

Language Proficiency among Non-Native Chinese Language Learners: A Discriminant Analysis [†]

Sian Hoon Teoh ^{1,*}, Buai Chin Heng ², Yee Feng Neo ² and Julia Yin Yin Tan ³

¹ Faculty of Education, Universiti Teknologi MARA, Cawangan Selangor, Bandar Puncak Alam 42300, Selangor, Malaysia

² Akademi Pengajian Bahasa, Universiti Teknologi MARA, Shah Alam 40450, Selangor, Malaysia

³ Faculty of Language Studies and Human Development, Universiti Malaysia Kelantan, Bachok 16300, Kelantan, Malaysia

* Correspondence: teohsian@uitm.edu.my

[†] Presented at the International Academic Symposium of Social Science 2022, Kota Bharu, Malaysia, 3 July 2022.

Abstract: The Chinese Language (CL) is difficult to learn, and CL non-native learners are finding it increasingly difficult. Therefore, additional information on how to guide students is critical. As a result, the goal of this research is to find out what criteria distinguish high and low proficiency students in Chinese learning. The data was collected through a Google form questionnaire from 79 CL non-native students who had previously studied Chinese. The findings demonstrate that the only difference between the two groups is the writing strategies used. This study concludes that various teaching methods should emphasise writing skills in order to become proficient in Chinese.

Keywords: Chinese Language (CL); discriminant analysis; non-native CL learners; proficiency



Citation: Teoh, S.H.; Heng, B.C.; Neo, Y.F.; Tan, J.Y.Y. Language Proficiency among Non-Native Chinese Language Learners: A Discriminant Analysis. *Proceedings* **2022**, *82*, 42. <https://doi.org/10.3390/proceedings2022082042>

Academic Editor: Mohamad Rahimi Mohamad Rosman

Published: 14 September 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The Chinese Language (CL) seems to have drawn more people to learn it seriously. Seen in this light, it has been observed that parents pay close attention to their children's CL learning settings. Several Malaysian studies have investigated the learning situations of these non-native CL students, including research into the causes that motivate parents to send their children to CL medium schools (known as *Sekolah Jenis Kebangsaan Cina* (SJKC)). These studies were also conducted on the pupils' academic achievement and learning issues [1–6].

These studies' conclusions have unintentionally highlighted a few problems faced by CL non-native learners. The problems highlighted were the non-native students' poor academic performance at SJKCs [2] and issues with the CL being the instructional medium of mathematics and science, hindering the learning of the latter as well [3]. The other much-discussed problems were the inadequate teaching methods used by teachers [1], an inconducive learning environment at home, and low expectations and support from parents [1–3].

This investigation was prompted by the public's interest in learning the CL effectively. Hence, the goal of this study was to investigate the learning elements that induced the learners' CL mastery of the 'high' and 'poor' proficiency learners in CL classrooms at SJKC schools.

2. Literature Review

2.1. Learning Chinese Language

Learning the CL is becoming increasingly popular as a result of education globalisation [7], which has resulted in an increasing interconnectedness of societies in economics, technology, politics, culture, and language [8]. Hence, language plays a vital role in maximising these interactionist relationships.

Among the spoken languages in the world, it was reported that CL ranks the most favoured [9]. Correspondingly, non-native CL learners have benefited from economic development by meeting global needs via learning this language. Although difficult, the CL has become an important subject in the Malaysian Certificate of Education, which is taken by all fifth-form secondary school pupils in Malaysia [10]. Hence, non-native learners' CL proficiency and challenges in achieving proficiency have been gaining attention.

2.2. The Difficulties in Learning Chinese Characters

When learning any language, voice is frequently used during the hearing and reading stages. Listening and reading require phonological awareness, which is concerned with the sounds of the spoken language. Phonological awareness, which is based on phonologically constructed speaking practises in Western (alphabetic) languages, plays a major role in a learner developing their reading. However, the CL is a unique medium where logographic letters are used to convey meaning rather than phonological speech.

Studies also highlighted that phonological speech, which is based on sound similarities of different characteristics, has caused difficulties for CL learners [11,12]. In classrooms, learners are often asked to identify the correct characters they see based on the meaning of a sentence. However, some CL characters have similar sounds, although they are different. This adds to the learners' confusion, which intensifies the failure to identify similar sounds (but having different meanings as they are different characters).

Figure 1 shows phonological speech representing different words with different meanings. For example, the phonological speech of /shàng/ is often used in reading for different characters with different meanings. Figure 1 also shows that a character will have a different meaning when it is attached to another (character). For example, the character /shàng/ is attached to /mǎ/ to form /mǎ shàng/, which means 'immediately'. On the other hand, if it is attached to /miàn/ to form /shàng miàn/, it means 'above'. Also, a single character of the different meaning (e.g., /shàng/) holds a specific meaning which is 'still'. This example shows that the same phonologic speech can have different characters with different phonologic speech.

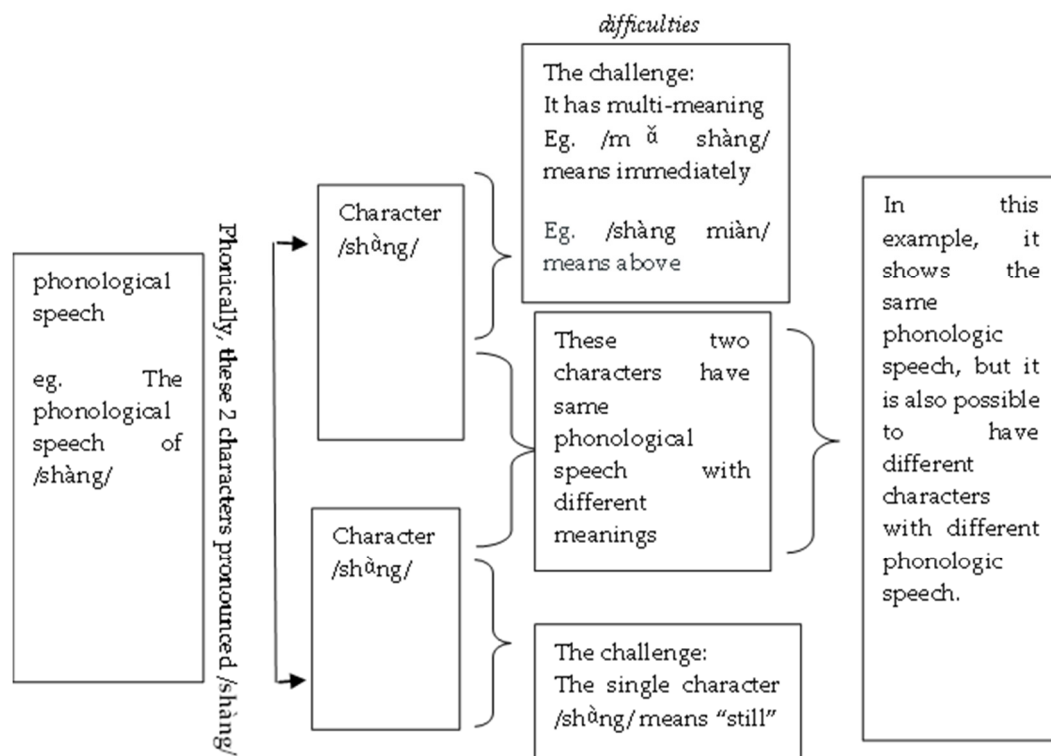


Figure 1. Complexity of Learning Chinese Language Characters: An Illustration.

Secondly, the complexity of learning this language is also challenged by the multi-meaning carried by a specific character and having different phonology for the same character with different meanings. As a result, a student's ability to read the CL is closely linked to his or her writing skills [6]. In short, writing skills are inextricably associated with CL proficiency.

2.3. Some Considerations in Learning the CL

It is noted that students' diverse origins, which may include strategies in learning, have an impact on their accomplishments. Hence in learning languages, students may be guided with specific strategies. It is important that students engage in the guided strategies when practising the four skills (reading, writing, listening, and speaking). As learning improves, learners may apply more strategies to improve acquisition and, thus, confidence. According to [13] students who apply appropriate strategies tend to be seen as good language learners. Applying appropriate strategies also requires appropriate knowledge for the learners to proceed with their ability. According to [14], carrying out appropriate strategies involves thought and behaviour. To this end, students' activities are equally vital in determining the success of language learning, which can be altered by motivation. Figure 2 shows the steps and links involved in achieving linguistic competency, showing the importance of strategy as a significant factor.

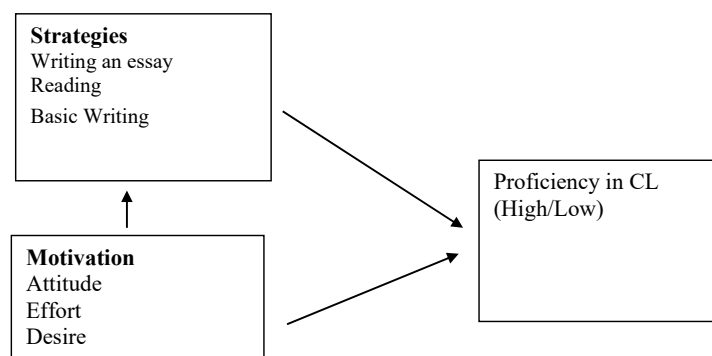


Figure 2. Proficiency in CL: Factors Involved.

3. Materials and Methods

3.1. Research Design and Sampling

This study employed a survey research design. Its sample was 79 non-native CL students who had experience in learning the CL at school. The focus of this study was on those who had shown good achievement in learning the Chinese language. Hence, the sample was former SJKC non-native CL high achievers in Malaysia. Taking note of the low percentage of high-achieving learners, which was an average of 800 students yearly, this study took 10% of 800 to respond to the questionnaires. Thus, a total of 79 students participated. The participants were then placed in two groups, namely 'low proficiency' and 'high proficiency'.

A sampling method called snowballing with volunteerism was employed in the selection of the samples. A list of samples was first identified by a few educators in Malaysia. Then, the participants' permissions were obtained to access their data. The participants who did not agree to participate were excluded. The inclusion criteria to select the samples were based on a volunteer basis, and most importantly, they consented to complete a Google form questionnaire.

3.2. Instrumentation

A questionnaire consisting of three parts was constructed following two references [15,16]. The three parts were CL proficiency, motivation, and strategies. Two experts were consulted to establish the content validity of the instrument. A pilot test was conducted on

79 participants. Its results showed a convincing and acceptable level of Cronbach's Alpha values, ranging from 0.76 to 0.85. Even though the construct 'writing an essay' showed a lower value of Cronbach's Alpha, the items contributed fittingly to the overall construct of writing, which was the combination of basic writing and essay writing (Cronbach's Alpha = 0.76).

3.3. Data Analysis

Discriminant Function Analysis (DA) was used to analyse the obtained data. The DA serves the same purpose as multiple linear regression by predicting an outcome. Hence, the DA was employed in this study since multiple linear regression is restricted to scenarios where the dependent variable is an interval variable. The regression equation also provides an estimated mean population numerical dependent variable value for specified weighted combinations of independent variable values. The 'proficiency level' was the dependent variable.

This study used two proficiency levels, 'low' and 'high', to address research question one. To address research question two, six proficiency levels were measured. The first independent variable was motivation, which consisted of attitude, effort, and desire. Another independent variable was strategy, which consisted of basic writing, essay writing, and reading. These independents were the discriminators (in regression analysis, the independent variables are predictors). This study focused on high achievers divided into two categories ('low proficiency' and 'high proficiency') in order to observe the factors that affect their proficiency.

In the DA, the independent variables are combined in weighted combinations to produce a single new composite variable, namely the discriminant score. Thus, the significant portions of the discriminant score reflect misclassifying cases into respective groups (low/high proficiency). A good DA model shows minimal misclassification, so the analysis detects the variables that primarily contribute to differentiating groups.

However, this was a simple discriminant analysis with two groups in the dependent variable. The simple discriminant analysis is provided with one set of eigenvalues: Wilks' Lambda and beta coefficients. The number of sets is always one less than the number of DV groups. Therefore, in this analysis, the data obtained were the respondents' demographic data and the answers given by them. Further, the 'proficiency level' was a nominal variable to indicate whether the learner was of high or low proficiency. The other variables were attitude, effort, desire, and writing strategies.

To reiterate, the aim of the analysis was to identify if these variables discriminate the participants' proficiency (low or high proficiency) and examine whether there were any significant differences between the 'low' and 'high' proficiency groups on each of the independent variables using group means and ANOVA.

The 79 respondents were divided into two groups based on their self-reported CL competency, which was based on the Common European Framework of Reference (CEFR) categorization. Low proficiency refers to those who rated themselves A1, A2, or B1, while high proficiency refers to the respondents rating themselves B2, C1, or C2.

4. Results

4.1. The First Finding

Research Question 1: Are the factors (motivation, namely attitude, effort, and desire; strategies, namely basic writing, writing an essay, and reading) significantly discriminate between the two groups ('low' and 'high' proficiency)?

If the Group Statistics and Tests of Equality of Group Means show that there are no significant group differences, it is not worthwhile to proceed any further with the analysis. In this study, the group statistics (group means and standard deviations) suggest that these may not be good discriminators as the separations are small. Nevertheless, to determine the significant discriminator, a Test of Equality of Group Means is used. Table 1 shows the descriptive statistics of all independent variables in the two groups (LP and HP) with

overall scores of 5.20, 4.81, 5.43, 3.19, 3.23, and 3.05 for attitude, effort, desire, basic writing strategy, essay writing, and reading, respectively. Specifically, for the LP, the mean scores are 5.18 (attitude), 4.76 (effort), 5.43 (desire), 3.13 (basic writing strategy), 3.15 (writing an essay), and 3.00 (reading).

Table 1. Group Statistics.

	Level2	Mean	Std. Deviation	Valid N (Listwise)	
				Unweighted	Weighted
1.00	Attitude	5.18	0.78	62	62.00
	Effort	4.76	0.92	62	62.00
	Desire	5.43	0.63	62	62.00
	Basic Writing Strategy	3.13	0.71	62	62.00
	Writing an Essay	3.15	0.73	62	62.00
	Reading	3.00	0.75	62	62.00
2.00	Attitude	5.26	0.94	17	17.00
	Effort	4.99	0.61	17	17.00
	Desire	5.41	0.71	17	17.00
	Basic Writing Strategy	3.37	0.50	17	17.00
	Writing an Essay	3.56	0.53	17	17.00
	Reading	3.21	0.59	17	17.00
Total	Attitude	5.20	0.81	79	79.00
	Effort	4.81	0.87	79	79.00
	Desire	5.43	0.64	79	79.00
	Basic Writing Strategy	3.19	0.67	79	79.00
	Writing an Essay	3.23	0.71	79	79.00
	Reading	3.05	0.72	79	79.00

While for the HP, the mean scores are 5.26 (attitude), 4.99 (effort), 5.41 (desire), 3.37 (basic writing strategy), 3.56 (writing an essay), and 3.21 (reading). The Tests of Equality of Group Means show no significant group differences; hence, proceeding further with the analysis is not worthwhile.

Table 2 provides statistical evidence of significance in differences between the means of the two groups for all IV's, with only the 'essay writing' producing a higher F value ($F = 4.80$), with a significant value of $p (=0.03) < 0.05$.

Table 2. Tests of Equality of Group Means.

	Wilks' Lambda	F	df1	df2	Sig.
Attitude	1.00	0.13	1	77	0.72
Effort	0.99	0.97	1	77	0.33
Desire	1.00	0.01	1	77	0.92
Basic Writing Strategy	0.98	1.70	1	77	0.20
Writing an Essay	0.94	4.80	1	77	0.03
Reading	0.99	1.04	1	77	0.31

Table 3 shows the Box's M test to test the null hypothesis of whether there is any difference among the groups. Alternatively, it tests equal population covariance matrices. In this analysis, the Box's M is 2.30 with $F = 2.25$, which is not significant at $p > 0.05$.

Table 3. Test Results.

Box's M		2.30
F	Approx.	2.25
	df1	1
	df2	6215.96
	Sig.	0.13

For the assumption of equal variances to hold, the log determinants should also be equal. Table 4 shows that the groups have a reasonable Log Determinant with values close to each other, indicating mild variability of the groups.

Table 4. Log Determinants.

Level2	Rank	Log Determinant
1.00	1	−0.64
2.00	1	−1.28
Pooled within-groups	1	−0.74

The eigenvalues in Table 5 show the result of a function. The canonical correlation is the multiple correlations between the predictors (independent variables) and the discriminant function. The function provides an index of overall model fit, which is interpreted as being the proportion of variance explained (R²). In this study, a canonical correlation of 0.24 suggests that the model is explaining a 5.9% variation in the grouping variable.

Table 5. Eigenvalues.

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	0.062 ^a	100.0	100.0	0.24

^a. First 1 Canonical Discriminant Functions were used in the analysis.

Wilks' Lambda indicates the significance of the discriminant function. Table 6 shows a significant function ($p < 0.000$), providing a proportion of total variability not being explained, i.e., it is the converse of the squared canonical correlation. Therefore, a 94.1% variation is unexplained in the function.

Table 6. Wilks' Lambda.

Test of Function(s)	Wilks' Lambda	Chi-Square	df	Sig.
1	0.94	4.63	1	0.03

4.2. The Second Finding

Research Question 2: Do motivation (attitude, effort, and desire) and strategies (basic writing, writing an essay, and reading) significantly influence the learners' proficiency?

The result of Research Question 2 is further explained since the model was significant (Table 7). Table 7 shows the F-value = 15.05 with $p < 0.05$, signifying that the analysed model is significant. Nevertheless, not all the dependent variables significantly contribute to the model, as displayed in Tables 8–10.

Table 7. ANOVA ^a.

	Model	Sum of Squares	df	Mean Square	F	Sig.
1	Regression	14.85	1	14.85	15.05	<0.001 ^b
	Residual	76.01	77	0.99		
	Total	90.86	78			

^a. Dependent Variable: Respondent's Self-Report CL Proficiency Level. ^b. Predictors: (Constant), Strategies.

Table 8. Model Summary.

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	0.40 ^a	0.16	0.15	0.99	0.16	15.05	1	77	<0.001

^a. Predictors: (Constant), Strategies.

Table 9. Coefficients ^a.

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	0.50	0.61		0.83	0.41
	Strategies	0.74	0.19	0.40	3.88	<0.001

^a. Dependent Variable: Respondent's Self-Report CL Proficiency Level.

Table 10. Excluded Variables ^a.

Model	Beta In	t	Sig.	Partial Correlation	Collinearity Statistics
					Tolerance
1	Motivation	0.08 ^b	0.68	0.50	0.08

^a. Dependent Variable: Respondent's Self-Report CL Proficiency Level. ^b. Predictors in the Model: (Constant), Strategies.

This analysis involved the dependent variables of six proficiency levels and the independent variables of 'motivation', measured from attitude, effort, and desire. The independent 'strategies' were measured from the items in basic writing, essay writing, and reading. A regression analysis was then employed to investigate the relationship among these three variables. The Stepwise method in the regression analysis (Table 8) shows that the model (with $p < 0.05$) is significant for the independent variable 'strategies' (Table 9) only. Hence, the variable 'motivation' (Table 10) was excluded from the analysis.

Table 9 shows that only the factor 'strategies' contributed to the model. 'Strategies' was the essay writing strategy.

Table 10 shows that the model in Table 9 will only be significant if 'motivation' is excluded from the analysis. In Table 8, the result indicates that only 16.3% (namely $R^2 = 0.16$) variation in proficiency level is explained by 'strategies'. Hence, other factors might have to be considered in future research.

5. Discussion

The first finding indicated that only the 'writing an essay' strategy contributed to differentiating the two groups. It shows that just by looking into any student's essay writing skills, one can easily indicate the student's proficiency level as being 'low' or 'high'.

The second finding supported the first finding in the analysis of regression. The regression analysis indicated that 16.3% (namely, $R^2 = 0.163$) variation in proficiency

level was explained by the participants' writing strategies. Despite the low percentage of contributions to the proficiency level, the indicator provides some insight into the work required in essay writing to achieve proficiency by non-native CL speakers in learning the language.

These findings were in accordance with many other authors' findings that success in learning the CL requires certain writing talents [11,12]. Understanding vocabulary [11] is important, but so are the tactics and approaches to coordinate and integrate the CL characters. Students are expected to demonstrate the ability to use diverse characters in essays, such as combining the characters to bring new meanings. As a result, if a character is correctly specified, it will provide multiple inputs [17]. This situation was investigated in this study's literature review (Figure 1). The difficulties of handling characters in writing or any other form of communication have produced much uncertainty.

The acquisition of the CL characters has been the subject of investigation [18,19], which leads to the discovery that learners must be exposed to the larger workings of Chinese characters. This is particularly important in Malaysia as Malaysian tests emphasise writing rather than reading. Mindful that students thriving in the CL have a basic understanding of how to use the characters, therefore, in classrooms, the teaching and use of the characters should be enhanced. In this manner, the CL learners may be in a better position to gain a better overall mastery of the language if they concentrate on the written work.

6. Conclusions

In conclusion, knowing how to use the characters when writing essays is significant in helping non-native CL learners learn the language more effectively. In essence, the CL is naturally tough to master, which is compounded by the complexity of its characters. Achievement can only be enhanced by increasing the learners' writing ability using suitable or acceptable strategies, which, in this case, is the ability to combine and use the characters in essay writing. Although this study has a limited number of high proficiency achievers (17 learners) compared to low proficiency achievers (62 learners), it managed to increase the literature on the awareness of the role of the characters in language mastery. It was discovered that assisting CL teachers focus on improving students' abilities to express ideas in their essay writing is essential. Given these circumstances, non-native CL learners need to use relevant literature such as dictionaries and other resources to obtain acceptable knowledge and skills in handling CL characters to express their views.

On that front, discriminant analysis can only provide insights based on the limited data of high proficiency learners. Therefore, more research may be conducted on how to teach non-native CL learners to use proper characters to express their thoughts.

Author Contributions: Conceptualization, S.H.T.; methodology, S.H.T.; software, S.H.T.; validation, S.H.T., B.C.H., Y.F.N., and J.Y.Y.T.; formal analysis, Y.F.N.; investigation, B.C.H. and Y.F.N.; resources, S.H.T.; data curation, S.H.T. and J.Y.Y.T.; writing—original draft preparation, S.H.T.; writing—review and editing, B.C.H. and Y.F.N.; visualization, S.H.T.; supervision, S.H.T.; project administration, B.C.H.; funding acquisition, B.C.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Grant Inisiatif APB (GIA), Universiti Teknologi MARA, grant number 600-TNCPI 5/3/DDF (APB) (001/2021).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to privacy restrictions.

Conflicts of Interest: The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Bi, L.W.T. A study on SRJK (C) Yuk Yi (Sabah, Malaysia) Non-Chinese Language pupils' learning situation (note: A translated title from Chinese Language). Master's Thesis, Minzu University of China, Beijing, China, 2019.
2. Heng, B.C.; Neo, Y.F.; Chan, J.Y. Investigating the attributes of Chinese primary school adoption among the Malay community in Malaysia. *ESTEEM J. Soc. Sci. Humanit.* **2021**, *5*, 153–166.
3. Heng, B.C.; Neo, K.S. Pendidikan Cina di Kalangan Pelajar Melayu Kelantan. In Proceedings of the 3rd Academic Conference of UiTM Kelantan, Kota Bharu, Kelantan, Malaysia, 14–15 December 2005.
4. Tan, J.Y.Y. Malay parents' perspective on admission of their children to Chinese primary school in Kelantan, Malaysia. *J. Arts Sci. Commer.* **2015**, *6*, 26–30.
5. Tan, Y.L.; Ngu, I.T.; Chew, F.P. The phenomenon of increased number of non-native students in under-enrolled Chinese Primary schools in Kuala Pilah. *J. Chin. Lit. Cult.* **2017**, *5*, 79–100.
6. Tan, L.H.; Spinks, J.A.; Eden, G.F.; Perfetti, C.A.; Siok, W.T. Reading depends on writing, in Chinese. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 8781–8785. [CrossRef] [PubMed]
7. Odinye, S.I. The Spread of Mandarin Chinese as a Global Language. Nigerian Journal Online-Published Articles. Available online: https://www.researchgate.net/publication/305639211_The_spread_of_Mandarin_Chinese_as_a_global_language (accessed on 24 March 2022).
8. Agwuele, A. Globalization, dying language and futility of saving them. *Public Culture* **2010**, *16*, 97–118.
9. Statista. The World's Most Spoken Languages. Available online: <https://www.statista.com/chart/12868/the-worlds-most-spoken-languages/> (accessed on 1 March 2022).
10. Foon, W. No Easy End to SPM Chinese Saga. The Star. Available online: <https://www.thestar.com.my/news/nation/2016/03/13/trend-in-weaker-spm-chinese-subject-results-is-worrying> (accessed on 1 March 2022).
11. Kwarteng, M.; Acheampong, D.O. The perceived vocabulary strategies and motivation used by learners of Chinese as a foreign language. *J. World Engl. Educ. Pract.* **2021**, *3*, 22–45. [CrossRef]
12. Tahsildar, M.N. Chinese Language Complexities among International Students in China. *Educ. Q. Rev.* **2019**, *2*, 67–76. [CrossRef]
13. Hardan, A.A. Language learning strategies: A general overview. *Procedia-Soc. Behav. Sci.* **2013**, *106*, 1712–1726. [CrossRef]
14. O'Malley, J.M.; Chamot, A.U.; Stewner-Manzanares, G.; Kupper, L.; Russo, R.P. Learning strategies used by beginning and intermediate ESL students. *Lang. Learn.* **1985**, *35*, 21–46. [CrossRef]
15. Cohen, A.D.; Oxford, J.C. Language Strategy Use Inventory. Available online: https://carla.umn.edu/maxsa/documents/LanguageStrategyInventory_MAXSA_IG.pdf (accessed on 20 February 2021).
16. Gardner, R.C. *Social Psychology And Second Language Learning: The Role Of Attitudes And Motivation*; Edward Arnold: London, UK, 1985.
17. Nel, N.M.; Soezin, K.R.O.G. Factors influencing the acquisition of Mandarin Chinese as a second additional language focusing on phonetics. *Particip. Educ. Res.* **2021**, *8*, 1–27. [CrossRef]
18. Kuo, L.J.; Kim, T.J.; Yang, X.; Li, H.; Liu, Y.; Wang, H.; Park, J.H.; Li, Y. Acquisition of Chinese characters: The effects of character properties and individual differences among second language learners. *Front. Psychol.* **2015**, *6*, 986. [CrossRef] [PubMed]
19. Osborne, C.; Zhang, Q.; Zhang, G.X. Which is more effective in introducing Chinese characters? An investigative study of four methods used to teach CFL beginners. *Lang. Learn. J.* **2020**, *48*, 385–401. [CrossRef]