Young Learners' Bilingual Status and Cognitive Development in Foreign Language Aptitude Testing

Author 1:
Maria-del-Mar Suárez
Universitat de Barcelona

Author 2:
Charles W. Stansfield
Language Learning and Testing Foundation

Abstract

Young learners' L1s preference, cognitive development and bilingual status might influence their performance on language aptitude tests, particularly if these are language-dependent. The objective of this study was to test the validity, reliability and consistency across populations of two such tests: the Modern Language Aptitude Test-Elementary in Catalan (MLAT-EC) and in Spanish (MLAT-ES). 629 bilingual students from grades 3 to 7 took the MLAT-ES and the MLAT-EC for test comparison in a counterbalanced order. The results show that their performance on both tests presented hardly any significant differences considering students' L1 preference (Catalan, Spanish or no preference). In addition, these bilingual examinees outperformed the predominantly monolingual samples in the MLAT-ES norming study. The same score patterns related to young learner cognitive development stages were found across test versions. These results reinforce the confidence in the validity of the MLAT-E adaptations and support the hypothesis that bilingualism results in greater aptitude.
Keywords: language aptitude; bilingualism; language aptitude testing; test validity; young learners

1. Introduction

Foreign language aptitude has been defined as the specific talent associated with learning a foreign or second language (Carroll, 1981; Skehan, 2002). This is a traditional view on aptitude rooted in educational psychology. In recent years, what is meant by aptitude has been redefined considering aspects of cognitive psychology and cognitive neuroscience and second language acquisition (Wen et al., 2017). Thus, it has been found to correlate with working memory (e.g. Li, 2015) as well as to issues of explicit and implicit learning (e.g. Granena & Long, 2013). Aptitude has also been closely related to the concept of time, so it is also conceptualized as the "prediction of

how well, relative to other individuals, an individual can learn a foreign language in a given amount of time and under given conditions" (Carroll & Sapon, 2002, p. 23).

Carroll's aptitude model includes four major abilities, i.e. phonemic coding ability, grammatical sensitivity, inductive language learning ability and rote learning ability, although later on Skehan (1989) proposed a three-component model consisting of auditory ability (the former phonemic coding ability), linguistic ability (merging grammatical sensitivity and inductive language learning ability), and memory.

Different tests are used to measure aptitude depending on the model adopted. One of the recent tests often used in aptitude studies, given its free availability, is the LLAMA test (Meara, 2005), which is also considered language-neutral, although LLAMA E, the sound symbol correspondence test, approached significant differences in that sense for speakers of logographic languages (Rogers et al., 2017). Measuring aptitude with language-based tests has been an issue of debate as performance on the aptitude test might also be determined by the examinee's mastery of their L1. Indeed, individual differences play a major role in one's L1 rate of acquisition (Wells, 1985, 1986). Thus, if one's L1 is still developing when taking a foreign language aptitude test, the scores will probably be affected by one's L1 proficiency, and in the case of young learners, by one's L1 developmental stage, but also one's cognitive developmental stage as well (Suárez & Muñoz, 2011). Furthermore, Skehan (1986; 1988, as mentioned in Skehan 1989) found clear connections between syntactic first language development and foreign language analytic ability. However, it remains unknown how simultaneous bilingual learners might perform on language-based aptitude tests and if this performance is equally affected by their language preference, and by their developmental stage.

Likewise, bilingualism does not tend to negatively affect foreign language learning as this prior experience of learning more than one language leads to enhanced cognitive and metalinguistic abilities (e.g., Bialystok et al., 2004). Consequently, it might also be the case that simultaneous bilingual learners might outperform monolinguals on the same aptitude tests given their bilingual status.

The aim of this study is thus twofold: first, it aims to examine the comparability in terms of validity of two aptitude tests for young simultaneous bilingual learners depending on their first language preference (Catalan and Spanish or both), and to see if there is greater aptitude among simultaneous bilinguals, at least for this language pair, as compared to their L1 Spanish prominently monolingual peers.

2. Literature review

2.1 L1 and cognitive development, formal education, and foreign language aptitude

While everyone should be able to learn a language given the right conditions (e.g., Bley-Vroman, 2009; Chomsky, 1975), L2 acquisition is shaped by individual differences, such as aptitude, as well as one's L1 developmental stage, L1 mastery and ultimate attainment especially for adult learners (Dąbrowska, 2018). As for young learners, it has also been found that cognitive developmental stages coincide with young learners' performance on aptitude tests (Suárez & Muñoz, 2011). It was then hypothesized that young learners' literacy skills, along with their cognitive development

and their L1 development could also be reflected in their aptitude scores, which showed significant statistical differences between grades 3 and 4, coinciding with the end of the Piaget's (1964) preoperational stage and beginning of the concrete operational stage, as well as a plateau in the results between grades 6 and 7, when stepping into the formal operational stage.

Regarding one's L1 development and how L2 development relates to it, as Ellis (1996, p. 94), puts it, "the same patterns [in L1 acquisition] are found in SLA. Novice language learners are bound up in the orthographic and phonological aspects of vocabulary", which are, at the same time, quite dependent on the kind of literacy instruction received. Indeed, language aptitude has also shown to be relevant for FL and L2 learning, especially in formal education settings (Ehrman & Oxford, 1995; Li, 2015; Sparks et al., 2011). Correlations have also been found between early L1 development performance in children and their later L2 aptitude scores and L2 proficiency (e.g. Skehan, 1986; Sparks et al., 2012).

Sparks and Ganschow's seminal work, with their Linguistic Coding Deficit/Differences Hypothesis (Ganschow & Sparks, 1995; Sparks & Ganschow, 1993; Sparks et al., 1997) concluded that one's L1 development, especially in phonological processing, certainly influences L2 development as well as FL aptitude scores, on the one hand, and, on the other hand, that long-term instruction plays a significant role in the mix. These differences might also affect other learning areas due to the essential role that phonological processing, decoding and literacy play in learning practice.

2.2 FL aptitude and sociocultural factors

Education and literacy training might also be a predictor of L1 abilities (Chipere, 2003) as well as of L2 ones (Hakuta et al., 2003; Tarone et al., 2013). While language aptitude has often been conceived as innate and untrainable (e.g. Carroll, 1981; see Singleton, 2017, for an overview), correlational studies leave it unclear whether these effects are due to language aptitude only or to other factors such as language learning experience, awareness or educational level (Rogers et al., 2016, 2017; Sáfar & Kormos, 2008; Singleton, 2014). In fact, Larsen-Freeman and Long (1991) challenged the idea of innateness in language aptitude and, instead, attributed it to classroom experience. The question is, if aptitude is multicomponential, whether some components are more beneficial to language learning than others (Kormos, 2013).

Thus, the language learning experience might particularly affect FL aptitude test scores, for several reasons, such as familiarity with tests, literacy experience, memory training, etc. (Ganschow & Sparks, 1995; Sáfár & Kormos, 2008; Sparks et al., 1997), with learners with greater previous language learning experience outperforming untrained ones (e.g. Grigorenko et al., 2000). Special care needs to be taken, therefore, when interpreting scores on tests such as the MLAT or others, where it has been found that prior FL learning is a factor that increases scores (Ganschow & Sparks, 1995; Sparks et al., 1997).

Larger data corpora point in the direction that measures of aptitude can predict achievement in children learning in instructed settings (Kiss 2009; Kiss & Nikolov 2005; Muñoz 2014; Suárez & Gesa, 2019, 2022). Therefore, if formal instruction results in more favorable aptitude scores, the learners' socioeconomic status or the type of formal instruction received will inevitably affect their aptitude scores and, perhaps, their success in foreign language learning.

2.3 Bilingualism, language learning experience and FL aptitude

Bi/Multilingualism accounts for cognitive advantages such as brain flexibility, superior analysis and control abilities and metalinguistic abilities (e.g., Bialystok & Martin, 2004; Costa et al., 2008; Poarch & van Hell 2012), as well as further language learning, probably due to the transfer of certain automatized basic skills consolidated in previous language learning experiences, such as auditory- and pattern recognition skills and superior auditory memory (Nation & McLaughlin, 1986) as well as strengthened general non-linguistic abilities, such as selective attention skills (Bialystok, 2017). However, it has also been found that the bilingual's advantages are less potent without explicit formal training (Hopp et al., 2019).

Research to date has not yet reached clear results as to how bi/multilingualism may affect cognitive abilities including aptitude, considering its supposed trainability. While aptitude might be shaped by one's innate abilities, experience in language learning might also enhance language learnability, thus causing a domino effect (Biedron & Véliz-Campos, 2021). Indeed, Turker et al. (2018) found that, while high-aptitude students outperformed lower-aptitude students in all tasks, and considering that aptitude is an innate, yet developing capacity which remains stable in the long term, the number of languages spoken by these bi/multilinguals did not, however, correlate with aptitude, which goes against previous results (Eisenstein, 1980; Sternberg & Grigorenko, 2002). This finding also contradicts several studies that have shown that polyglots and hyperpolyglots present high aptitude scores, often along with high levels of motivation and metalinguistic awareness (e.g. Hyltenstam, 2018, 2021).

What seems to be certain is that the more languages one has learned, the more likely one is to include another language in one's repertoire, without this meaning that no effort is required (Hyltenstam, 2021).

The present study, therefore, intends to cover the gaps in the literature by offering cross-sectional results on how young learners' L1s preference and bilingual status might affect their aptitude score on two language-dependent tests.

3. The present study

3.1 Aims and research questions

The aim of this study is to examine the influence of L1 preference in the subjects' performance on two different aptitude tests (the MLAT-ES, Modern Language Aptitude Test – Elementary in Spanish, and the MLAT-EC, in Catalan), and to explore if the subjects' bilingualism involves an advantage in comparison with the mainly monolingual status of the subjects in the MLAT-ES norming study (Stansfield et al., 2004).

The three resulting research questions are:

1. Is the subjects' performance on the MLAT-ES affected by their L1 preference within the same grade and across grades?
2. Is the subjects' performance on the MLAT-EC affected by their L1 preference within the same grade and across grades?
3. Is there any advantage for Catalan/Spanish bilingual speakers as compared to those, mainly monolingual students who participated in the MLAT-ES norming study?

3.2 Participants and procedure

A convenience sample of 629 Catalan/Spanish bilingual students between grades 3 to 7 (age range 8.3 – 14.9) participated in this study. They were all receiving classes in English as a Foreign Language (EFL) at school and some of them were also taking EFL extracurricular classes. They were schooled in Catalonia, where the main language of instruction is Catalan (except for the Spanish and English classes), although the main language spoken in many other social exchanges is Spanish. They all took both the MLAT-ES and the MLAT-EC. The administration procedure was counterbalanced, that is, about half of the test takers (N=325) took the MLAT-E in Spanish first and 5 months later they took it in Catalan (Group 1) while the other half (N=304) took the tests in reverse order (Group 2). Only the first test score was considered for this study, as there were test-retest effects in the second sitting from three to five months after the first administration, even if the aptitude test of the second sitting was in a different yet closely-related language (Suárez, 2010, 2022). Both groups also took a dictation and a cloze test along with other EFL criterion measures and oral performances in the days following the administration of the first aptitude test (see further details in Suárez, 2010).

Table 1. Participants' age, grade and group.

	All subjects			**Group 1 (ES – EC)**			**Group 2 (EC-ES)**		
Grade	N	Age	SD	N	Age	SD	N	Age	SD
3	123	8.8	.52	66	8.8	.33	57	8.8	.66
4	137	9.8	.43	75	9.9	.48	62	9.7	.36
5	118	10.8	.33	57	10.9	.33	61	10.8	.34
6	120	11.8	.3	60	11.7	.34	60	11.9	.33
7	131	12.9	.45	67	12.9	.45	64	12.8	.44

The participants in Group 1 and Group 2 were all childhood bilingual speakers of Catalan and Spanish speakers with different preferences as to their dominant language, as shown in Table 2.

Table 2. Students self-report of their L1 dominant language: Catalan, Spanish or balanced bilinguals.

	All subjects			Group 1 (ES – EC)			Group 2 (EC-ES)		
Grade	Catalan	Spanish	Balanced	Catalan	Spanish	Balanced	Catalan	Spanish	Balanced
3	46	47	30	25	25	16	21	22	14
4	51	46	40	21	30	24	30	16	16
5	45	40	33	16	24	17	29	16	16
6	45	37	38	19	21	20	26	16	18
7	48	52	31	7	35	25	41	17	6
All	235	222	172	88	135	102	147	87	70

Language preference was coded using the answers to the questions as to which language participants preferred when speaking with their mother, father, brothers, sisters, rest of family, and friends, as well as which language they felt more comfortable with when speaking in any communicative situation. Due to the magnitude of the main study, which involved a massive data collection of measures in English as a foreign language (Suárez, 2010), no objective measurement of L1 proficiency or mastery could be collected. As these participants were rather young, they were asked the same questions twice (at the beginning of the data collection and at the end) to

verify the consistency in their answers. Despite their self-reported L1 preference for one or another language or both indistinctly, they were all considered fully proficient in both languages as they were all simultaneous bilinguals, i.e., they had learned both languages in early childhood. However, this should not be confused with their being equally proficient in both L1s, since not only language dominance but also cross-linguistic influence and frequency of use might affect their performance in either language and, by extension, in the aptitude tests in this study. Indeed, Catalan is supposed to be the vehicular language at school for all subjects except for Spanish and English according to the law 12/2009 d'Educació, but the metropolitan area, where most of these data were collected, presents a more frequent use of Spanish over Catalan, at least for the 2003-2018 period (Institut d'Estadística de Catalunya, 2018).

The data available in the *Manual* were used to compare the validity of the MLAT-ES and MLAT-EC when administered in a bilingual community as compared to its usage in a mainly monolingual kind of environment, as was the one in the MLAT-ES norming study (Stansfield et al., 2004). The number of participants in the MLAT-ES norming study per grade was Grade 3, n = 207; Grade 4, n = 206; Grade 5, n = 289; Grade 6, n = 306; Grade 7, n = 178.

3.3 Instruments and procedure

3.3.1 The MLAT-ES

The MLAT-ES is the Spanish version of the original MLAT-E, created by Carroll and Sapon and normed by the Psychological Corporation in 1967. Both tests were designed for the 8-13 age range. The MLAT-ES was developed by Stansfield and Reed in 2003, who first designed a theoretical framework to develop adaptations of the MLAT (the test version for adults; Carroll & Sapon, 1959) for test takers whose L1 is not English, and then started working on the framework for children. The MLAT-ES is for Spanish speakers who live in the United States or in other countries where Spanish is spoken as the L1. The Spanish variety used in the test is Latin American Spanish, which could cause some distraction for test takers in Spain, who speak a different variety of Spanish.

The original test contained 165 items and was administered to 235 students in Costa Rica. After running an item analysis for item discrimination ability, difficulty and reliability, 32 items were removed so the test ended up having a total of 133 items. This last version was administered to 1186 students, 745 of whom lived in South America (Colombia, Costa Rica and Mexico) and to 441 living in Spain, of whom 214 were Spanish monolinguals living in Madrid while the remaining 227 were bilingual Catalan/Spanish speakers from Catalonia. But for the Catalan students, nearly all of the students participating in the norming study were monolingual students, or at least clearly Spanish dominant, and with an age-appropriate level of mastery of literacy and of Spanish as an L1. Item difficulty, item discrimination, and test reliability analyses were run again. More items where then removed, with the final test version having 123 items. The tests results were then recalculated considering only the items in the final version and a predictive validity study was run using criterion variables including teachers' ratings of their students' aptitude for foreign language learning. The final version was published in 2005 (Stansfield et al., 2004). It contained items that showed some differential functioning in the bilingual Catalan/Spanish community, as extensively explained in Suárez (2010).

Like the MLAT-E, the MLAT-ES is a paper and pencil test that takes approximately one hour to complete. Three of its four parts are derived from the MLAT. Each part is meant to focus on an ability, while also tapping into one or more other abilities more weakly, as indicated below.

Part 1: Hidden Words (*Palabras Ocultas*). This part corresponds to the MLAT's Spelling Clues but using easier vocabulary. It measures L1 vocabulary as well as sound-symbol association. (30 items)

Part2: Words in Sentences (*Palabras que se Corresponden*). This part is meant to measure grammatical sensitivity without using formal grammatical terms. Test takers are expected to find the word in a sentence that performs the same function as the one in capital letters in the item's stem. (30 items)

Part 3. Rhyming Words (*Palabras que Riman*). This part does not emanate from the MLAT. It taps into the ability to hear speech sounds while selecting words that rhyme. (38 items)

Part 4. Learning Numbers (*Números en Otro Idioma*). The test taker must learn six numbers and how to combine them in an artificial language. This part taps into rote memory learning as well as vocabulary learning and the ability to form and remember associations between speech sounds. (25 items)

All items were scored in the same way. That is, one point for right answer, zero points for wrong or no answer. An Excel spreadsheet specially provided by the test creators (Stansfield et al., 2004) was used to score the tests. Although some cross-cultural and cross-linguistic issues were detected when running the item analysis with this bilingual population (Suárez, 2022), all the items were accounted for in the test total score for comparability purposes with the main norming study.

The content validity of this test for the participants in this study was calculated using several analyses: index of facility, discrimination power of items, point-biserial correlation coefficient and the corrected item-total correlation. The construct validity was also determined using correlations with several criterion variables. Nearly all the item results yielded positive results, and in the case of those items that didn't, the results were mainly due to the Spanish language variety used in the test. The only weak point in the test was the construct validity when calculated using the correlations with the teachers' ratings of the students' language aptitude, and these ratings were themselves not considered accurate and reliable.

3.3.2 The MLAT-EC

The MLAT-EC (Suárez, 2010) stems from the Spanish version of the MLAT-E, the MLAT-ES. It contains the same parts as the MLAT-ES. Item discrimination, item difficulty, and reliability analysis were run with all the items originally in the test. In the end, the test yielded 122 items (Part 2 contained 29 items instead of 30). After removing the faulty items or those that did not come up to the standard of others, the MLAT-EC turned out to be slightly more difficult than the MLAT-ES across grades (Suárez, 2010.) not only in the raw scores but also in the distribution of percentiles. In addition, it presents a similar pattern in the mean scores across grades, with the Catalan version starting with a lower mean score in grade 3, catching up with those on the MLAT-ES at grades 4 and 5, and having a lower mean score once again in grades 6 and 7, although the difference never reached statistical significance between the two tests. What is remarkable, though, is that both the scores on the MLAT-ES and the MLAT-EC seem to be clearly related to the participants' cognitive development (Suárez, 2011), proven by the substantial differences in the scores between grades 3 and 4 and the plateau of scores in grades 6 and 7.

For the sake of consistency, the decision was also made to keep the same font size and type, and layout format of the MLAT-ES for test comparability purposes, even though some literacy

difficulties and confusions due to the layout were detected when administering the tests in the lower grades, mainly caused by the font type and size used, Times New Roman 12, which may have been smaller than optimal. Further and extensive details regarding item discrimination, difficulty and reliability of this version of the test can be found in Suárez (2010). In that study, given the relatively balanced distribution of the population regarding their language preference, the participants were all placed in one pool. The tests were scored following the same system as the one used for the MLAT-ES, that is, using an Excel spreadsheet.

Running the same analyses as those applied to the MLAT-ES, after removing the faulty items, the MLAT-EC was shown to be valid and reliable against the EFL criterion measures used.

4. Analysis

Each part of the MLAT-ES and MLAT-EC was analyzed separately, since they are meant to measure different abilities, although it is usually the total score that tends to consistently show the highest correlations with criterion measures (Li, 2016). All parts are considered to tap into the same abilities regardless of the language of the test, at least in this case, as both languages at work are Romance languages with very similar overall characteristics (although with their particularities, too).

Since the distribution of the scores was not normal across all grades, according to Kolmogorov-Smirnov results, and the number of participants was also limited in certain groups, non-parametric tests of statistical significance were used. Consequently, medians, instead of means, were reported. For research questions 1 and 2, which compared the performance of the three groups on each part depending on their preferred language (Spanish, Catalan or both equal), Kruskal Wallis tests were run across grades for both the MLAT-ES and the MLAT-EC, in order to compare the influence of L1 preference in the scores across the same grade as well as to compare the evolution of scores across grades determining one's L1 preference as the dependent variable. Mann-Whitney U-tests were then run when significant differences either across grades or L1 preference were found. To answer research question 3, that is, to see if there was an advantage for the bilingual population as compared to the mainly monolingual one in the norming study, the data provided in the manual were used and contrasted with the results obtained by the participants in this present study. To contrast the results obtained by Group 1 with those of the norming study, Cohen's *d* formula for t-tests was applied to calculate how many standard deviations lie between the two means compared and the relative effect size. The percentage of change and its direction were also calculated.

The criteria to qualify Cohen's *d* size from t-tests was the following: negligible effect (>= -0.15 and <.15); small effect (>=.15 and <.40); medium effect (>=.40 and <.75); large effect (>=.75 and <1.10); very large effect (>=1.10 and <1.45); huge effect >1.45.

The ranges applied to describe change percentage were: huge decrease <-75; very large decrease (<=-50 and >-75); large decrease (<=-30 and >-50); medium decrease (<=-15 and >-30); small decrease (<=-5 and >-15); negligible change (>= -5 and <5); small increase (>=5 and <15); medium increase (>=15 and <30); large increase (>=30 and <50); very large increase (>=50 and <75); huge increase >75.

5. Results

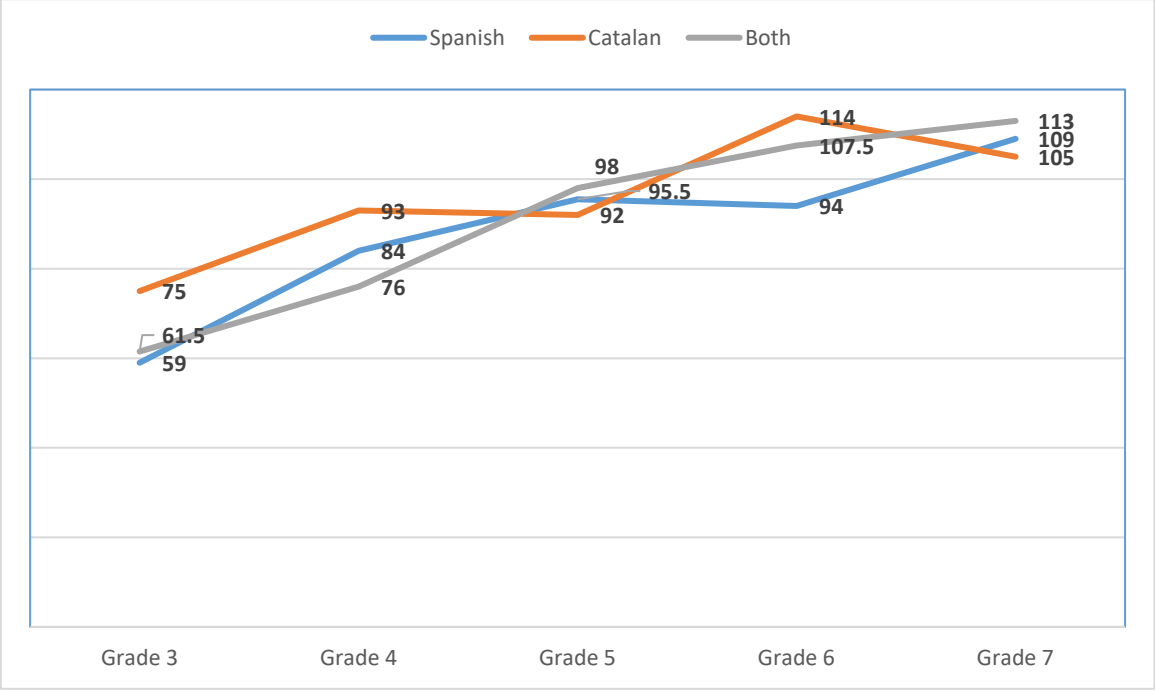
5.1 MLAT-ES depending on L1s preference per grade

Descriptive results computed for the MLAT-ES parts and total score depending on one's L1 preference appear in Table 3 (per parts and total) and Figure 1 (only total). As in Suárez and Muñoz (2011), there appears a considerable difference between the scores in grades 3 and 4 as compared to those across the rest of grades.

Table 3. Descriptive results on the MLAT-ES of Group 1 depending on L1 preference.

Grade	Lg pref.	N	MLAT-ES Part 1		MLAT-ES Part 2		MLAT-ES Part 3		MLAT-ES Part 4		MLAT-ES Total	
			Mdn	SD	*Mdn*	SD	*Mdn*	SD	*Mdn*	SD	*Mdn*	SD
3	Spanish	25	10.00	5.74	9.00	6.24	20.00	7.88	15.00	7.01	59.00	18.47
	Catalan	25	13.00	6.52	12.00	5.76	27.00	8.87	21.00	5.49	75.00	18.27
	Both	16	10.00	8.18	10.00	5.46	22.00	11.45	20.00	6.59	61.50	24.94
4	Spanish	30	21.50	7.45	15.00	7.31	26.50	10.21	21.00	6.80	84.00	24.90
	Catalan	21	25.00	5.86	21.00	7.59	29.00	7.84	23.00	3.38	93.00	18.63
	Both	24	16.00	5.79	18.00	6.87	22.00	9.62	20.50	5.10	76.00	19.57
5	Spanish	24	25.00	5.79	17.50	8.30	31.50	9.53	23.50	6.05	95.50	22.95
	Catalan	16	21.50	6.23	22.50	8.43	28.00	9.8	24.00	5.52	92.00	23.87
	Both	17	25.00	3.38	18.00	7.07	33.00	5.62	21.00	4.96	98.00	15.07
6	Spanish	21	23.00	7.48	24.00	7.69	30.00	8.27	24.00	2.35	94.00	19.32
	Catalan	19	27.00	6.61	26.00	5.53	36.00	6.54	25.00	3.68	114.00	18.54
	Both	20	23.50	5.53	25.50	5.29	35.50	7.32	23.50	4.11	107.50	17.8
7	Spanish	35	28.00	3.46	26.00	7.04	35.00	5.12	24.00	5.13	109.00	15.28
	Catalan	7	27.00	4.39	26.00	4.25	36.00	8.43	25.00	0.78	105.00	11.22
	Both	25	28.00	2.51	26.00	5.11	35.00	4.69	23.60	2.67	113.00	10.52

Figure 1. Descriptive results on the MLAT-ES total results of Group 1 depending on L1 preference.



To measure the influence of the participants' L1 preference on the students' performance on the MLAT-ES, Kruskal-Wallis tests were run for each grade (Table 4). Statistically significant values appear in bold.

Table 4. Group 1: Kruskal-Wallis H tests on the MLAT-ES.

Grade	Kruskal-Wallis df (2)	MLAT-ES Part 1	MLAT-ES Part 2	MLAT-ES Part 3	MLAT-ES Part 4	MLAT-ES Total
3	*H*	4.043	2.169	4.552	9.528	7.507
	Asymp. Sig.	.132	.338	.103	**.009**	**.023**
4	*H*	8.522	1.217	1.949	5.912	4.849
	Asymp. Sig.	**.014**	.544	.377	.052	.089
5	*H*	.504	.289	2.571	1.449	.752
	Asymp. Sig.	.777	.865	.277	.485	.687
6	*H*	3.825	2.242	8.966	2.276	6.570
	Asymp. Sig.	.148	.326	**.011**	.320	**.037**
7	*H*	.063	.510	.020	2.021	.723
	Asymp. Sig.	.969	.775	.990	.364	.697

The Kruskal-Wallis tests showed five significant differences. In Grades 3 and 6, both parts (4 and 3, respectively) showed that language preference might play a role on the participants' scores on the MLAT-ES, and the rhymes in Part 3 are also language-dependent, even though the pronunciation of Catalan and Spanish differs in a few consonant and vowel phonemes. These significant differences being very powerful ($H(2) = 9.528$, $p = .009$ and $H(2) = 8.966$, $p = .011$) and between the three groups were probably what made the total score be affected by the participants' language preference, as revealed by the post-hoc Mann-Whitney U-tests. In the case of grade 4, however, part 1, which is also language-bound, showed a significant difference ($H(2) = 8.522$, $p = .014$) that did not cause any differences in the Total score.

Post-hoc Mann-Whitney U-tests using a Bonferroni-adjusted alpha level of .017 (0.5/3) were used to compare all pairs of groups with significant differences. The *p* significant values ranged from .003 to .023 and the pairs compared never proved a correspondence between the participants' language preference and a higher score in that version of the MLAT or, for that matter, a higher performance from those who preferred Spanish over Catalan or both languages indifferently.

Table 5. Significant Mann-Whitney U-Tests results on the MLAT-ES grouped by language preference.

Grade	Part	Language pair	Z	Significance level
3	4	Spanish < both	2.280	.023
	Total	Spanish < Catalan	2.855	.004
	Total	Spanish < both	2.844	.004
4	1	Both> Catalan	2.988	.003
6	3	Spanish < Catalan	2.763	.005
	3	Spanish < Both	2.292	.022
	Total	Spanish < Catalan	2.431	.015

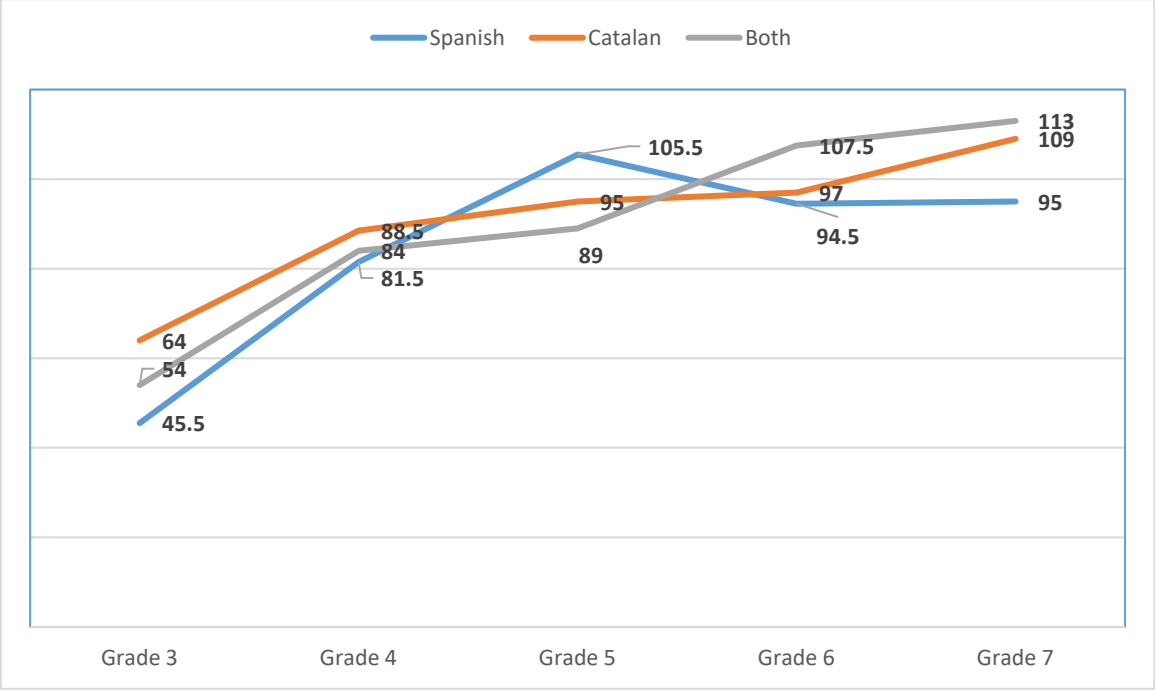
5.2 MLAT-EC depending on L1s preference per grade

The descriptive results on the MLAT-EC for Group 2, i.e., those who took the MLAT-EC in the first place, show a very similar distribution to that of Group 1 (see Table 6 for parts and total scores and Figure 2 for total scores across grades).

Table 6. Descriptive results on the MLAT-EC of Group 2 depending on L1 preference.

Grade	Lg pref.	N	MLAT-EC Part 1		MLAT-EC Part 2		MLAT-EC Part 3		MLAT-EC Part 4		MLAT-EC Total	
			Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
3	Spanish	22	9.00	9.47	10.00	4.24	15.00	9.53	16.00	7.76	45.50	20.06
	Catalan	21	14.00	7.32	9.00	7.55	16.00	10.24	13.00	8.04	64.00	27.99
	Both	14	12.00	6.59	7.50	5.78	14.50	7.63	20.50	6.22	54.00	18.05
4	Spanish	16	21.00	5.83	14.50	7.65	30.50	9.69	20.00	7.57	81.50	24.78
	Catalan	30	23.00	5.73	15.00	7.49	31.50	7.15	23.00	4.27	88.50	19.18
	Both	16	21.00	7.81	12.00	5.29	27.00	11.54	22.00	3.05	84.00	22.60
5	Spanish	16	26.50	4.89	23.50	5.97	32.50	5.73	24.00	2.97	105.50	14.04
	Catalan	29	23.00	5.55	22.00	7.09	32.00	8.09	24.00	4.59	95.00	18.34
	Both	16	23.00	5.58	13.00	6.49	29.50	7.29	23.50	6.10	89.00	20.91
6	Spanish	16	22.50	6.65	17.50	7.99	31.50	9.30	23.00	2.92	94.50	20.90
	Catalan	26	24.50	4.10	20.00	6.50	33.00	6.88	20.00	5.39	97.00	13.83
	Both	18	25.50	5.77	22.00	7.02	34.00	4.73	25.00	4.68	107.50	17.69
7	Spanish	17	25.00	6.46	17.00	6.65	32.00	6.77	24.00	6.63	95.00	21.73
	Catalan	41	27.00	3.45	26.00	6.18	34.00	7.54	25.00	1.26	109.00	15.41
	Both	6	27.00	6.83	22.00	8.09	32.50	4.14	24.50	3.92	103.50	21.39

Figure 2. Descriptive results on the MLAT-EC total results of Group 2 depending on L1 preference.



The same procedure as with Group 1 was followed to determine if there were significant differences caused by language preference in the participants' performance on the MLAT-EC by Group 2. Significant *p value*s appear in bold letters in Table 7 below.

Table 7. Group 2: Kruskal-Wallis H tests on the MLAT-EC.

Grade	Kruskal-Wallis df(2)	MLAT-EC Part 1	MLAT-EC Part 2	MLAT-EC Part 3	MLAT-EC Part 4	MLAT-EC Total
3	*H*	1.887	1.753	1.435	1.977	.946
	Asymp. Sig.	.389	.416	.488	.372	.623
4	*H*	3.109	1.970	.929	2.128	1.950
	Asymp. Sig.	.211	.373	.628	.345	.377
5	*H*	.2799	6.644	3.222	.113	3.183
	Asymp. Sig.	.247	**.036**	.200	.945	.204
6	*H*	3.825	2.242	8.966	2.276	6.570
	Asymp. Sig.	.148	.326	**.011**	.320	.057
7	*H*	.063	.510	.020	2.021	.723
	Asymp. Sig.	.444	**.015**	.544	**.004**	.068

Two grades (6 and 7) present significant differences in part 3 (grade 6: $H(2) = 8.966$, $p = .011$) and in part 4 (grade 7: $H(2) = 2.021$, $p = .004$) which are more, language-dependent than Part 2, in which significant differences can be found in grades 5 ($H(2) = 6.644$, $p = .036$) and 7 ($H(2) = .510$, $p = .015$.

Post-hoc Mann-Whitney tests using a Bonferroni-adjusted alpha level of .017 (0.5/3) were used to compare all pairs of groups with significant differences (Table 8).

Table 8. Significant Mann-Whitney U-Tests results on the MLAT-EC grouped by language preference.

Grade	Part	Language pair	Z	Significance level
5	2	Spanish > Both	-2.589	.010
6	3	Spanish > Both	-2.292	.022
		Catalan > Both	-2.033	.042
7	2	Catalan > Spanish	-2.881	.004
	4	Catalan > Spanish	-3.294	.001

In this case, the language preference of the group with the highest score corresponded to that of the MLAT-EC on three occasions. One of them appeared in Part 2, which is not so language-dependent as Parts 3 or 4. In the other cases, this time it was those preferring Spanish that significantly outperformed those who showed no preference for either of the languages. Consequently, there are no significant differences that correspond to the language preference and the language of the aptitude test.

5.3 MLAT-EC and MLAT-ES across grades depending on participants' L1 preference

All the Kruskal-Wallis tests run for Group 1 and Group 2 across grades turned out to be significant ($H(4)$; ps = <.043). Post-hoc Mann-Whitney analyses with the corresponding Bonferroni correction

for multiple comparisons also showed that the pattern of the differences across grades is very similar to the one found in Suárez (2010), this time with L1 preference as grouping variable in both the MLAT-ES and the MLAT-EC. These differences are especially recurrent between grades 3 and 4 (see Table 6, significant values in bold). Remarkably, these differences are not so consistent in these grades in Group 1, in parts 3 and 4 (language-bound) and part 2 (not so language-bound). Significant differences are also present in Part 1 between grades 6 and 7 on two occasions in Group 1 and 2, where the groups were somewhat unbalanced in number of participants. All in all, it is difficult to conclude that these significant differences are due to the participants' language preference only.

Table 9. Post-hoc Mann-Whitney analyses across grades depending on language preference.

Gr. 1	**MLAT-ES Part 1**			**MLAT-ES Part 2**			**MLAT-ES Part 3**			**MLAT-ES Part 4**			**MLAT-ES Total**		
Grade	**Sp.**	**Cat.**	**Both**	**Sp.**	**Cat.**	**Both**	**Sp.**	**Cat.**	**Both**	**Sp.**	**Cat.**	**Both**	**Sp.**	**Cat.**	**Both**
3 – 4	**.000**	**.004**	**.093**	.218	.152	**.020**	.112	.279	.796	**.038**	.237	.698	**.004**	**.038**	.093
4 – 5	.429	.517	**.000**	.584	.469	.938	.273	.842	.003	.059	.550	.540	.273	.886	.003
5 – 6	.967	.229	.630	.102	.051	.098	.632	**.019**	.591	.053	.437	.591	.873	.060	.134
6 – 7	.056	.973	**.036**	.408	.390	.592	.128	.334	.894	.835	.569	.081	**.013**	.658	.286
Gr. 2	**MLAT-EC Part 1**			**MLAT-EC Part 2**			**MLAT-EC Part 3**			**MLAT-EC Part 4**			**MLAT-EC Total**		
Grade	**Sp.**	**Cat.**	**Both**	**Sp.**	**Cat.**	**Both**	**Sp.**	**Cat.**	**Both**	**Sp.**	**Cat.**	**Both**	**Sp.**	**Cat.**	**Both**
3 – 4	**.009**	**.002**	**.028**	**.011**	.061	**.028**	**.001**	**.047**	**.003**	.189	**.004**	.261	**.009**	**.003**	**.003**
4 – 5	**.005**	.879	1.00	.157	**.027**	.480	.476	.365	.480	.288	.138	.480	**.034**	.519	.480
5 – 6	.077	.522	.716	.288	.485	.071	.723	.701	**.015**	.480	.135	.292	.157	.898	**.039**
6 – 7	.112	.031	1.00	.866	**.002**	1.000	.881	.605	.478	.598	*	*	.866	**.016**	.478

* Unable to compute because all sample medians in this pair are less than or equal to the hypothesized median

5.4 MLAT-ES and MLAT-EC distributed among bilinguals compared to monolinguals

As the raw data of the MLAT-ES norming study were not available, the percentage of difference as well as the effect size was calculated using the means, standard deviations and number of participants available in the MLAT-ES Manual (Stansfield et al., 2004) (see Table 8) applied to Cohen's *d* formula for relative effect size from t-tests. When comparing the MLAT-EC results with the results of the MLAT-ES norming study, one must remember that the MLAT-EC has one less item. Nevertheless, the means were in all cases higher than in the MLAT-ES norming study.

Table 10. Means and standard deviations of total score on the MLAT-ES and the MLAT-EC across grades.

Grade	**MLAT-ES *Manual***		**MLAT-ES Group 1**		**MLAT-EC Group 2**	
	Mean	**SD**	**Mean**	**SD**	**Mean**	**SD**
3	51.2	25.3	65.09	21.09	55.12	22.75
4	65.9	28.0	81.84	22.16	83.77	21.77
5	75.6	25.9	91.12	21.12	93.13	18.44
6	86.5	23.0	99.93	19.04	94.20	17.32
7	94.0	19.4	106.69	13.24	99.95	18.37

After applying the formula for Cohen's *d*, it can be seen in Table 11 that except for Grade 3, where all tests were very difficult in all versions, and two occasions when comparing performance on the MLAT-ES in the *Manual* and the MLAT-EC in grades 6 and 7, when scores also reach a plateau, for the rest of comparisons, there was a medium effect size. The small effect size in grade 7 is actually close to medium effect size, i.e. 0.40. Also, 6 out of the 10 comparisons present a medium

increase, with percentages ranging from 21 to -27 and one of the small increases, in grade 7, actually borders the medium status (14%).

Table 11. Cohen's *d* effect size and percentage of increase between comparisons.

Grade	Tests comparison	Cohen's *d*	Effect size	Percentage change	Direction
3	MLAT-ES *Manual* vs MLAT-ES Group 1	0.57	medium	27	medium increase
	MLAT-ES *Manual* vs MLAT-EC Group 2	0.16	small	8	small increase
4	MLAT-ES *Manual* vs MLAT-ES Group 1	0.6	medium	24	medium increase
	MLAT-ES *Manual* vs MLAT-EC Group 2	0.67	medium	27	medium increase
5	MLAT-ES *Manual* vs MLAT-ES Group 1	0.62	medium	21	medium increase
	MLAT-ES *Manual* vs MLAT-EC Group 2	0.71	medium	23	medium increase
6	MLAT-ES *Manual* vs MLAT-ES Group 1	0.6	medium	16	medium increase
	MLAT-ES *Manual* vs MLAT-EC Group 2	0.35	small	9	small increase
7	MLAT-ES *Manual* vs MLAT-ES Group 1	0.71	medium	14	small increase
	MLAT-ES *Manual* vs MLAT-EC Group 2	0.39	small	7	small increase

6. Discussion

The first and second research questions aimed at shedding light on the issue of whether language preference could affect the participants' performance on the MLAT-ES and the MLAT-EC. In the case of the MLAT-ES, the results show that there are very few significant differences that could be due to language preference. While significant differences do appear in the parts of the test that are language-dependent, i.e., Parts 1 and 4, in grades 3, 4 and 6 (and the Total Score in Grade 3), significant differences also appear in Grade 7 in Part 2. However, at no point do these differences favor those who self-report a preference for Spanish. In the case of the MLAT-EC, those learners showing a preference for Catalan do outperform the other groups. However, that happens in the upper grades, where in principle the learners' L1s have already been fully acquired. Consequently, it remains unknown if those leaning towards Catalan as a language of preference outperformed the other two groups due to this language preference or due to other factors as could be, for instance, the main language of education in their school. What might have happened in the case of the MLAT-ES is that participants in the lower grades did not have their own L1s systems developed enough to permit them to handle linguistic material in Latin American Spanish with which they were not as familiar as with that spoken in the Balearic peninsula. Therefore, in grades 3 and 4 the variety of Spanish used in the MLAT-ES might be a confounding factor in the lower grades. Although the overall results were not affected by the Spanish variety of the test (Suárez, 2010), it might be worth developing a version of the test using the peninsular variety of Spanish.

Regarding the MLAT-EC, the differences appear from Grade 5 on. In this case, the participants' attention as well as their academic performance might have influenced the results. Focusing on

6th-graders' performance on the MLAT-EC, except for Part 4, the group showing a preference for Spanish perform lower on three out of the four parts, and so do those in Grade 7, where it is Part 2 the one receiving a score significantly lower than in the other two groups. It should also be highlighted that students in Spain receive formal instruction on syntax (and word functions) at grade 5, so by Grade 7 they could infer what Part 2 was about even if it was not explicitly instructed. Consequently, the group from Spain performing lower on the MLAT-EC could be explained by not applying formal concepts that other groups might have applied. In any case, it seems that the approach adopted towards this part of the test from Grades 5 on changes probably due to the explicit metalinguistic knowledge that the students could then apply when answering it.

Despite the significant differences exposed above, the same pattern of scores is found across grades for all language preferences for both tests, with most significant differences lying between grades 3 and 4. This certainly reinforces the equivalence between both tests and their strength in terms of validity, as the participants' L1 preference did not affect the evolution of scores. The results obtained are also comparable to those obtained in Suárez and Muñoz (2011). Therefore, it can once again be stated that language aptitude might be innate and, therefore, very much determined in its origins and further potential of development, but it also keeps developing hand in hand with the individual's cognitive development.

All in all, the results obtained do not seem to be influenced by one's L1 preference or by the language of the test except for grade 3 in the case of the MLAT-ES. Consequently, it could be said that the fact that the MLAT-ES and MLAT-EC are not language neutral is not a factor that challenges its validity for this language pair. It remains unknown, though, if scores were affected by the participants' L1s mastery, which is one of the reasons why nowadays there is a preference towards aptitude tests that are language neutral (Rogers et al., 2017).

While the validity of this pair of tests might not be put into question, except for the case of the MLAT-ES in grade 3, there is a clear advantage for the cohort in this study as compared to the cohort in the norming study. One of the clear differences between both groups is the simultaneous bilingual status of the Catalan cohort, which would reinforce those studies who have found advantages for bilinguals in language learning. However, it is not clear if it is bilingualism that enhances language aptitude or vice versa. Besides, this statement should be taken with caution, given the lack of detailed data regarding this aspect. Another factor that could certainly have influenced the results favoring the Catalan population is formal instruction, as in Muñoz (2014). The participants had been exposed to formal instruction in both Spanish and Catalan since Grade 1 and that instruction became grammar oriented from Grade 3 on, with students learning the syntactic function of sentence elements in Grades 5 and 6. We understand that in the countries that participated in the field testing and validation of the MLAT-ES, instruction in formal grammar is less rigorous than in Spain.

7. Conclusions and limitations

This study presented the comparability of two already normed aptitude tests, the MLAT-ES and the MLAT-EC. The results across grades were robust enough to affirm that both tests are equivalent in validity despite the test taker's language of preference. The differences within grades could not be explained by language preference either. However, the fact that grade 3 showed a somewhat different performance on the MLAT-ES leads us to opt for the Catalan version for this grade when dealing with speakers of this language pair. These tests, which also proved to be valid

and reliable in the corresponding norming studies, are therefore a robust tool to measure aptitude in young learners following Carroll's model. Naturally, new views of aptitude are appearing which include other cognitive abilities as well as factors such as motivation and personality, but this should not be an undermining factor for the traditional measurement of aptitude or for the inclusion of the MLAT-E in any subsequent aptitude studies with young learners. Indeed, the MLAT-EC has been found to correlate as positively with learning vocabulary in a "modern" learning context as with learning vocabulary through subtitles (Suárez & Gesa, 2022). Nevertheless, this study is not without limitations: a measure of the participants' L1 mastery could have been used instead of their self-reported linguistic preference or dominance and more accurate details on the norming-study participants' education could have helped to explain more precisely the reasons why the bilingual Catalan/Spanish group outperformed them.

8. References

Bialystok, E. (2017). The bilingual adaptation: How minds accommodate experience. *Psychological bulletin, 143*(3), 233-262. http://dx.doi.org/10.1037/bul0000099

Bialystok, E., Craik, F.I., Klein. R., & Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: evidence from the Simon task. *Psychology and Aging*, *19*(2), 290-303. https://doi.org/10.1037/0882-7974.19.2.290

Bialystok, E. & Martin, M. M. (2004). Attention and inhibition in bilingual children: Evidence from the dimensional change card sort task. *Developmental Science, 7*(3), 325–339. https://doi.org/10.1111/j.1467-7687.2004.00351.x

Biedroń, A., & Véliz-Campos, M. (2021). Trainability of foreign language aptitudes in children. In J. Rokita-Ja´skow & A. Wolanin (Eds.), Facing diversity in child foreign language education, Second language learning and teaching (pp. 39-53). https://doi.org/10.1007/978-3-030-66022-2_3

Bley-Vroman, R. (2009). The evolving context of the fundamental difference hypothesis. *Studies in Second Language Acquisition, 31*(2), 175-198. http://dx.doi.org/10.1017/S0272263109090275

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). Newbury House.

Carroll, J. B., & Sapon, S. (1959). *Modern Language Aptitude Test (MLAT): Manual.* Psychological Corporation.

___. (1967). *Modern Language Aptitude Test-Elementary.* New York: Psychological Corporation.

___. (2002). *Modern Language Aptitude Test Manual*. Second Language Testing, Inc.

Chipere, N. (2003). *Understanding complex sentences: Native speaker variations in syntactic competence.* Palgrave.

Chomsky, N. (1975). *Reflections on language.* Pantheon.

Costa, A., Hernández, M., Sebastián-Gallés, N. (2008). Bilingualism aids conflict resolution: evidence from the ANT task. *Cognition*, *106*(1), 59-86. https://doi.org/10.1016/j.cognition.2006.12.013

Dąbrowska, E. (2019). Experience, aptitude, and individual differences in linguistic attainment: A comparison of native and nonnative speakers. *Language Learning, 69*(1), 72-100. https://doi.org/10.1111/lang.12323

Ellis, N. C. (1996). Sequencing in SLA: phonological memory, chunking, and points of order. *Studies in Second Language Acquisition, 18*(1), 91-126. http://dx.doi.org/10.1017/S0272263100014698

Ehrman, M. E., & Oxford, R. L. (1995). Cognition plus. Correlates of language learning success. *The Modern Language Journal*, *79*(1), 67-89. http://dx.doi.org/10.1111/j.1540-4781.1995.tb05417.x

Eisenstein, M. (1980). Childhood bilingualism and adult language learning aptitude. *Applied Psychology, 26*(1-2), 159-172. https://doi.org/10.1111/j.1464-0597.1980.tb00888.x

Ganschow, L. & Sparks, R. (1995). Effects of direct instruction in Spanish phonology on the native language skills and foreign language aptitude of at-risk foreign language learners. *Journal of Learning Disabilities, 28*, 107–120. https://doi.org/10.1177%2F002221949502800205

Gesa, F., & Suárez, M.M. (2022). Individual differences in young learners' L2 vocabulary development: the case of language aptitude and type of exposure. In V. De Wilde & C. Goriot (Eds.), *Individual differences in L2 learning of children and adolescents* (pp. 65-85). DeGruyter.

Granena, G. & Long, M. (2013). Age of onset, length of residence, language aptitude, and ultimate L2 attainment in three linguistic domains. *Second Language Research 29*(3), 311–343. http://dx.doi.org/10.1177/0267658312461497

Grigorenko, E. A., Sternberg, R. J., & Ehrman, M. (2000). Theory-based approach to the measurement of foreign language learning ability: The Canal-F theory and test. *Modern Language Journal, 84*(3), 390-405. http://www.jstor.org/stable/330568

Hakuta, K., Bialystok E., & Wiley, E. (2003). Critical evidence: A test of the critical-period hypothesis for second language acquisition. *Psychological Science, 14*(1), 31-38. https://doi.org/10.1111%2F1467-9280.01415

Hopp, H. Vogelbacher, M., Kieseier, T., & Thoma, D. (2019). Bilingual advantages in early foreign language learning: Effects of the minority and the majority language. *Learning and Instruction, 69*, 99-110. http://dx.doi.org/10.1016/j.learninstruc.2019.02.001

Hyltenstam, K. (2018). Polyglotism. A synergy of abilities and predispositions. In K. Hyltenstam, I. Bartning, & L. Fant (Eds.), *High-level language proficiency in second language and multilingual contexts* (pp. 170–195). Cambridge University Press.

Hyltenstam, K. (2021). Language aptitude and language awareness: Polyglot perspectives. *Annual Review of Applied Linguistics, 41*, 55-75. https://doi.org/10.1017/S0267190521000027

Institut d'Estadística de Catalunya (2018). *Enquesta d'usos lingüístics de la población (EULP, 2018).* https://llengua.gencat.cat/web/.content/documents/dadesestudis/altres/arxius/EULP2018_barcelona-presentacio.pdf

Kiss, C. (2009). The role of aptitude in young learners' foreign language learning. In M. Nikolov (Ed.), *The age factor and early language learning* (pp. 253-276). De Gruyter Mouton. https://doi.org/10.1515/9783110218282.253

Kiss, C., & Nikolov, M. (2005). Developing, piloting, and validating an instrument to measure young learners' aptitude. *Language Learning, 551*(1), 99-150. https://doi.org/10.1111/j.0023-8333.2005.00291.x

Kormos, J. (2013). New conceptualizations of language aptitude in second language attainment. In G. Granena & M. Long (Eds.), *Sensitive periods, language aptitude and ultimate attainment* (pp. 131-152). John Benjamins.

Larsen-Freeman, D., & Long, M. H. (1991). *An introduction to second language research*. Longman.

Li, S. (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. http://dx.doi.org/10.1093/applin/amu054

___. (2016). The construct validity of language aptitude. *Studies in Second Language Acquisition, 38*(4), 801-342. http://dx.doi.org/10.1017/S027226311500042X

Meara, P. (2005). *LLAMA Language Aptitude Tests*. Lognostics.

Muñoz, C. (2014). The association between aptitude components and language skills in young learners. In M. Pawlak & L. Aronin (Eds.), *Essential Topics in Applied Linguistics and Multilingualism, Second Language Learning and Teaching* (pp. 51-68). Springer International Publishing.

Nation, R. & Mclaughlin, B. (1986). Experts and novices: An information processing approach to the good language learner. *Applied Psycholinguistics 7*(1), 41-55. https://doi.org/10.1017/S0142716400007177

Piaget, J. (1964). Cognitive development in children: development and learning. *Journal of Research in Science Teaching, 2,* 176-186. https://doi.org/10.1002/tea.3660020306

Poarch, G. J., & Van Hell, J. G. (2012). Cross-language activation in children's speech production: Evidence from second language learners, bilinguals, and trilinguals. *Journal of Experimental Child Psychology, 111*(3), 419-438. https://doi.org/10.1016/j.jecp.2011.09.008

Rogers, V. E., Meara, P., Aspinall, R., Fallon, L., Goss, T., Keey, E. & Thomas, R. (2016). Testing aptitude. *EUROSLA Yearbook, 16*(1), 179–210. https://doi.org/10.1075/eurosla.16.07rog

Rogers, V., Meara, P., Barnett-Legh, T., Curry, C., & Davie, E. (2017). Examining the LLAMA aptitude tests. *Journal of the European Second Language Association, 1(*1), 49–60. https://doi.org/10.22599/jesla.24

Safar, A., & Kormos, J. (2008). Revisiting problems with foreign language aptitude. *International Review of Applied Linguistics in Language Teaching (IRAL), 46*(2),113-136.

Singleton, D. (2014). Apt to change: The problematic of language awareness and language aptitude in age-related research. *Studies in Second Language Learning and Teaching, 4*(3), 557-571.

___. (2017). Language aptitude: Desirable traitor acquirable attribute? *Studies in Second Language Learning and Teaching*, *7*(1), 83-109. http://dx.doi.org/10.14746/ssllt.2017.7.1.5

Skehan, P. (1986). Where does language aptitude come from? In P. Meara (ed.), *Spoken language* (pp. 86-92). Centre for Information on Language Teaching.

___. (1989). Individual differences in second language learning. Edward Arnold.

___. (2002). Theorising and updating aptitude. In P. Robinson (ed.), *Individual differences and instructed language learning* (pp. 69–95). John Benjamins Publishing Company.

Sparks, R. L., & Ganschow, L. (1991). Foreign language learning difficulties: Affective or native language aptitude differences? *The Modern Language Journal, 75*(1)*,* 3-16. https://doi.org/10.1111/j.1540-4781.1991.tb01076.x

___. (1993). The impact of native language learning problems on foreign language learning: Case study illustrations of the Linguistic Coding Deficit Hypothesis. *Modern Language Journal, 77*(1), 58–74. https://doi.org/10.2307/329559

Sparks, R., Ganschow, L., Artzer M., & Patton, J. (1997). Foreign language proficiency of at-risk and not-at-risk learners over two years of foreign language instruction. *Journal of Learning Disabilities, 30*(1), 92-98. https://doi.org/10.1177/002221949703000108

Sparks, R. L., 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

Stansfield, C.W., Reed, D., & Velasco, A. M. (2004). *Prueba de aptitud para lenguas extranjeras-Versión de primaria*. Language Learning and Testing Foundation.

Sternberg, R. J., & Grigorenko, E. L. (Eds.). (2002). *The general factor of intelligence: How general is it?* Lawrence Erlbaum Associates Publishers.

Suárez, M.M. (2010). Language aptitude in young learners: The Elementary Modern Language Aptitude Test in Spanish and Catalan. Doctoral dissertation, Universitat de Barcelona. https://www.tesisenred.net/handle/10803/38244#page=1

___. (2022). A need for an aptitude test for young learners in Catalan: The case of the Modern Language Aptitude Test - Elementary in Catalan. *Language Teaching Research Quarterly, 31*, 101-118. https://doi.org/10.32038/ltrq.2022.31.08

Suárez, M.M., & Gesa, F. (2019). Learning vocabulary with the support of sustained exposure to captioned video: do proficiency and aptitude make a difference?. *The Language Learning Journal, 47*(3), 497-517. http://doi.org/10.1080/09571736.2019.1617768

Suárez, M.M., & Muñoz, C. (2011). Aptitude, age and cognitive development: The MLAT-E in Spanish and Catalan. *EUROSLA Yearbook, 11*, 5–29. https://doi.org/10.1075/eurosla.11.03sua

Tarone, E., Bigelow, M., & Hansen, K. (2013). *Literacy and second language oracy*. Oxford University Press.

Turker, S., Reiterer, S., Schneider, P., & Seither-Preisler, A. (2018). The neuroanatomical correlates of foreign language aptitude. In S. M. Reiterer (Ed.), *Exploring language aptitude: Views from psychology, the language sciences, and cognitive neuroscience* (pp. 119–149). Springer.

Wells, C. G. (1985). *Language development in the pre-school years*. Cambridge University Press.

___. (1986). Variation in child language In P. Fletcher & T. Garman (eds.), *Language acquisition* (pp. 109-139). Cambridge University Press.

Wen, Z., Biedroń, A., & Skehan, P. (2017). Foreign language aptitude theory: Yesterday, today and tomorrow. *Language Teaching*, *50*(1), 1-31. https://doi.org/10.1017/S0261444816000276

Postal and e-mail addresses

Maria-del-Mar Suárez

Pg. Vall d'Hebron, 171. CP: 08035.
Faculty of Education. Universitat de Barcelona
Barcelona, Catalonia, Spain

Charles W. Stansfield
Language Learning and Testing Foundation, Inc.
10713 Mist Haven Terrace

Rockville, MD 20852
United States of America