

Critical Period Constraints In The Second Language Acquisition Of Syntax: The Case Of EFL Arabic Learners Of English

Abdulrahman Alzamil

Assistant Professor of Second Language Acquisition of Syntax, Department of Foreign Languages, Faculty of Arts, Taif University, Kingdom of Saudi Arabia

(Received 11/1/1441H; Accepted for publication 1/4/1441H)

Abstract: Age is believed to be a determining factor in achieving native-like competence in second language (L2) acquisition. The Critical Period Hypothesis (Lenneberg, 1967) states that native competence cannot be attained after puberty. Previous research has reached diverse conclusions regarding the ability of L2 learners to attain the level of native speakers. This study addresses whether native-like competence is attainable by English as a Foreign Language (EFL) learners who began learning English beyond the critical period. Data were gathered from 32 male EFL learners in Saudi Arabia whose first language was Arabic, alongside six English-speaking controls. The Arabic speakers were fourth-year students specialising in English at a Saudi university. Two tests were administered: a) a grammaticality judgement test; and b) an Oxford Quick Placement Test of proficiency. The participants were categorised into four proficiency levels (elementary, lower intermediate, upper intermediate and advanced). The findings revealed that the participants in the advanced group performed similarly to the native English speakers; therefore, EFL learners can achieve native-like competence.

Keywords: Critical period, EFL learners, Arabic, native-like competence, syntax.

قيود الفترة الحرجة على اكتساب اللغة الثانية: دراسة حالة للناطقين باللغة العربية

عبد الرحمن الزامل

الأستاذ المساعد بقسم اللغات الأجنبية، كلية الآداب جامعة الطائف، المملكة العربية السعودية

(قدم للنشر في ١١ / ١ / ١٤٤١هـ، وقبل للنشر في ١ / ٤ / ١٤٤١هـ)

ملخص البحث: يُعتقد أن العمر عامل حاسم في تحقيق كفاءة الناطقين باللغة الأم عند اكتساب اللغة الثانية. تنص فرضية لينبرج (١٩٦٧) عن الفترة الحرجة، على أنه لا يمكن تحقيق كفاءة الناطقين باللغة الأم بعد البلوغ. توصلت البحوث السابقة إلى استنتاجات متنوعة فيما يتعلق بقدرة متعلمي اللغة الثانية على تحقيق مستوى الناطقين باللغة الأم. تتناول هذه الدراسة إمكانية تحقيق متعلمي اللغة الإنجليزية بوصفها لغة أجنبية كفاءة الناطقين باللغة الأم إذا بدأوا تعلم اللغة الإنجليزية بعد الفترة الحرجة. جمعت البيانات من ٣٢ ناطقا باللغة العربية يتعلمون الإنجليزية لغة أجنبية، إلى جانب مجموعة ضابطة تتألف من ستة ناطقين باللغة الإنجليزية. كان الناطقون بالعربية من طلاب السنة الرابعة في قسم اللغة الإنجليزية بإحدى الجامعات السعودية. أُجري اختباران اثنان: (أ) اختبار الحكم على المقبولية النحوية و (ب) اختبار أكسفورد السريع لتحديد مستوى الكفاءة. صنف المشاركون إلى أربعة مستويات (مبتدئ، ومتوسط منخفض، ومتوسط عال، ومتقدم). أوضحت النتائج أن أداء المشاركين من المجموعة المتقدمة كان مماثلا للناطقين باللغة الإنجليزية. وعليه يمكن لمتعلمي اللغة الإنجليزية لغةً أجنبية أن يحققوا كفاءة الناطقين باللغة الأم.

الكلمات المفتاحية: الفترة الحرجة، متعلمو اللغة الإنجليزية لغةً أجنبية، اللغة العربية، كفاءة الناطقين باللغة الأم.

1.1 Introduction

The role of age in second language (L2) native-likeness has been the central focus of numerous researchers over the last 50 years (Slabakova, 2016). The Critical Period Hypothesis is the time when native-like fluency can be achieved (Lenneberg, 1967). This hypothesis is supported by the story of Victor, a 12 year-old boy who was found in the woods and had experienced no human contact. Victor was unable to make a significant progress with regard to his language ability. Lenneberg's critical period assumption does not relate to L2 acquisition, as L2 learners already have native-like competence in one or more languages. Nevertheless, second language acquisition (SLA) researchers extended his assumption to L2 acquisition. Two positions are supported by several researchers regarding the effects of this critical period on L2 acquisition (Slabakova, 2016): a) native-like mastery is not possible after puberty (e.g., Abrahamsson, 2012; Birdsong and Molis, 2001; Bley-Vroman, 1990; DeKeyser, 2000; Granena and Long, 2013; Johnson and Newport, 1989); and b) L2 learners can reach native-like mastery after puberty with exposure to L2 input (e.g., Donaldson, 2011; Montrul and Slabakova, 2003; Slabakova, 2006). Studies testing the Critical Period Hypothesis differed concerning what language components can or cannot be acquired. There is no consensus in the SLA literature on when the critical period ends (Munoz and Singleton, 2011). Researchers in the literature have even proposed different ends for the critical period depending on language components. For example, Ruben (1997) suggests the following critical periods of: a) phonology (up to 12 months); b) syntax (up to 4 years); c) semantics (up to 15-16 years).

Findings in the SLA literature varied regarding whether L2 learners can attain native-like performance. This could be attributed to previous research adopting different methodologies and recruiting participants from diverse first language (L1) backgrounds and settings (EFL vs. English as a Second Language (ESL)). This study aims to examine the ultimate attainment of EFL participants studying in the English department of a university in Saudi Arabia to address the following question:

1. Does age have an impact on the L2 acquisition of English syntax?
2. Can instructed EFL learners whose L1 is Arabic have native-like competence in English syntax?

1.2 Literature review

Researchers who examined the impact of age on the level of success in L2 acquisition have focused

on ESL learners, as they tend to receive ample L2 input. The researchers examined primarily two factors (Birdsong, 2005): a) age of arrival in the L2 country; and b) length of residence. Most studies that examined the two factors found that age is a greater determiner of successful acquisition than length of residence (Munoz and Singleton, 2011). The length of residence can indicate the amount of input L2 learners receive. However, counting the number of years may not be reliable without considering the frequency with which L2 learners use L2 compared with their L1 (Munoz and Singleton, 2011). Since the current study is concerned with the L2 acquisition of grammar, some previous research will be discussed in the following paragraphs.

Johnson and Newport (1989) conducted one of the most influential studies that supports the role played by age in L2 acquisition. The study recruited 46 Korean and Chinese ESL learners living the United States. The participants varied in terms of: a) age of arrival in the United States (3-39 years); and b) how long they had lived in the United States (3-26 years). Their syntactic and morphological knowledge was tested using a grammaticality judgement test comprising 276 sentences. The results revealed that those who arrived in the United States at an early age (between 3 and 7 years) outperformed the other participants; thereby supporting the existence of the critical period. The findings were supported by DeKeyser's (2000) study, which was conducted with 57 Hungarian ESL learners in the United States. The participants' length of residence was 34 years and their age of arrival was between 1 and 40 years. The researcher replicated the grammaticality judgement task designed and administered by Johnson and Newport (1989) but shortened it (200 sentences). The findings supported Johnson and Newport's (1989) study and the existence of a critical period.

Conversely, Montrul and Slabakova's (2003) findings contradict the Critical Period Hypothesis. The study investigated whether 64 Spanish learners (aged 19-56 years), whose L1 was English, could perform at a similar level to the control group of 20 Spanish native speakers (aged 18-33 years). The participants were not residing in a Spanish-speaking context; therefore, the researchers opted to recruit university professors and instructors. Two written tasks were administered to investigate the syntactic and semantic Spanish knowledge of participants. They found that around 30% of participants performed similarly to the native-speaker group. These findings do not support the Critical Period

Hypothesis. Montrul and Slabakova (2003) mentioned that they did not consider the age of arrival in the L2 context because most of their participants were EFL learners. However, some participants were Spanish language instructors who studied in Spanish-speaking countries for an average of 6 years and 2 months; thereby affecting the homogeneity.

Studies that addressed the effects of age on L2 native-likeness differed regarding the criterion of native-likeness. Some used native speakers as the criterion (e.g., Abrahamsson and Hyltenstam, 2008; Montrul and Slabakova, 2003; Saito, 2013). Conversely, a few researchers criticised the use of native speakers (e.g., Davies, 2004; Munoz and Singleton, 2011; Piller, 2002) and promoted the use of very advanced L2 learners as the criterion. I propose that the use of native speakers may be more reliable than the use of advanced L2 learners as the definition of 'advanced' in this context may differ between proficiency tests. Moreover, administering diverse proficiency tests may assign the same learners to different proficiency levels (Hulstijn, 2012).

1.3 Methodology

1.3.1 Participants

The study was conducted on randomly selected 32 male EFL learners in Saudi Arabia who were speaking Saudi-Arabic, and six English-speaking controls. The controls were recruited in the UK and paid £15 for their efforts. The L1 Arabic speakers were fourth-year university-level students in an English Department. One of the sampling procedures followed in the present study was to ensure none of the participants began learning English before the age of six years (the youngest was 10 years old). This is in accordance with Ruben (1997) proposal that the critical period for syntax ceases at the age of four years. To strengthen the sample's homogeneity, none of the participants enrolled on English courses in English-speaking countries. Originally, there were 40 participants. However, eight participants were eliminated due to their exposure to English at a very young age or residence in an English-speaking country. As the goal of the present study was to assess whether L2 learners can achieve native-like performance, the researcher aimed to discover as many advanced learners as possible; however, only three were recruited. Typically, studies examining the Critical Period Hypothesis recruit ESL learners who have spent a specified period in the country in which the L2 is spoken (Larson-Hall, 2008). The present study

is different as the participants were selected on the basis that they had not visited an English-speaking country. The 32 participants were categorised into four proficiency levels according to the Oxford Quick Placement Test (OQPT). Their information is summarised below.

Table 1. Participants' biographical information

	Elementary	Lower intermediate	Upper intermediate	Advanced	NS
Number of participants	15	8	6	3	6
Age	Mean = 23.3 SD = 1.2	Mean = 22.3 SD = .7	Mean = 21.2 SD = .4	Mean = 21.3 SD = .6	Mean = 21.8 SD = 1.2
Age of first exposure	Mean = 11.7 SD = .7	Mean = 11.8 SD = .8	Mean = 11.8 SD = .4	Mean = 12.3 SD = .6	-
OQPT score	Mean = 24.1 SD = 2.9	Mean = 34.8 SD = 2.2	Mean = 41.3 SD = 1.2	Mean = 50.7 SD = 2.1	-

1.3.2 Instruments

Two tests were administered: a) a grammaticality judgement test (GJT), which was used by DeKeyser (2000); and b) an OQPT (Syndicate U.C.L.E., 2001). The GJT administered by DeKeyser (2000) consisted of 196 items. This was a shorter version of the test developed by Johnson and Newport (1989), which comprised 276 items. DeKeyser (2000) shortened the test to make it less tiring. The test comprised 98 pairs, which were not adjacent. Of the 196 items, half were grammatical while the other half were not. The participants were required to decide whether or not they are grammatical. The items were ordered at random, and 11 rule types were tested, as follows:

1. Past tense
2. Plural
3. Third-person singular
4. Present progressive
5. Determiners
6. Pronominalization
7. Particle movement
8. Subcategorisation
9. Yes/no questions
10. Wh-questions
11. Word order

Below are two examples of past-tense items:

1. Sandy fill a jar with cookies last night. _____
2. John sang for the church choir yesterday. _____

1.4 Procedure

Ethical concerns were considered for the participants. All the participants were asked to sign a consent form. They were assured that their personal information would be kept confidential. All the participants completed the OQPT first, and the test was timed (30 minutes). During the following week, all the participants completed the GJT, most of whom finished in around one hour. They were asked if the GJT was too long or demanding and none of them complained. However, some pointed out that it should not be any longer; thus, supporting the decision to administer DeKeyser’s (2000) version.

1.5 Results

The overall results of all the five groups are presented in the following graph. The average score (out of 196) of accurate answers for each group were converted into percentages.

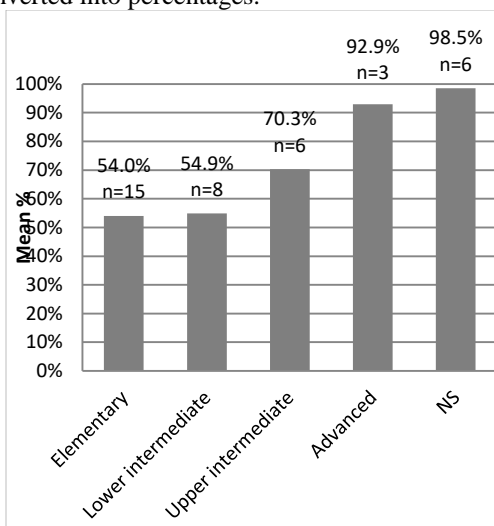


Figure 1. The overall means for the five groups (converted into percentages)

The graph above highlights that the elementary and lower intermediate groups scored the lowest. The percentage scores rise with proficiency.

Non-parametric tests were implemented as the data was not normally distributed ($p < .05$) according to the Kolmogorov-Smirnov test. The table below presents the results for each group (elementary [Elem], lower intermediate [LI], upper intermediate [UI], advanced [Adv] and native speakers [NS]) in each rule type. The table includes mean scores and standard deviations (SDs). Kruskal-Wallis tests were run between the groups and their results are presented in the table.

Table 2. Mean scores, standard deviations and Kruskal-Wallis results for the groups

Rule types	Maximum Score	Elem	LI	UI	Adv	NS	K-W results
Past tense	18	9.8 SD=2.1	10.2 SD=1.5	13.0 SD=1.7	17.0 SD=1.0	18.0 SD=0.4	$p < .01$
Plural	18	9.2 SD=2.3	7.8 SD=0.7	12.3 SD=1.8	16.5 SD=0.6	18.0 SD=0.8	$p < .01$
Third person singular	16	10.1 SD=2.0	9.2 SD=0.7	12.5 SD=1.9	15.0 SD=1.2	16.0 SD=0.4	$p < .01$
Present progressive	12	6.8 SD=1.6	6.5 SD=1.8	8.3 SD=1.6	11.5 SD=0.6	12.0 SD=0.4	$p < .01$
Determiners	14	7.8 SD=2.3	7.8 SD=1.8	10.5 SD=1.9	13.5 SD=1.0	14.0 SD=0.4	$p < .01$
Pronominalization	16	9.0 SD=1.9	8.7 SD=1.4	10.3 SD=1.2	15.0 SD=1.0	16.0 SD=0.5	$p < .01$
Particle movement	16	10.2 SD=2.3	9.5 SD=1.5	11.0 SD=2.3	14.5 SD=0.6	16.0 SD=0.0	$p < .01$
Subcategorization	20	9.5 SD=3.1	10.2 SD=1.5	14.5 SD=1.9	19.5 SD=1.5	20.0 SD=0.5	$p < .01$
Yes/no questions	24	11.6 SD=2.7	11.5 SD=2.4	18.8 SD=3.1	23.0 SD=1.5	23.5 SD=0.5	$p < .01$
Wh-questions	12	8.0 SD=1.5	8.2 SD=1.8	9.3 SD=1.0	11.0 SD=0.0	12.0 SD=0.4	$p < .01$
Word order	30	12.5 SD=3.2	16.7 SD=2.3	20.5 SD=2.4	27.5 SD=2.6	30.0 SD=0.8	$p < .01$

The results reveal that the five groups differed significantly in every rule type, as found by the Kruskal-Wallis tests. Mann-Whitney tests were conducted as post-hoc tests between each group pair. Bonferroni correction adjusted the p value to .005. The adjustment was made to avoid Type 1 error (rejecting a true null hypothesis), as proposed by Field (2013). The first table presents the comparisons between the experimental groups, while the second presents the comparisons between the experimental groups and the native English speakers. Only significant differences ($p < .005$) will be reported.

Table 3. Mann-Whitney results for the comparisons among the experimental groups

Ruletypes	Elem vs. LI	Elem vs. UI	Elem vs. Adv	LI vs. UI	LI vs. Adv	UI vs. Adv
Pasttense	-	$p=.002$	$p=.002$	-	-	-
Plural	-	-	$p=.002$	$p=.001$	-	-
3rd-person singular	-	-	$p=.002$	-	-	-
Present progressive	-	-	$p=.002$	-	-	-
Determiners	-	-	$p=.002$	-	-	-
Pronominalization	-	-	$p=.002$	-	-	-
Particle movement	-	-	-	-	-	-
Subcategorisation	-	$p=.005$	$p=.002$	$p=.001$	-	-
Yes/no questions	-	$p=.002$	$p=.002$	-	-	-
Wh questions	-	-	$p=.005$	-	-	-
Word order	-	$p<.001$	$p=.002$	$p=.005$	-	-

The findings reveal that the elementary group is less accurate than the upper intermediate and advanced groups. The upper intermediate group outperformed the lower intermediate in three rule types. Other comparisons did not reveal significant differences.

Table 4. Mann-Whitney results for the comparisons between the experimental groups and the native English speakers

Rule types	Elem vs. NS	LI vs. NS	UI vs. NS	Adv vs. NS
Past tense	$p < .001$	$p = .001$	$p = .002$	-
Plural	$p < .001$	$p = .001$	$p = .002$	-
3rd person singular	$p < .001$	$p = .001$	$p = .002$	-
Present progressive	$p < .001$	$p = .001$	$p = .002$	-
Determiners	$p < .001$	$p = .001$	$p = .002$	-
Pronominalization	$p < .001$	$p = .001$	$p = .002$	-
Particle movement	$p < .001$	$p = .001$	$p = .002$	-
Subcategorisation	$p < .001$	$p = .001$	$p = .002$	-
Yes/no questions	$p < .001$	$p = .001$	$p = .002$	-
Wh-questions	$p < .001$	$p = .001$	$p = .002$	-
Word order	$p < .001$	$p = .001$	$p = .002$	-

No significant difference was found between the native English-speaker group and the advanced group; however, differences were noted in the other groups.

1.6 Discussion

The argument for the potential for achieving native-like competence beyond the boundaries of the critical period is supported by the present study (e.g., Donaldson, 2011; Montrul and Slabakova, 2003; Slabakova, 2006). In the present study, only the advanced group displayed a similar performance to the native English speakers. I acknowledge the low number of participants in this group (three participants), but this does not undermine the findings. Since there are L2 learners who can produce native-like performance, this may be possible for other L2 learners. For some people, this seems like a predictable finding; however, given several researchers (e.g., Abrahamsson, 2012; Birdsong and Molis, 2001; Bley-Vroman, 1990; DeKeyser, 2000; Granena and Long, 2013; Johnson and Newport, 1989) found that native-like competence cannot be achieved renders the findings of this study interesting. The rationale for recruiting participants at elementary and lower intermediate level was a cautious step as the intention was not to rely solely on the results of the proficiency test (OQPT). This provides each participant, regardless of their proficiency level, an opportunity to participate and be compared with the English native speakers given that the participants were fourth-year students at an English department. It should be mentioned that the OQPT seems to be a reliable test as the participants' proficiency levels were reflected in their performance in the GJT.

The findings are not compatible with two influential SLA studies in the literature (i.e.,

DeKeyser, 2000; Johnson and Newport, 1989). Johnson and Newport (1989) recruited their Korean and Chinese participants based on their exposure to English (minimum of five years). Moreover, the decision to examine the speakers of these two languages was based on their dissimilarity to English. Similarly, DeKeyser's (2000) recruited his Hungarian participants at random. They were then asked about their residence duration in the United States and the age at which they arrived. In other words, neither study administered a proficiency test. If the present study had not administered a proficiency test and relied only on the participants being university-level students who were studying English, the advanced speakers may have been missed; only three of the 32 participants were advanced. That is, the study aimed to find as many advanced speakers as possible based on OQPT. The two studies (i.e., DeKeyser, 2000; Johnson and Newport, 1989) recruited ESL learners because they receive greater opportunities to practise the language and receive considerable input. The present study opted for EFL learners who specialised in English, have never visited an English-speaking country and began their formal study of English after the age of 11. This demonstrates that EFL learners can reach native-likeness, unlike ESL learners in other studies.

The findings support the study conducted by Montrul and Slabakova (2003). The present study is similar in that it administered a proficiency test and the recruitment of participants was not done at random. Some would assume that being a fourth-year student in an English department would mean that you would be an advanced speaker of English. However, this study found that such students may graduate with a low level of proficiency. This is not surprising given that participants in other studies (e.g., DeKeyser, 2000; Johnson and Newport, 1989) who spent a long time in English-speaking countries did not demonstrate native-like performance. In other words, relying on educational backgrounds or length of residence is insufficient to gain knowledge of proficiency levels.

1.7 Conclusion

Previous research varied regarding whether the critical period exists (e.g., Abrahamsson, 2012; Bley-Vroman, 1990; DeKeyser, 2000; Johnson and Newport, 1989) or not (e.g., Donaldson, 2011; Montrul and Slabakova, 2003; Slabakova, 2006). The presented study took a different approach to examining the critical period. EFL students majoring in English were selected for the purpose of determining whether they were able to attain native-

like competence. The study found that L1 Arabic speakers studying English at an advanced level had the capacity to perform at a similar level to native English speakers.

One limitation of the present study is the small sample size for each proficiency level, especially the advanced group. This is due to the fact that it is difficult to find EFL learners with high proficiency levels. It is recommended that future research should recruit a larger number of participants, as this would provide deeper insight into the performance of L2 learners. Moreover, it is recommended to recruit participants from different L1 backgrounds to assess whether they play a role in attaining native-likeness.

References

- Abrahamsson, N. (2012) 'Age of onset and nativelike L2 ultimate attainment of morphosyntactic and phonetic intuition', *Studies in Second Language Acquisition*, 34(2), pp. 187-214.
- Abrahamsson, N. and Hyltenstam, K. (2008) 'The robustness of aptitude effects in near-native second language acquisition', *Studies in Second Language Acquisition*, 30(4), pp. 481-509.
- Birdsong, D. (2005) 'Interpreting age effects in second language acquisition', in J. Kroll and A. de Groot (eds.) *Handbook of Bilingualism: Psycholinguistic Approaches*. pp. 109-127.
- Birdsong, D. and Molis, M. (2001) 'On the evidence for maturational constraints in second-language acquisition', *Journal of Memory and Language*, 44(2), pp. 235-249.
- Bley-Vroman, R. (1990) 'The logical problem of foreign language learning', *Linguistic Analysis*, 20, pp. 3-49.
- Davies, A. (2004) 'The native speaker in applied linguistics', in Davies, A. and Elder, C. (eds.) *The Handbook of Applied Linguistics*. Malden: Blackwell Publishing, pp. 431-450.
- DeKeyser, R. (2000) 'The robustness of critical period effects in second language acquisition', *Studies in Second Language Acquisition*, 22(4), pp. 499-533.
- Donaldson, B. (2011) 'Left dislocation in near-native French', *Studies in Second Language Acquisition*, 33(3), pp. 399-432.
- Field, A. (2013) *Discovering Statistics Using IBM SPSS Statistics*. Sage Publications: London.
- Granena, G. and Long, M. (2013) 'Age of onset, length of residence, language aptitude, and ultimate L2 attainment in three linguistic domains', *Second Language Research*, 29(3), pp. 311-343.
- Hulstijn, J.H. (2012) 'The construct of language proficiency in the study of bilingualism from a cognitive perspective', *Bilingualism: Language and Cognition*, 15(2), pp. 422-433.
- Johnson, J.S. and Newport, E.L. (1989) 'Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language', *Cognitive Psychology*, 21(1), pp. 60-99.
- 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), pp. 35-63.
- Lenneberg, E.H. (1967) *Biological Foundations of Language*. New York: Wiley.
- Montrul, S. and Slabakova, R. (2003) 'Competence similarities between native and near-native speakers: An investigation of the preterite-imperfect contrast in Spanish', *Studies in Second Language Acquisition*, 25(3), pp. 351-398.
- Munoz, C. and Singleton, D. (2011) 'A critical review of age-related research on L2 ultimate attainment', *Language Teaching*, 44(1), pp. 1-35.
- Piller, I. (2002) 'Passing for a native speaker: Identity and success in second language learning', *Journal of Sociolinguistics*, 6(2), pp. 179-208.
- Rubén, R.J. (1997) 'A time frame of critical/sensitive periods of language development', *Acta Otolaryngol*, 117(2), pp. 202-205.
- Saito, K. (2013) 'Age effects on late bilingualism: The production development of /ʁ/ by high-proficiency Japanese learners of English', *Journal of Memory and Language*, 69(4), pp. 546-562.
- Slabakova, R. (2006) 'Is there a Critical Period for the acquisition of semantics?', *Second Language Research*, 22(3), pp. 302-338.
- Slabakova, R. (2016) *Second Language Acquisition*. Oxford: Oxford University Press.
- Syndicate, U.C.L.E. (2001). *Quick Placement Test*. Oxford: Oxford University Press.