The Impact of Language Aptitude on Grammar Attainment of EFL Learners with Different Ages of Onset in the Chinese Context

Zhao Haiyong[1]*, Luo Shaoqian[2]

[1]Capital University of Economics and Business, Beijing, China

[2]Beijing Normal University, Beijing, China

Abstract

The effect of language aptitude on grammar attainment of second language learners with different ages of onset became one of the hot topics in the last decades. However, most previous studies were conducted in a second language context with inconsistent results, and few studies were conducted in a foreign language context. To fill this gap, the current study investigated the impact of language aptitude on grammar attainment of EFL learners with different ages of onset in the Chinese context. A Timed Grammatical Judgment Test and an Untimed Grammatical Judgment Test, a Metalinguistic Knowledge test, and a Language Aptitude test were used in the study, and 112 participants were sampled. The results show that the aptitude scores of the three groups of EFL learners with different ages of onset have no significant difference; language aptitude moderates performance on the Untimed Grammatical Judgment Test and the Metalinguistic Knowledge test, which was mainly reflected in the earlier starters; and higher aptitude learners outperform lower aptitude learners in all three grammar tests. Overall, this study provides evidence to support the impact of language aptitude on EFL learners in the Chinese context.

Keywords: *Age of Onset, Grammar Attainment, Language Aptitude*

Introduction

The past decades have witnessed the development of second language acquisition which

investigated the individual difference (ID) factors influencing the learning process and attainment (e.g., Dörnyei & Ryan, 2015; Li et al., 2022). These ID factors include many attributes, such as the age of onset, learning strategies, motivations, learning style, language aptitude, involvement, anxiety, etc. Among them, the age of onset (AO) consistently accounts for the greatest proportion of variance (typically around 30%) in ultimate second language attainment, while language aptitude, as the second strongest predictor variable, typically accounts for 10%-20% of the variance (Granena & Long, 2013a).

Language aptitude is a complex ability, which combines individual perceptual and cognitive abilities and refers to the specific talent or potential allowing an individual to learn a foreign or second language easier and faster than most peers (Carroll, 1981). Research into foreign language aptitude has been prevalent in the United States since the 1950s and 1960s, which was followed by four decades of silence before picking up renewed steam around the turn of the twenty-first century (Wen et al., 2019). During this period, scholars proposed different models of language aptitude (Grigorenko et al., 2000; Robinson, 2007; Skehan, 1998; Sparks & Ganschow, 2001) which provided implications for the development of language aptitude tests. Grammar is one of three language factors in a second language, and the relationship between language aptitude and grammar learning is a central topic in second language learning (Li, 2015). Meanwhile, it is useful to reveal the second language process by discussing the relationship between language aptitude and language attainment (Skehan, 2015). In the past two decades, with more studies investigating the age effects on L2 grammar attainment, the effect of language aptitude on grammar attainment of second language learners with different ages of onset became one of the hot topics. However, most studies were conducted in a second language context with inconsistent results, and few studies were conducted in a foreign language context. To fill this gap, the current study investigated the impact of language aptitude on grammar attainment of EFL learners with different ages of onset in the Chinese context.

Literature Review

Although many studies were conducted on the relationship between language aptitude, age of onset, and language attainment in the past decades in a second language context, they did not get consistent results. Studies reveal that language aptitude influences not only the language attainment of later starters (DeKeyser, 2000; DeKeyser et al., 2010; Huang, 2014) but also earlier starters (Abrahamsson & Hyltenstam, 2008; Bylund et al., 2012; Granena, 2013a, 2014). Granena (2013b) found that although language aptitude can influence language attainment of later starters, not all learners with higher language aptitude can achieve native speakers' language proficiency. However, Granena and Long (2013b), who used the auditory Timed Grammatical Judgment Test (TGJT), found that language aptitude failed to moderate the second language attainment of both earlier and later starters. Although DeKeyser (2000), DeKeyser et al. (2010), Abrahamsson & Hyltenstam (2008), and Granena & Long (2013b) all used the auditory Grammatical Judgement Test (GJT), their testing conditions are quite different. In DeKeyser's studies (DeKeyser, 2000; DeKeyser et al., 2010), each sentence was read twice the time with 3-second intervals between each time, and 6-second intervals between every sentence. In

Abrahamsson & Hyltenstam (2008), both the auditory and visual GJTs were used, the grammar structures were quite complex, and the participants were successful second language learners and reached near-native levels. Still, they had one thing in common, that is, the participants had time to think and judge the sentences consciously. In their research, the sentences were played one by one within a given time without stopping.

Different from the above studies which were conducted in a second language context, Harley and Hart (1997, 2002) conducted studies in an immersion context. Harley and Hart (1997) found that memory was the main predicting factor of French attainment, while language analysis ability was the only predicting factor for later starters. In this study, the early and later starters faced different teaching styles, so other researchers doubted that the relationship between language aptitude and age may result from the interaction between aptitude and teaching approach. To further test the result, they conducted another study. In Harley and Hart (2002), participants participated in a three-month exchange program. The result showed that language analysis ability still could significantly predict their French attainment. The participants' study time was so short that it was considered as the rate of study, not attainment. While in a foreign language context, Larson-Hall (2008) discussed the effect of language aptitude on L2 grammar attainment. The study showed that language aptitude could significantly predict the auditory GJT of early starters ($r=0.32$, $p<0.05$, power=0.71).

The above studies mainly used the Modern Language Aptitude Test (MLAT), LLAMA, and other language aptitude tests to measure language aptitude, while Huang (2014) measured participants' aptitude by self-assessment, and DeKeyser et al. (2010) measured aptitude with the psychological measurement tool in participants' native language, though the reliability of the instruments needs to be verified furtherly. For example, Bokander & Bylund (2020) examined the evidence for the internal validity of the LLAMA test battery. They found only one out of the four subtests (LLAMA B) produced scores that fit a latent trait model with sufficient accuracy. LLAMA D and LLAMA F have low reliability, and LLAMA E might not unequivocally target phonetic encoding. And they also found LLAMA battery does not tap into the three components of aptitude as conceptualized by Skehan (1998). Meanwhile, Suzuki (2021) also proposed several recommendations to redesign and extend LLAMA D as a potential aptitude test for proceduralization. Therefore, for the time being, it might not be appropriate to employ LLAMA as a high-stake test for education or placement purposes, and any practical application of knowledge generated by the LLAMA is premature in the absence of appropriate validation procedures.

The research topic also attracted the attention of Chinese scholars, however, only a few empirical studies were conducted in the Chinese context. Chen & Chen (2019) investigated the role of explicit and implicit aptitudes in the acquisition of second language grammar. They compared both children's and adults' aptitudes and their explicit and implicit grammar learning performance. It was found that children learned implicitly, and the implicitness/explicitness distinction in their aptitude did not correlate with their learning performance; adults learned explicitly, and the implicitness/explicitness distinction in their aptitude was more obvious than in children, and the explicitness was correlated with their learning performance. Li et al. (2019)

investigated the relationship between AO, language aptitude, and L2 grammar attainment. They found that there is a significant correlation between aptitude and L2 grammar attainment in earlier and later starters. They also found that implicit language aptitude was related to GJT scores in the younger learners, while explicit language aptitude was related to GJT scores in the older learners. But they did not discuss the impact of language aptitude on various stimulus tasks and different grammar rules. Furthermore, the language aptitude test they used was not developed for Chinese foreign language learners.

The Current Study

Given these research gaps in previous research, the current study was designed to use three different grammar tests and one language aptitude test developed specifically for Chinese EFL learners, to investigate the potential effect of language aptitude on grammar attainment among EFL learners in China.

Research Questions

(1) Are there any significant differences in language aptitude of foreign language learners with different ages of onset in China?

(2) To what extent does language aptitude affect the grammar attainment of EFL learners with different ages of onset in the Chinese context?

Participants

112 participants were conveniently sampled in the current study who were from 17 different cities in the same province. The participants: (1) learned English for at least 5 years continuously; (2) had never taken part in any extracurricular English classes and never lived or studied abroad; (3) never or occasionally received parents' English tutoring; (4) had no language or study difficulties. Participants respectively finished an English learning background questionnaire, a Timed Grammatical Judgment Test (TGJT), an Untimed Grammatical Judgment Test (UTGJT), a Metalinguistic Knowledge (MK) test, and a language aptitude test. The participants' AO is 3-14 years old, their age of testing is 17-20 years old, and they have learned English for 6-16 years. Based on participants' AO, they were divided into three groups: Primary one (AO=3-7), Primary three (AO=8-10), and Secondary one (AO=11-14). The reasons for the dividing are: (1) the criterion was used by many similar studies (Qureshi, 2016); (2) this criterion is useful to compare the language attainment of learners with different AO.

Table 1

Participants' Information

AO	Number		Age of Testing (AOT)	Years of Exposure	Hours of Studying English
	Male	Female			
R=3-7	17	16	M=18.39 R=17-20	M=12.82 R=10-16	M=1748.12 R=1280-2408
R=8-10	21	44	M=18.48 R=17-20	M=9.42 R=7-12	M=1472 R=1008-1728
R=11-14	9	5	M=19 R=18-20	M=6.71 R=6-8	M=903.43 R=600-1088

Instruments

All the instruments used in the current study were in written form and piloted before formal testing. The result shows that their reliability and validity are acceptable.

Questionnaire

The questionnaire used in the current study included 12 items to get personal information, English learning background, and extracurricular English learning experience.

Grammatical Judgment Test (GJT)

The GJT, designed by Ellis et al. (2009), measured 17 basic English grammar rules (Ellis, et al., 2009). These rules usually bring difficulties for learners who usually make mistakes when using these rules, are learned in different learning stages, and cover both syntax and morphology. Both TGLT and UTGJT were used in the current study.

The TGJT was used to measure implicit grammar knowledge and included 68 sentences among which 34 are grammatical and 34 are ungrammatical. The sentences are ordered randomly. All the words in the test are from the most common 2000 words (Nation, 1990). The participants finished the test within the given time and chose "Right" if the sentence is grammatical and "Wrong" if the sentence is ungrammatical. One point for each sentence and the total score is 68 points. Meanwhile, grammatical and ungrammatical sentences were computed respectively, and the total score of each part is 34 points. The given time was decided according to the performance of native speakers. Ellis et al. (2009) piloted the test in America with English native speakers. They computed the time for each sentence by every native speaker firstly and then got the base time for each sentence by averaging each native speaker. Considering the lower processing speed of the second language learners, they added 20% time for second language learners based on the base time. In the current study, we followed a similar procedure. After testing, the internal reliability of this test is calculated, and the result is 0.9.

The items in UGJT were the same as those in TGJT, but the UGJT measured explicit grammar knowledge and didn't have a time limit. The internal reliability of the test is 0.82.

Metalinguistic Knowledge (MK) Test

The MK test, measuring explicit knowledge, was used to measure learners' mastery and application of grammar rules. The first part of this test covered 17 multiple-choice items. The ungrammatical section in each sentence was underlined. The participants were required to choose one sentence from the choices to explain the wrong section (Ellis et al., 2009). One point is for each item, and the total score is 17 points. The second part of the test covered 17 grammar error-correction items. The participants are required firstly to judge whether each sentence is grammatical or not, then underline the wrong section, and then correct it. Two points are for each item, among them, one point for underlining the ungrammatical part, and one point for error correction. The sentences in this part meet the following criteria: all ungrammatical structures can be corrected, participants are familiar with all the words in each sentence, and the grammar rules in this part are the same as those in the GJT.

Although the MK test didn't have a time limit, all participants could finish the test within 20 minutes. The total score is 51 points. The internal reliability of the MK test is 0.76.

Language Aptitude (LA) Test

The language aptitude test used in the current study was the Foreign Language Aptitude Test for Chinese Learners of Foreign Languages (FLAT-C), which was developed based on Carroll's aptitude theory (Carroll, 1965) and Skehan's second language learning theory (Skehan, 1998) with the Rasch model by Li & Luo (Li & Luo, 2019). The test reliability was above 0.9, and the infit values of all items are between 0.7 and 1.3. The test includes number learning, phonetic script, paired associates, words in sentences, and language analysis (Li & Luo, 2019). DeKeyser (2000) used MLAT IV to investigate the relationship between language aptitude and age effects, and Li et al. (2019) considered the relationship between language analysis ability and grammar to be important, therefore, the current study used the fourth part of FLAT-C, *words in sentences*, to measure grammatical sensitivity (one of language analysis abilities according to Peter Skehan) and covered 24 items. Language analysis ability treats language as an object of analysis and arrives at linguistic generalizations and is at the core of the constructs of language learning aptitude and metalinguistic awareness (Roehr-Brackin & Tellier, 2019). The score of the current language aptitude test is from 0 to 24 points.

Data Collection and Analysis Procedure

The participants finished the questionnaire, TGJT, UGJT, MK test, and LA test in the office or classroom when their time was available, and the data collection lasted one month. After collection, the researchers assessed the grammar tests with two rounds, and then all the data were input into IBM SPSS Version 22 for the statistical analysis. Stata 14.0 was used to compute the effect size.

Results

Language Aptitude

Table 2 shows the language aptitude of learners in three groups. It shows that with the increase of AO, language aptitude becomes higher, but the SD becomes larger. The Skewness, kurtosis and Levene test (p=0.683>0.05, Levene=0.38) reveal that One-way ANOVA can be used which shows F=0.13 (p=0.877>0.05). The effect size is gotten with Stata, that is, η^2=0.002, 95% CI [0, 0.03]. The above data show that the language aptitude of the three groups is not significantly different.

Table 2

Description of Language Aptitude

Group	Mean	SD	Range	skewness	kurtosis
Primary 1	14.42	2.32	10-21	0.68	0.83
Primary 3	14.68	2.46	7-20	-0.31	0.74
Secondary	14.71	2.76	9-19	-0.55	-0.01

Based on the standard score, participants are divided into the higher aptitude group (Z-score>0.5), intermediate aptitude group (-0.5<Z-score<0.5), and lower aptitude group (Z-score<-0.5) (Granena, 2012). There are 8 and 11 participants in primary one respectively with higher

and lower aptitude, 24 and 18 participants in the primary-three group respectively with higher and lower aptitude, and 3 and 4 participants in the secondary-one group respectively with higher and lower aptitude. Among them, the primary-three group has more participants with higher language aptitude, but fewer participants with lower aptitude.

The Effect of Language Aptitude on L2 Grammar Tests

Table 3 shows language aptitude and TGJT have no significant difference in three sub-groups and the whole group, however, language aptitude is significantly correlated with UGJT (r=0.29, p=0.002<0.01) and MK (r=0.32, p=0.001<0.01) in the whole group. In UGJT, language aptitude has a significant correlation with L2 grammar attainment only in Primary three (r=0.32, p=0.010<0.05), while in the MK test, there is a significant correlation between language aptitude and L2 grammar attainment in Primary one (r=0.38, p=0.028<0.05) and Primary three (r=0.31, p=0.013<0.05).

Table 3

Correlation Between Aptitude and L2 Grammar

	Whole group	Primary one	Primary three	Secondary one
LA-TGJT	0.16 (p=0.085)	0.31 (p=0.080)	0.21(p=0.094)	-0.20 (p=0.485)
LA-UGJT	0.29** (p=0.002)	0.29 (p=0.099)	0.32* (p=0.01)	0.17 (p=0.57)
LA-MK	0.32** (p=0.001)	0.38* (p=0.028)	0.31*(p=0.013)	0.33 (p=0.249)

When TGJT, UTGJT, and MK test are used as dependent variables, and language aptitude is used as a covariate, the covariance analysis shows that the main effect of language aptitude on TGJT attainment is not significant (F=0.82, p=0.366>0.05, η^2=0.01), while the main effects of language aptitude on UTGJT and MK test attainment are significant (F=5.81, p=0.018<0.05, η^2=0.052; F=10.78, p=0.001<0.05, η^2=0.092). That is, language aptitude can explain respectively 5.2% and 9.2% variance of UTGJT and MK attainment, and both belong to intermediate effect size. When language aptitude effect is controlled, three sub-groups have no significant difference in TGJT, UTGJT and MK test (F=0.85, p=0.429>0.05, η^2=0.02; F=0.35, p=0.704>0.05, η^2=0.01; F=0.39, p=0.681>0.05, η^2=0.01). The interaction effect of language aptitude and AO is also not significant on three grammar tests (F=1.49, p=0.231>0.05, η^2=0.03; F=0.36, p=0.701>0.05, η^2=0.01; F=0.47, p=0.629>0.05, η^2=0.01).

When language aptitude is controlled, the results show that with the increase of AO, the TGJT and MK test attainments decrease, while in UTGJT, the Primary-three group gets the highest score, and the Primary-one group gets the lowest score, but there are no significant differences among three groups in all three grammar tests. The results support the fact that earlier starters do not have a significant advantage in L2 grammar attainment in a foreign language context. In the current study, learners in the Secondary-one sub-group started English learning 3-6 years later than learners in the other two groups, but they could catch up with the other two groups after 6-7 years of learning or a shorter time.

Table 4 shows the performance of learners with higher and lower language aptitude in three

different L2 grammar tests. In Primary one, learners with higher aptitude get higher scores than those with a lower aptitude in all three tests, but there is a significant difference only in the MK test (t=3.48, p=0.003<0.05). In Primary-three, learners with higher aptitude perform better than those with a lower aptitude in all three grammar tests, but they have a significant difference only in UTGJT and MK test (t=2.04, p=0.048<0.05; t=2.06, p=0.046<0.05). Different from the last two groups, in the Secondary-one group, learners with higher aptitude get lower scores than those with a lower aptitude in TGJT, but perform better in UGJT and MK tests, however, they aren't significantly different in all three grammar tests (t=-0.02, p=0.987>0.05; t=.054, p=0.610>0.05; t=1.41, p=0.251>0.05).

Table 4

L2 Grammar Attainment of Learners with Higher and Lower Aptitude

	Primary one		Primary three		Secondary one	
Test	Higher (n=8)	Lower (n=11)	Higher (n=24)	Lower (n=18)	Higher (n=3)	Lower (n=4)
TGJT	49.13 (SD=11.18)	40.18 (SD=4.14)	49.21 (SD=7.82)	43.11 (SD=12.56)	42.33 (SD=15.01)	42.50 (SD=11.48)
UGJT	55.75 (SD=7.67)	47.55 (SD=10.40)	56.79 (SD=5.46)	52.11 (SD=9.33)	56.00 (SD=5.29)	53.25 (SD=7.37)
MK	33.50 (SD=5.76)	23.18 (SD=6.69)	30.92 (SD=6.21)	27.17 (SD=5.31)	31.00 (SD=1.00)	24.50 (SD=9.15)

The Effect of Language Aptitude on Different Stimulus Tasks

Table 5 shows the correlation between language aptitude and stimulus tasks. It shows that there is a significant correlation between language aptitude and ungrammatical sentences in Primary one, Primary three, and the whole group, but their language aptitude is not significantly correlated with grammatical sentences. However, there is no significant correlation between Secondary one learners' aptitude and all the stimulus tasks.

Table 5

Correlation Between Language Aptitude and Stimulus Tasks

Test	Stimulus tasks	Whole group	Primary one	Primary three	Secondary one
TGJT	Grammatical sentence	0.05 (p=0.614)	-0.03 (p=0.889)	.013 (p=0.316)	-0.05 (p=0.855)
	Ungrammatical sentence	0.24* (p=0.011)	0.52** (p=0.002)	0.25* (p=0.044)	-0.35 (p=0.224)
UGJT	Grammatical sentence	0.14 (p=0.155)	-0.002 (p=0.991)	0.16 (p=0.213)	0.39 (p=0.172)
	Ungrammatical sentence	0.33** (p=0.000)	0.45** (p=0.009)	0.37** (p=0.002)	-0.28 (p=0.336)

In most stimulus tasks, higher aptitude learners perform better than lower aptitude learners, and only higher aptitude learners in the Secondary-one group get lower scores than lower

aptitude learners in ungrammatical sentences. No matter whether it is a timed test or an untimed test, higher aptitude learners in three groups perform better in ungrammatical sentences than that in grammatical sentences. Lower aptitude learners in Primary one and Primary three perform better in grammatical sentences than that in ungrammatical sentences, while lower aptitude learners in Secondary one get higher scores in ungrammatical sentences than that in grammatical sentences. These suggest that higher aptitude learners can perform better in grammatical judgment tests, and they have an advantage, especially in explicit grammar tests, while lower aptitude learners perform better in implicit grammar tests.

The Independent t-test shows that for Primary one and Primary three learners, there is a significant difference between higher aptitude and lower aptitude learners in ungrammatical sentences in timed (t=3.66, p=0.002; t=2.39, p=0.022) and untimed conditions (t=3.54, p=0.005; t=2.98, p=0.005), while there is no significant difference between them in grammatical sentences in timed (t=0.20, p=0.84; t=1.07, p=0.293) and untimed conditions (t=0.27, p=0.791; t=0.79, p=0.433). For learners in Secondary one, there is no significant difference between higher and lower aptitude learners in ungrammatical sentences (t=-0.21, p=0.844; t=-0.51, p=0.632) and grammatical sentences (t=0.13, p=0.900; t=1.31, p=0.247) in both timed and untimed conditions.

The Effect of Language Aptitude on L2 Grammar Rules

In TGJT, the language aptitude of participants in Primary one is only significantly correlated with Dative alternation (r=0.39, p=0.025); for Primary three, participants' language aptitude is significantly correlated with since/for (r=0.33, p=0.008), possessive -s (r=0.32, p=0.01), and comparatives (r=0.31, p=0.012); but for Secondary one, participants' language aptitude has no significant correlation with any grammar rules. But taking all the participants as a whole, it is found that language aptitude has a significant correlation with comparatives (r=0.26, p=0.005) and adverb placement (r=0.19, p=0.046).

In UTGJT, the language aptitude of participants in Primary one is only significantly correlated with adverb placement (r=0.38, p=0.03). Language aptitude of participants in Primary three has a significant correlation with possessive -s (r=0.25, p=0.049), comparatives (r=0.33, p=0.008), and adverb placement (r=0.31, p=0.012). However, language aptitude in Secondary one still has no significant correlation with any grammar rules. But taking all the participants as a whole, it is found that language aptitude has a significant correlation with verb complements (r=0.23, p=0.014), comparatives (r=0.24, p=0.011), and adverb placement (r=0.31, p=0.001).

In MK, more grammar rules have a significant correlation with language aptitude, especially in the whole group and the Primary-three sub-group. In the Primary-one sub-group, only Ergative Verb is significantly correlated with language aptitude (r=0.49, p=0.004), and in Secondary-one, all the L2 grammar rules have no significant correlation with language aptitude. But in Primary three, more grammar rules have significant correlation with language aptitude: since/for (r=0.38, p=0.002), possessive -s (r=0.27, p=0.028), plural -s (r=0.26, p=0.035), embedded questions (r=0.29, p=0.019) and adverb placement (r=0.31, p=0.013). There are six L2 grammar rules which are significantly correlated with language aptitude in the whole group: since/for (r=0.24, p=0.01), Ergative verbs (r=0.20, p=0.036), possessive -s (r=0.27, p=0.004), plural -s (r=0.21, p=0.03), embedded questions (r=0.25, p=0.008) and adverb placement (r=0.25,

p=0.007).

In TGJT, higher language aptitude learners get higher scores in since/for, modal verbs, and yes/no questions, while lower aptitude learners get higher scores in modal verbs, indefinite article, and since/for, but both groups get the lowest scores in relative clauses. In UTGJT, both higher and lower aptitude learners get the highest grammar score in the indefinite article (M=3.79; M=3.58), and both groups master the same rules in the lowest seven grammar structures, that is, third-person -s, possessive -s, unreal conditionals, ergative verbs, comparatives, embedded questions, and dative alternation. Different from TGJT, in UTGJT, both higher and lower aptitude learners master the same grammar structures in the untimed condition. In MK, the higher aptitude learners get the highest score in since/for and plural -s, and the lowest score in Question tags and modal verbs; however, lower aptitude learners get the highest score in plural -s and regular past tense, and the lowest score in modal verbs and question tags. In MK, the plural -s is the easiest grammar rule, but question tags and modal verbs are the most difficult with which the scoring rate of lower aptitude learners is only 12% and 11% respectively.

Discussion

To recapitulate, there are several major findings from the current study. First, with the increase of AO, participants' aptitude increases as well, but the three sub-groups' aptitude has no significant difference, which is the same as the result in Larson-Hall (2008), however, this result has both similarities and differences with Granena and Long (2013b) and Harly and Hart (1997). Granena and Long used LLAMA to measure aptitude and found that with the increase of AO, participants' aptitude decreased, but there is no significant difference between them. Based on the computer, the LLAMA aptitude test, independent of participants' native language, with artificial language, pictures and verbal materials, measures learners' aptitude. Harly and Hart found that later learners outperformed earlier learners in aptitude, but there is no significant difference between them. In their study, they used MLAT, Wechsler scale, and PLAB to measure aptitude. Different from them, Granena (2012) found that earlier starters performed better than later starters in aptitude, but there is a significant difference between them. According to Granena, the significant difference results from many factors, such as positive cognition of earlier starters, later starters being older than earlier starters when testing, and most earlier starters being successful learners. The language aptitude test (language analysis) in the current study requires participants to make a deduction with explicit knowledge, which measures learners' language inference ability. The sub-group learners' aptitude has no significant difference in the current study may result from participants' similar chronological age and cognition when testing, participants having enough time to finish the test and monitor their performance, which to some extent supports the stability of language aptitude that is difficult to be trained or adjusted and insusceptible to environmental change.

Second, language aptitude moderates the UTGJT and MK test, especially in Primary one and Primary three. The three grammar tests in the current study require participants to focus on language correctness and make a judgment. Generally, tasks, focusing on language form and correctness, mainly measure learners' analysis ability and metalinguistic knowledge. The current

study shows that language aptitude cannot predict TGJT in which participants finish the task within the given time and have fewer opportunities to effectively use language analysis ability. Under this condition, second language learners may meet processing difficulties which provide more influence on older learners (McDonald, 2000, 2006). However, UTGJT allows participants to have enough time to think, consciously control testing performance and use explicit knowledge to process language. The UTGJT requires participants to make a judgment on sentence correctness, while the MK test requires participants to identify, correct grammar errors, and explain the violation of grammar rules, which requires participants to highly use explicit consciousness to focus on language forms. Higher language aptitude learners outperform lower language aptitude learners on these tasks because higher language aptitude learners can better use language analysis ability and more effectively focus on sentence form. In the current study, the language aptitude of participants in Secondary one has no significant correlation with three grammar tests, because this group has fewer samples, which results in that there is smaller variance in learners' language aptitude and grammar test score, therefore, there is a higher similarity among participants, and learners with different L2 proficiency can't be distinguished. According to the Ministry of Education of the People's Republic of China, students should start to learn English in Primary three from 2001 (Ministry of Education, 2001). Secondary one sub-group started to learn English in grade one in junior middle school (2008-2009). Meanwhile, participants who had taken extra-curriculum classes or had studied or lived abroad were removed, so participants meeting the current requirement were fewer. Further analysis shows that language aptitude of Primary three sub-group is significantly correlated with UTGJT, and language aptitude of Primary three sub-group and language aptitude of Primary one sub-group are significantly correlated respectively with the MK test, which suggests that language aptitude moderates younger learners, and may show that there is a certain continuity in second language learning ability (Skehan, 2015), and the younger the AO of learners is, the stronger the L2 grammar attainment relies on language aptitude (Li, 2015).

The current results are different from that in DeKeyser (2000) and Huang (2014). DeKeyser et al. (2010) suggest that younger and older starters have different learning mechanisms. Younger learners master more knowledge, but they rely on less language aptitude; however, older learners must depend on more language aptitude to promote learning, which is consistent with the Fundamental Difference Hypothesis (DeKeyser, 2000). The differences between the current study and the above studies may result from various research methods. (1) The number of earlier starters in DeKeyser (2000), and DeKeyser et al. (2010) is 15 and 20 respectively. In these two studies, the language aptitude of earlier starters is not significantly correlated with GJT, which may result from fewer samples so that the variance of learners' GJT is small. In fact, the earlier starters' accuracy rate of GJT is above 90% in DeKeyser (2000). (2) Both DeKeyser (2000) and DeKeyser et al. (2010) suggest that although adults' implicit learning ability decreases continuously with the increase of AO, their explicit learning mechanism can make up for this. The current study shows that the language aptitude of the Primary-three sub-group is significantly correlated with UTGJT, while the language aptitude of Primary one and Primary three subgroups is significantly correlated with the MK test, which suggests that the younger

learners can also rely on explicit analysis ability to get better performance in grammar tests. This may be consistent with Paradis's description (2009) that only learners who start to learn a second language before 4-5 years old can learn a second language with the implicit mechanism. (3) Both Granena (2012) and Granena and Long (2013b) used LLAMA to measure aptitude which is a computer-based test and language-independent, while the current study used a Chinese-version language aptitude test developed for Chinese foreign language learners.

Although the result in the current study is similar to that of Granena (2013a, 2014) and Abrahamsson and Hyltenstam (2008), they applied different language structures and cognition measurements. In Abrahamsson and Hyltenstam (2008), they used both written and auditory GJTs to measure L2 grammar knowledge, their participants were second language learners with advanced second language proficiency, and the sentences in GJT were composed of longer sentences with complex semantics. That is, high aptitude learners have an advantage in distinguishing grammar errors with processing and analysis ability. Granena (2014) found that the language aptitude of earlier starters only plays moderating role in grammatical sentences in untimed auditory GJT. The current study found that language aptitude is only significantly correlated with UTGJT and MK, that is, language aptitude is not significantly correlated with language abilities with implicit knowledge. Another explanation may be that other types of language aptitude are correlated with the automatic use of second language knowledge. For example, Granena (2013a) found that implicit aptitude was significant with implicit grammar knowledge.

Third, in stimulus tasks, almost all the higher language aptitude learners outperform lower language aptitude learners in ungrammatical sentences, and all the higher language aptitude learners in three sub-groups perform better in ungrammatical sentences than in grammatical sentences. Because ungrammatical sentences measure explicit grammar knowledge (Ellis et al., 2009), higher language aptitude learners perform better in tasks using explicit knowledge, and language aptitude has a greater influence on the tasks measuring explicit grammar knowledge, which is similar to Granena (2013a), while in Granena (2013a), the participants are higher proficient second language learners and have started to learn the second language when being adults. In UTGJT and MK, learners have enough time to control their second language knowledge, analyze sentence structure, and monitor their performance. Under this condition, higher language aptitude learners can use their analytical ability and metalinguistic knowledge better, which is also needed in the current grammar structures. The grammar structures in the MK test are affected most strongly, in which higher language aptitude learners perform better in most of the structures, but higher language aptitude learners and lower language aptitude learners have significant differences only in some structures, and their scoring trend in different grammar structure is similar, which suggests to some extent that there is no guarantee of absolute advantage for higher language aptitude learners. However, so far, only a few studies have explored the influence of language aptitude on L2 grammar structures, such as Granena (2012), Farshi & Tavakoli (2021), therefore, further studies could be conducted in this field that can provide more information about the learning process and learning attainment in different learning stages.

Conclusion

The current study has revealed that there is a close relationship between the language aptitude of learners with different ages of onset and the influence of language aptitude on various L2 grammar tests. It was also found that second language learners' language aptitude tends to increase as their AO increases, but there is no significant difference between them. It was also found that language aptitude moderates the untimed grammar judgment test and the metalinguistic knowledge test, especially in earlier starters. Finally, language aptitude only predicts ungrammatical sentences, and higher language aptitude learners perform better than lower language aptitude learners.

As such, the study could be used as a reference for future research in this field, and also lends implications and support to developing novel language aptitude tests for Chinese foreign language learners. That said, there are some limitations in the current study. For example, the current study only used the "words in sentences" in FLAT-C to measure explicit language aptitude. Future studies can be conducted to explore the relationship between implicit language aptitude and L2 grammar attainment, as some recent studies probing implicit language aptitude have provided strong evidence (Li & Qian, 2021; Suzuki, 2021). Then, the sample size in the current study was quite small, so future studies can enlarge it by involving more participants. In addition, the current study just used the GJT and the MK test to measure grammar knowledge which may not reflect a full picture of participants' grammar knowledge and second language proficiency, so future studies can use other language aptitude tests or other components of language aptitude tests and more second language tests to further explore the influence of language aptitude on L2 grammar attainment and second language proficiency, as some studies have provided strong evidence (Sparks, 1995; Sparks et al, 2011). Finally, the current study just focused on the impact of language aptitude on the second language learners' grammar attainment, future studies can investigate the effect of language aptitude on the learning process, as language aptitude may play differential roles at distinct stages of SLA (Wen & Skehan, 2021).

References

Abrahamsson, N., & Hyltenstam, K. (2008). The Robustness of Aptitude Effects in Near-native Second Language Acquisition. *Studies in Second Language Acquisition, 30*(4), 481-509. https://doi.org/10.1017/S027226310808073X

Bokander, L., & Bylund, E. (2020). Probing the Internal Validity of the LLAMA Language Aptitude Tests. *Language Learning, 70*(1), 11-47. https://doi.org/10.1111/lang.12368

Bylund, E., Abrahamsson, N., & Hyltenstam, K. (2012). Does First Language Maintenance Hamper Nativelikeness in a Second Language? A Study of Ultimate Attainment in Early Bilingual. *Studies in Second Language Acquisition, 34*(2), 215-241. https://doi.org/10.1017/S0272263112000034

Carroll, J. B. (1965). The prediction of success in intensive foreign language training. In R. Glaser (Ed.), *Training, Research and Education*. New York: Wiley.

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*. Rowley, MA: Newbury House.

Chen, Y. P., & Chen, X. (2019). Explicit and Implicit Language Aptitudes in Second Language Acquisition. *Foreign Language Teaching and Reaching, 51*(5), 723-734. http://www.fltr.ac.cn/WKG/WebPublication/paperDigest.aspx?paperID=db8f7bed-3dda-44f5-9dfc-0d6477f08f37

DeKeyser, R. M. (2000). The Robustness of Critical Period Effects in Second Language Acquisition. *Studies in*

Second Language Acquisition, 22(4), 499-533. https://doi.org/10.1017/S0272263100004022

DeKeyser, R. M., Alfi-Shabtay, I., & Ravid, D. (2010). Cross-linguistic Evidence for the Nature of Age-effects in Second Language Acquisition. *Applied Psycholinguistics, 31*(3), 413-438. https://doi.org/10.1017/S0142716410000056

Dörnyei, Z., & S. Ryan. (2015). *The Psychology of the Language Learner Revisited.* New York & London: Routledge.

Ellis, R., Loewen, S., Elder, C., Erlarn, R., Philp, J., & Reinders, H. (2009). *Implicit and Explicit Knowledge in Second Language Learning, Testing and Teaching*. Bristol: Multilingual Matters. https://doi.org/10.1017/S0272263110000586

Farshi, N., & Tavakoli, M. (2021). Effects of Differences in Language Aptitude on Learning Grammatical Collocations under Elaborated Input Conditions. *Language Teaching Research, 25*(3), 476-499. https://doi.org/10.1177/1362168819858443

Granena, G. (2012). *Age Differences and Cognitive Aptitudes for Implicit and Explicit Learning in Ultimate Second Language Attainment*. Unpublished doctoral dissertation. Maryland: University of Maryland. http://hdl.handle.net/1903/12743

Granena, G. (2013a). Individual Differences in Sequence Learning Ability and Second Language Acquisition in Early Childhood and Adulthood. *Language Learning, 63*(4), 665-703. https://doi.org/10.1111/lang.12018

Granena, G. (2013b). Reexamining the robustness of aptitude in second language acquisition. In G. Granena & M. Long (Eds.), *Sensitive Periods, Language Aptitude, and Ultimate L2 Attainment*. Amsterdam: John Benjamins Publishing Company. https://doi.org/10.1075/lllt.35

Granena, G. (2014). Language Aptitude and Long-term Achievement in Early Childhood L2 Learners. *Applied Linguistics, 35*(4), 483-503. https://doi.org/10.1093/applin/amu013

Granena, G., & Long, M. (2013a). *Sensitive Periods, Language Aptitude, and Ultimate L2 Attainment*. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Granena, G., & Long, M. (2013b). Age of Onset, Length of Residence, Language Aptitude, and Ultimate L2 Attainment in Three Linguistic Domains. *Second Language Research, 29*(3), 311-343. https://doi.org/10.1177/0267658312461497

Grigorenko, E. L., Sternberg, R. J., & Ehrman, M. E. (2000). A Theory-based Approach to the Measurement of Foreign Language Learning Ability: The Canal-F Theory and Test. *The Modern Language Journal, 84*(3), 390-405. https://psycnet.apa.org/doi/10.1111/0026-7902.00076

Harley, B., & Hart, D. (1997). Language Aptitude and Second Language Proficiency in Classroom Learners of Different Starting Ages. *Studies in Second Language Acquisition, 19*(3), 379-400. https://doi.org/10.1017/S0272263197003045

Harley, B., & Hart, D. (2002). Age, aptitude and second language learning on a bilingual exchange. In P. Robinson (Ed.), *Individual Differences and Instructed Language Learning*. Amsterdam: John Benjamins Publishing Company. https://doi.org/10.1075/lllt.2

Huang, B. H. (2014). The Effects of Age on Second Language Grammar and Speech Production. *Journal of Psycholinguistic Research, 43*(4), 397-420. https://doi.org/10.1007/s10936-013-9261-7

Larson-Hall, J. (2008). Weighing the Benefits of Studying a Foreign Language at a Younger Starting Age in a Minimal Input Situation. *Second Language Research, 24*(1), 35-63. https://doi.org/10.1177/0267658307082981

Li, L. R., & Luo, S. Q. (2019). Development and preliminary validation of a foreign language aptitude test for Chinese learners of foreign languages**.** In Z. Wen, P. Skehan, A. Biedron, S. Li & R. Sparks (Eds.), *Language Aptitude: Advancing Theory, Testing, Research and Practice Rethinking*. London: Routledge, SLA Research Series.

Li, S. F. (2015). The Associations between Language Aptitude and Second Language Grammar Acquisition: A Meta-analytic Review of Five Decades of Research. *Applied Linguistics, 36*(3), 385-408. https://doi:10.1093/applin/amu054

Li, S. F., Ellis, R. & Zhu Y. (2019). The Associations between Cognitive Ability and L2 Development under Five Different Instructional Conditions. *Applied Psycholinguistics, 40*(3), 693-722. https://doi.org/10.1017/S0142716418000796

Li., S., Hiver, P., & Papi, M. (2022). *The Routledge Handbook of Second Language Acquisition and Individual Differences.* New York: Routledge.

Li, S. F., & Qian J. C. (2021). Exploring Syntactic Priming as a Measure of Implicit Language Aptitude. *Studies in Second Language Acquisition*, 1-32. http://dx.doi.org/10.1017/s0272263120000698

Li, H., Ma, L., & Zhang, X. H. (2019). On the Relationship Between Age of Onset, Language Aptitude and L2 Grammar Attainment. *Modern Foreign Languages, 42*(4), 527-539. https://xdwy.cbpt.cnki.net/WKG/WebPublication/paperDigest.aspx?paperID=9ce8787d-0850-4d20-b974-7946af98d175

McDonald, J. L. (2000). Grammaticality Judgments in a Second Language: Influences of Age of Acquisition and Native Language. *Applied Psycholinguistics, 21*(3), 395-423. http://dx.doi.org/10.1017/S0142716400003064

McDonald, J. L. (2006). Beyond the Critical Period: Processing-based Explanations for Poor Grammaticality Judgment Performance by Late Second Language Learners. *Journal of Memory and Language, 55*(3), 381-401. https://doi.org/10.1016/j.jml.2006.06.006

Ministry of Education, 2001. http://www.moe.gov.cn/s78/A26/jces_left/moe_714/tnull_665.html

Nation, I. S. P. (1990). *Teaching and Learning Vocabulary*. New York: Newbury House.

Paradis, M. (2009). *Declarative and Procedural Determinants of Second Language*. Amsterdam: Benjamins. https://doi.org/10.1075/sibil.40

Qureshi, M. A. (2016). A Meta-analysis: Age and Second Language Grammar Acquisition. *System*, (60), 147-160. https://doi.org/10.1016/j.system.2016.06.001

Robinson, P. (2007). Aptitudes, abilities, contexts, and practice. In R. M. DeKeyser (Ed.), *Practice in a Second Language: Perspectives from Applied Linguistics and Cognitive Psychology* (pp. 256-286). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511667275.015

Roehr-Brackin, K., & Tellier, A. (2019). The Role of Language-analytic Ability in Children's Instructed Second Language Learning. *Studies in Second Language Acquisition, 41*(5), 1111-1131. https://doi.org/10.1017/S0272263119000214

Skehan, P. (1998). *A Cognitive Approach to Language Learning*. Oxford: Oxford University Press.

Skehan, P. (2015). Foreign Language Aptitude and Its Relationship with Grammar: A Critical Overview. *Applied Linguistics, 36*(3), 367-384. https://doi.org/10.1093/applin/amu072

Sparks, R. (1995). Examining the Linguistic Coding Differences Hypothesis to Explain Individual Differences in Foreign Language Learning. *Annals of Dyslexia*, 45, 187-214. https://doi.org/10.1007/BF02648218

Sparks, R., & Ganschow, L. (2001). Aptitude for Learning a Foreign Language. *Annual Review of Applied Linguistics*, 21, 90-111. http://dx.doi.org/10.1017/S026719050100006X

Sparks, R., Patton, J., Ganschow, L., & Humbach, N. (2011). Subcomponents of Second Language Aptitude and Second Language Proficiency. *The Modern Language Journal, 95*(2), 253-273. http://dx.doi.org/10.1111/j.1540-4781.2011.01176.x

Suzuki, Y. (2021). Probing the Construct Validity of LLAMA-D as a Measure of Implicit Learning Aptitude: Incidental Instructions, Confidence Rating, and Reaction Time. *Studies in Second Language Acquisition, 43*(3), 663-676. https://doi.org/10.1017/S0272263120000704

Wen, Z., & Skehan, P. (2021). Stages of Acquisition and the P/E Model of Working Memory: Complementary or Contrasting Approaches to Foreign Language Aptitude? *Annual Review of Applied Linguistics*, *41*(2), 6-24. http://dx.doi.org/10.1017/S0267190521000015

Wen, Z. S., Skehan, P., Biedroń, A. Li, S. F., & Sprarks, R. L. (2019). *Language Aptitude: Advancing Theory, Testing, Research and Practice*. New York and London: Routledge. https://doi.org/10.4324/9781315122021

Ethics Declarations

Competing Interests

No, there are no conflicting interests.

Rights and Permissions
Open Access