation) might be playing a more important role at this stage. Further studies are needed to clarify this issue.

13.6 Conclusion

This study aims to test Skehan's hypotheses regarding the relationship between different aptitude components and L2 proficiency at different developmental stages. On the whole, most of Skehan's (1998) hypotheses were supported. phonetic coding ability was shown to play a major role at earlier stages. However, different from Skehan's hypotheses, memory seemed to be important at all stages except at the rather advanced level, and inductive language learning ability began to play a major role beyond the intermediate level. None of the aptitude components measured by the aptitude test used in this study could predict L2 proficiency at the higher advanced level. Of course, the findings are closely related to the specific aptitude measures used in the study.

Despite these findings, there were some severe limitations. Firstly, the PLAB was designed for native English speakers aged between 13 and 19. Though the level of English language proficiency required for taking the test is assumed to be very low, it is unavoidable that some students with lower levels of English proficiency might have been biased. Therefore, in future studies a language aptitude test suitable for Chinese learners of foreign languages is in urgent need to be developed. Secondly, this study only explored a limited number of aptitude components, it is hoped that in future studies, more aptitude components like grammatical sensitivity and working memory could be included. Thirdly, in terms of language proficiency, this study only took the composite scores. However, it is expected that the results will be more informative if more refined measures of language proficiency like measures of different aspects of language skills and knowledge are taken in future studies.

Appendix: Part 2 Paired Words

1. haf	A. say	B. land	C. sea	D. walk	E. money
2. dv?l	A. sing	B. start	C. bed	D. stop	E. eye
3. lesa	A. land	B. hope	C. read	D. hand	E. write
4. fugl	A. tree	B. bird	C. ant	D. machine	E. light
5. elda	A. fish	B. altar	C. cook	D. cold	E. juice
6. synda	A. speak	B. swim	C. eye	D. heart	E. cake
7. hróp	A. roof	B. fall	C. good	D. call	E. flower
8. andlit	A. fly	B. sing	C. push	D. ground	E. face
9. áta	A. quiet	B. food	C. water	D. clothes	E. shoot
10. hl?ja	A. sky	B. left	C. house	D. laugh	E. shake
11. byssa	A. gun	B. road	C. jump	D. pen	E. night
12. dyr	A. dog	B. lash	C. die	D. animal	E. bed
13. kl?ei	A. cold	B. clothes	C. key	B. night	E. hair
14. skrifbore	A. kite	B. letter	C. hope	D. can	E. desk
15. rétt	A. turn	B. eat	C. people	D. grass	E. pull
16. skera	A. ice	B. break	C. cut	D. ship	E. egg
17. tungl	A. climb	B. moon	C. island	D. jump	E. lamp
18. st?kk	A. march	B. close	C. door	D. jump	E. hear
19. eyja	A. wind	B. strength	C. ship	D. island	E. flow
20. maeur	A. person	B. see	C. cool	D. cry	E. milk
21. vinna	A. travel	B. salt	C. picture	D. job	E. leave
22. útlit	A. top	B. look	C. knife	D. smile	E. drink
23. hlaupa	A. sport	B. day	C. laugh	D. face	E. run
24. vinur	A. map	B. buy	C. friend	D. music	E. cook

References

Carroll, John B. (1962). The prediction of success in intensive foreign language training. In R. Glaser (Ed.), *Training research in education* (pp. 87–136). Pittsburgh, PA: University of Pittsburgh Press.

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

Carroll, John B. (1990). Cognitive abilities in foreign language aptitude: Then and now. In T. Parry & C. Stansfield (Eds.), *Language aptitude reconsidered* (pp. 11–19). Englewood Cliffs, NJ: Prentice Hall.

Carroll, D. W. (2008). *Psychology of language* (5th ed.). Beijing: Foreign Language Teaching and Research Press.

Carroll, J. B., & Sapon, S. M. (1959/2002). *Modern language aptitude test*. Bethesda: Second Language Testing.

Dai, Y. (2006). The effects of language aptitude on second language acquisition. *Foreign Language Teaching and Research, 6*, 451–459.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19*, 450–466.

DeKeyser, R. (2000). The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition, 22*, 493–533.

Dornyei, Z. (2005). *The psychology of the language learner: Individual differences in second language acquisition*. Mahwah: Lawrence Erlbaum.

Dornyei, Z., & Skehan, P. (2003). Individual differences in second language learning. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 589–630). Oxford: Blackwell.

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

Erlam, R. (2005). Language aptitude and its relationship to instructional effectiveness in second language acquisition. *Language Teaching Research, 9*, 147–172.

Ganschow, L., Spark, R., & Jovorsky, J. (1998). Foreign language learning problems: An historical perspective. *J Learn Disabil, 31*, 248–258.

Gardner, R., & MacIntyre, P. (1992). A student's contribution to second language learning. Part 1: Cognitive variables. *Language Teaching, 25*, 211–220.

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

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

Hummel, K. M. (2009). Aptitude, phonological memory, and second language proficiency in nonnovice adult learners. *Applied Psycholinguistics, 30*, 225–249.

Ioup, G., Boustagui, E., Tigi, M., & Moselle, M. (1994). Reexamining the critical period hypothesis: A case study of successful adult SLA in a naturalistic environment. *Studies in Second Language Acquisition, 16*, 73–98.

Kiss, C., & Nilolov, M. (2005). Developing, piloting, and validation an instrument to measure young learners' aptitude. *Language Learning, 55*, 99–150.

Ma, Z., & Wang, T. (2011). The effects of language aptitude and working memory on L2 reading comprehension. *Shandong Foreign Language Teaching, 32*, 41–47.

Miyake, A., & Friedman, N. P. (1998). Individual differences in second language proficiency: Working memory as language aptitude. In A. F. Healy,& L. E. Bourne Jr (Eds.), *Foreign language learning:Psycholinguistic studies on training and retention* (pp. 339–364). Mahawah, NJ: Lawrence Erlbaum.

Pimsleur, P., D. M. Sunderland, and R. McIntyre. 1963. *Underachievement in foreign language learning* (ERIC Document Reproduction Service No. ED018160). Retrieved from ERIC database.

Pimsleur, P., Reed, D. J., & Stansfield, C. W. (2004). *Pimsleur language aptitude battery: Manual*. Bethesda: Second Language Testing, Inc.

Robinson, P. (2002). Effects of individual differences in intelligence, aptitude and working memory on adult incidental SLA: A replication and extension of Reber, Walkenfeld and Hernstadt, 1991. In P. Robinson (Ed.), *Individual differences and instructed language learning* (pp. 211–266). Amsterdam: John Benjamins.

Robinson, P. (2005). Aptitude and second language acquisition. *Annual Review of Applied Linguistics, 25*, 46–73.

Rysiewicz, J. (2008). Cognitive profiles of (un)successful FL Learners: A cluster analytical study. *The Modern Language Journal, 92*, 87–99.

Sawyer, M., & Ranta, L. (2001). Aptitude, individual differences, and instructional design. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 319–353). Cambridge: Cambridge University Press.

Skehan, P. (1986). Cluster analysis and the identification of learner types. In V. Cook (Ed.), *Experimental approaches to second language learning* (pp. 81–94). Oxford: Pergamon.

Skehan, P. (1989). *Individual differences in second language acquisition*. London: Edward Arnold.

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.

Skehan, P. (2002). Theorizing and updating aptitude. In P. Robinson (Ed.), *Individual differences and instructed language learning* (pp. 69–93). Amsterdam/Philadelphia: John Benjamins.

Sparks, R., & Ganschow, L. (1991). Foreign language learning differences: Affective or native language aptitude differences? *The Modern Language Journal, 75*, 3–16.

Sparks, R., & Ganschow, L. (1993). Searching for the cognitive locus of foreign language learning difficulties: Linking first and second language learning. *The Modern Language Journal, 77*, 289–302.

Sparks, R., & Ganschow, L. (2001). Aptitude for learning a foreign language. *Annual Review of Applied Linguistics, 21*, 90–111.

Sparks, R., Ganschow, L., Javorsky, J., Pohlman, J., & Patton, J. (1992). Test Comparisons among students identified as high-risk, low-risk, and learning disabled in high school foreign language courses. *The Modern Language Journal, 76*, 142–159.

Wesche, M. B. (1981). Language aptitude measures in streaming, matching students with methods, and diagnosis of learning problems. In K. C. Diller (Ed.), *Individual differences and universals in language learning aptitude* (pp. 119–154). Rowley, MA: Newbury House.

Winke, P. 2005. *Individual differences in adult Chinese second language acquisition: the relationships among aptitude, memory, and strategies for learning*. Unpublished doctoral dissertation, Georgetown University.