



The development of the cognate advantage from elementary to middle school years in French-English bilinguals attending a dual language program in France

Erin Quirk, Cathy Cohen

► To cite this version:

Erin Quirk, Cathy Cohen. The development of the cognate advantage from elementary to middle school years in French-English bilinguals attending a dual language program in France. *International Journal of Bilingual Education and Bilingualism*, 2022, pp.1-16. 10.1080/13670050.2022.2087468 . hal-03701205

HAL Id: hal-03701205

<https://hal.science/hal-03701205>

Submitted on 21 Jun 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

The development of the cognate advantage from elementary to middle school years in French-English bilinguals attending a dual language program in France

Erin Quirk and Cathy Cohen

Abstract

The cognate advantage in bilingual children varies in strength across groups and individuals, in particular on receptive measures (e.g. picture-identification). This variation may be due to children's developing ability to benefit from cognates in such tasks, yet longitudinal studies of this phenomenon, especially with older children, are rare. Using longitudinal and cross-sectional analyses of picture-identification performance in French and English from 37 bilingual children attending a dual language program in France, this study investigates the cognate advantage across a wide range of ages (6 to 13 years old). The influence of children's relative exposure on the cognate advantage is also investigated. We find that significant differences in cognate versus non-cognate performance emerge only after the first years of elementary school. Children show marked growth in late elementary and middle school years, but only on the English task. The strength of the cognate advantage in picture-ID is inversely related to exposure to that language. These findings are discussed in relation to the nature of these children's exposure and properties of cognates in the two languages. Implications for dual language pedagogy are also discussed.

1. Introduction

Cognates are words that share both phonological form and meaning across a language pair (e.g. *tradition* /trə'diʃən/ and *tradition* /tradisjõ/ in English and French) (Proctor & Mo, 2009). Adult bilinguals show a robust advantage on various language tasks with cognate over non-cognate words (e.g. Costa, et al., 2000). From a very young age, children also show sensitivity to cognates in word comprehension (Pérez et al., 2010; Quirk, 2020) and production (Floccia et al., 2018; Sheng et al., 2016). However, the effect on word comprehension in particular shows considerable variation across individual children (Kelley & Kohnert, 2012; Potapova et al., 2016) and studies (e.g., Floccia et al., 2018; Umbel et al., 1992). The cognate advantage is typically measured by comparing performance on cognates and non-cognates in picture-identification (picture-ID) for word comprehension and picture-naming for production. The current study investigates two possible sources of variation in the cognate advantage in picture-ID – children's relative exposure to each language and their age – through cross-sectional and longitudinal analyses.

1.1 Relative exposure and the cognate advantage in vocabulary tasks

The facilitating effect of cognates in picture-naming is stronger in a non-dominant than dominant language in unbalanced bilingual adults and children, and in the language less widely spoken in the community in balanced bilingual children (Poarch & Van Hell, 2012; Rosselli et al., 2014). In young children, relative exposure, or the relative amount of time children are exposed to each of their languages, is a widely used means of determining dominance (Unsworth et al., 2018).

Most studies that investigate the mediating effect of dominance (via relative exposure) on the cognate advantage have focused on L2 vocabulary skills, typically in English. Two studies of L2 English learners, aged 4 to 7, found a stronger cognate advantage in children with less exposure to the language (Pérez, et al., 2010; Robinson Anthony et al., 2020). The effect of dominance on the cognate advantage has rarely been studied in older children and teens; however, one study of Spanish (L1) – English (L2) bilinguals (aged 8-13) found that neither Spanish nor English proficiency scores predicted the cognate advantage in English picture-ID, suggesting that dominance did not mediate the cognate effect (Kelley & Kohnert, 2012). Thus, the influence of dominance on the cognate advantage in picture-ID in older bilingual children remains unclear. The current study addresses this gap by investigating picture-ID performance as a function of cognate status in French and English in elementary to middle-school age bilingual children with varying levels of relative exposure to these languages.

1.2 Development of the cognate advantage

While young adult Spanish-English bilinguals show a small but reliable cognate advantage in picture-ID (Stadthagen-González et al., 2013), children show considerable variability, with roughly 40% of Spanish-English bilingual children showing no advantage at all on cognates versus non-cognates (Potapova et al., 2016; Kelley & Kohnert, 2012). This asymmetry may derive from children's developing ability to recognize and use cognates in word comprehension, which has been reported for cognate identification in the written modality (Hancin-Bhatt & Nagy, 1994; Malabonga, et al., 2008).

While children as young as 6 can identify cognates with minimal instruction (Hipfner-Boucher, et al., 2016), their ability to use cognates in word comprehension may require time to develop. In line with this possibility, a longitudinal study Frisian-Dutch bilinguals aged 5 to 8 found that Dutch-dominant children's sensitivity to cognates grew over time, in particular for cognates with lower phonological overlap (Bosma et al., 2019). Spanish-English bilinguals aged 9 to 11 also showed growth in their ability to recognize cognates, but crucially even in these older children, cognate recognition was not at ceiling (Nagy et al., 1993); thus, development may extend into middle school years.

Sensitivity to cognates may continue to grow through middle school as children become better at recognizing morphological regularities between the two languages (Hancin-Bhatt & Nagy, 1994). However, to date most studies of the cognate advantage in bilingual children have either focused on younger children (e.g., Bosma et al., 2019; Hipfner-Boucher, et al., 2016; Pérez, et al., 2010; Sheng, et al., 2016) or viewed the effect only cross-sectionally (Goriot et al., 2021; Kelly & Kohnert, 2012), possibly introducing other sources of variation. In particular, L2 learners in pre-K to early elementary years may show large shifts in language exposure, which could be confounded with the effect of age on the cognate advantage. Thus, to better isolate the effect of age, variation in children's language exposure should be considered and, if possible, children's development should be viewed longitudinally into late elementary and middle-school years.

2. The current study

The current study investigates the effect of cognate status, i.e., phonological similarity of translation equivalents (TEs), on 37 French-English bilingual children's picture-ID performance in elementary and middle school years in French and English. Children's ability to identify the meaning of words with different degrees of phonological similarity are compared in two age groups cross-sectionally (Time 1) and longitudinally, three years later (Time 2). Children's

current and cumulative relative exposure to English and French was also assessed yearly with language background questionnaires. Children belonged to two age groups: Group 1 were followed from 1st to 4th grade (aged 6 to 9). Group 2 were followed from 5th to 8th grade (aged 10 to 13). Children in both groups ranged widely in their relative cumulative language exposure.

Our research questions were:

- 1) Is bilingual children's accuracy on French and English picture-ID influenced by cognate status?
- 2) Is the relationship between cognate status and accuracy modulated by children's age?
- 3) Is the relationship between cognate status and accuracy modulated by children's relative exposure?
- 4) How does the relationship between cognate status and accuracy on French and English picture-ID change over time in bilingual children from early to late elementary years (Group 1) and from late elementary to middle school years (Group 2)?

Regarding question 1, we expect that a cognate advantage in picture-ID will be present in these children, in particular in Group 2 at both time points and at Time 2 for Group 1. Regarding question 2, we expect that this effect will be stronger in Group 2 generally and for both groups at Time 2. Regarding question 3, we expect the strength of the cognate advantage in picture-ID in a language to be inversely related to the amount of exposure to that language. Regarding question 4, we hypothesize that as children move through elementary and middle school, their cognate advantage on picture-ID will grow.

3. Methods

This study uses a subset of data from a larger five-year longitudinal project, the INEXDEB project (INput et EXpérience dans le DEveloppement Bilingue) (Cohen, 2015), which tracks dual language development in 50 French-English bilinguals, 20 from 1st to 5th grade and 30 from 5th to 9th grade, attending a French state school with a bilingual program, referred to here as the School. INEXDEB aims to better understand emergent bilingual students' linguistic development in order to better respond to their academic needs. In this paper we report only on language exposure and picture-ID data from the 37 children for whom data were available at Time 1 (T1) and Time 2 (T2).

3.1 Context

The School runs from elementary to high school. To attend the School, children take an entrance test to ensure that they have what the School considers to be native or near-native English proficiency. No English as a foreign language programs are provided. Non-Francophone children coming from abroad take classes in French as a second language for two years before integrating fully into the mainstream French class. The standard French national curriculum is taught for three quarters of the week. The last quarter (six hours) is dedicated to teaching in English at native-speaker level. In elementary school, these hours are devoted to the English language and literature components of the British national curriculum. All other subjects are taught only in French. In middle school, four English hours are devoted to English language and literature while two are devoted to the French national curriculum History/Geography program, taught in English by native-speaker teachers. All other school subjects in middle school are taught only in French. A strict separation between French and the other language characterizes this type of bilingual program, resulting in "a form of double monolingualism" (Hélot, 2011, p. 42).

3.2 Participants

Participants were 37¹ French-English bilingual children (Group 1 - 1st grade: $n=19$, $M\ age=6;4$, $SD=;3$; Group 2- 5th grade: $n=18$, $M=10;4$, $SD=;4$). No language disorders or hearing impairment were reported by parents. Participants came from middle-to-high SES homes as measured by parents' years in education, ($M\ maternal\ education=17$, $SD=1.4$; $M\ paternal\ education=16.6$ years, $SD=1.6$). Their family language backgrounds were as follows:

- Ten children had two native English-speaking parents. Four had lived only in France, acquiring English from birth and French from between ages 1 and 3. Six had lived in an English-speaking country at some point before the study began and were first exposed to French between ages 2 and 9.
- Sixteen children had one native English- and one native French-speaking parent and were exposed to both languages from birth. Thirteen had lived only in France or a French-speaking country while three had lived in an English-speaking country prior to the start of the study.
- Eleven children had two native French-speaking parents and had lived in an English-speaking country for between two and five years before returning to France. All children were exposed to French from birth while exposure to English started between ages 2 and 8.

Children's mean English cumulative exposure was 54% ($SD=23\%$) at the study outset and 51% ($SD=21\%$) at the study's end. The distribution of relative exposure is shown in Figure 1.

[INSERT FIGURE 1]

3.3 Materials

Language background questionnaire

Children's French and English cumulative language exposure at the study outset was estimated by parents in an extensive language background questionnaire completed during year one of the study. It enquired about participants' yearly language input in varied contexts inside and outside the home (with each close family member, with the childminder, at daycare, at preschool and elementary school), from birth to the year preceding the study onset on a five-point scale (only English/more English than French/English and French equally/more French than English/only French). Parents were contacted for clarification if their responses were unclear. From this information, an estimated percentage of exposure to each language was generated for each year of the child's life, and then converted to a ratio. For instance, if a child was estimated to have had 65% English exposure and 35% French exposure in a particular year, the ratios would be 0.65 and 0.35, respectively. Summing up yearly readings for a five-year-old child, whose exposure to French and English remained unchanged year on year, her cumulative exposure would be 3.25 years for English and 1.75 years for French (method adapted from Unsworth, 2013). To assess children's current exposure, parents estimated the number of hours their child was in contact with each language in a typical week, first during school time (including weekends), and secondly during the school holidays². The total number of hours spent yearly in each language was then

¹ All participants providing data for the first 4 years of INEXDEB were included except for one child, excluded as she received more than 10% exposure to a third language.

² See Cohen and Mazur-Palandre, 2018, for full details of how current exposure was calculated.

calculated, converted to ratios and added to cumulative exposure readings, providing children's total cumulative exposure at the study outset.

From year two of the study onwards, parents completed a shorter language background questionnaire in which they estimated their child's current exposure as explained above. The resulting figures were then added to the previous year's readings to yield the to-date cumulative exposure in each language.

Picture-ID

Picture-ID performance was assessed yearly using standardized receptive vocabulary tests: *The British Picture Vocabulary Scale* for English (BPVS; Dunn, Dunn, Whetton & Burley, 1987), and *l'Échelle de Vocabulaire en Images Peabody* for French (EVIP; Dunn, Thérault-Whalen & Dunn, 1993). Standard instructions were followed. The researcher (this paper's second author) read a word and the child indicated which of the four pictures on the page corresponded best to it. Standardized and raw scores for the group as a whole are reported in Table 1.

Items are ordered in terms of difficulty, determined via word frequency. Thus, item number is an index of item difficulty with higher numbers indicating higher difficulty. The first 12 items of the BPVS are not included in our analyses because these items were not administered to any participants. Additionally, one BPVS item was omitted because no suitable TE could be found³. Thus, in total 155 BPVS items and 170 EVIP items were analyzed.

[INSERT TABLE 1]

Cognate Status

BPVS and EVIP items were first analyzed for their cognate status, which involved 1) finding a TE, 2) transcribing the word and its TE and in a subset, recording their pronunciation, 3) assessing the phonological similarity of each word pair.

A search of the online Collins French-English Dictionary provided possible translations of all BPVS and EVIP items, selected based on closeness to the meaning shown in the prompt image. If multiple TEs were available after this step, word frequencies and cognate status were considered. If no cognates were in the list of TEs, the word with the closest frequency to its equivalent, given on a 5-point scale in the Collins dictionary entry, was chosen. If the possible TEs included a cognate which did not differ in its frequency from the test item word by more than 2 points, it was preferred.

Words were phonologically transcribed and then checked for accuracy by the second author, a near-native French speaker and native English speaker. British pronunciation was used as this is the dialect spoken by the researcher who administered the picture-ID tasks. A handful of corrections were made to match her pronunciation.

Studies differ in how they define cognate status. First, most studies define cognates categorically. However, given that the cognate advantage varies across degrees of phonological similarity (Bosma et al., 2019), we chose a continuous measure which may more closely reflect how speakers perceive phonological similarity. Secondly, the use of subjective versus objective measures can yield different results for the cognate advantage in picture-ID (Potapova et al., 2016). We used a composite score based on both subjective and objective measures to blend the saliency captured by a subjective measure with the reliability of an automated objective measure.

³ The omitted BPVS item was *terpsichorean*, for which no suitable one-word translation could be found in French.

The subjective measure of phonological overlap is based on ratings of phonological similarity on a scale from 1 to 7, with 1 indicating the lowest level of phonological similarity and 7 the highest⁴. All phonetic transcriptions of French-English TEs ($n=303$) were given a subjective rating of phonological overlap on this scale by the two authors and half of the items ($n=152$) were also rated by an additional 13 adult English speakers with varying proficiency in French (average of 4 on a scale of 5, $SD=1.3$). Via an interface created with the PsyToolkit online platform (Stoet, 2017) participants heard TEs spoken by the same researcher who administered the picture-ID tasks in this study and were asked to assess their phonological overlap on the scale described above. The intra-class correlation coefficient for all 15 raters was 0.77, which is considered good (Koo & Li, 2016). A mean rating was then generated for all items.

For the objective measure, we used a normalized version of Levenshtein distance (Schepens et al., 2013), i.e. the number of substitutions, omissions and additions required to go from one form to another divided by the length of the longer word subtracted from one.

For example, the equivalents *bateau* – *boat* transcribed phonetically in X-SAMPA are bato – b@ut. To arrive at b@ut from bato, one deletes o, substitutes @ for a and inserts u. Thus, the distance is $1-(3/4) = .25$.

Both subjective and objective measures were converted to z-scores and summed to create a composite phonological similarity score, referred to henceforth as *phonological similarity*.

The average subjective rating (scale 1 to 7) for the two tests was similar: BPVS 3.77 ($SD=2.12$); EVIP 3.62 ($SD=2.09$). The average objective measure was also similar: BPVS 0.25 ($SD=0.20$); EVIP 0.24 ($SD=.20$). The subjective and objective measures were highly inter-correlated ($r=.80$, $p<.0001$).

3.4 Procedures

Written informed consent was obtained from parents each year. Data collection sessions were recorded using a high-performance digital recorder (Marantz PMD620MKII). Children were tested individually in a quiet classroom during lunch break by the second author.

Testing sessions lasted around 10 minutes. At least one week separated the French and English picture-ID sessions and their order was counterbalanced, with half of participants randomly selected to start in French and the other half in English. Testing order changed each year. No significant group differences were found between order of testing groups in year one. Children received a book each year to thank them for their participation.

4. Analyses

Prior work has found that cognate status on standardized tests is correlated with item difficulty, with the presence of cognates increasing as difficulty increases (Kelley & Kohnert, 2012), which by design also influences children's accuracy. Indeed, in the BPVS, there is a positive relationship between item difficulty and cognate status ($r=.24$, $p<.01$), although not for the EVIP ($r=.11$, $p=.16$). Thus, in all analyses, the effect of item difficulty, which is indexed by item number, is accounted for.

We ran generalized linear mixed effects regression models in R using the *glmer* function of the *lmer4* package (Bates et al., 2015). The first set of regression models predicts accuracy on BPVS and EVIP at Year 1, with phonological similarity, age, item difficulty and relative exposure

⁴ Instructions were as follows: "How similar do [the word pairs] sound? Try to focus only on the way they sound, ignoring how they are written. Assign them a value from 1 to 7 with 1 being completely different in sound and 7 being completely similar in sound."

as fixed effects and item and participant as random effects. Then in the second set of analyses, longitudinal data from Groups 1 and 2 are analyzed in a similar fashion, but with an added variable, time (T1 and T2).

5. Results

5.1 Research Question 1

The first research question asked whether a cognate advantage could be seen in these children on French and English picture-ID at T1 (1st and 5th grade), and three years later at T2 (4th and 8th grade). To address this question, we ran generalized linear mixed effects regressions predicting accuracy on BPVS and EVIP items (coded as 0 or 1) with cognate status (a continuous measure of phonological similarity), relative exposure to English, and item difficulty (indexed by item number) as fixed effects. Item and participant were random effects. The odds ratios, confidence intervals and *p*-values for these models are shown in Tables 2 and 3. At T1, phonological similarity is a significant predictor of accuracy for Group 2 but not Group 1 for both English and French tests. At T2, phonological similarity is a significant predictor of accuracy for both groups for English but only for Group 2 for French. For both groups and tests, the effect of phonological similarity is stronger at T2. The strongest effect is for English at T2 for Group 2.

[INSERT TABLE 2 and 3]

5.2 Research Questions 2 and 3

We hypothesized that the cognate effect would be modulated by children's age and relative exposure. To address this, we ran the same analyses as above, but instead of splitting the data by age group, we added age group as a fixed effect and tested for an interaction between phonological similarity and age group (RQ2) and between phonological similarity and relative exposure (RQ3). Results are shown in Tables 4 and 5.

[INSERT TABLE 4 and 5]

The cognate advantage was stronger in older children for the BPVS at both time points and the EVIP at T2, but this difference was only marginally significant at T1 for the EVIP ($p=.09$). These interactions are plotted in Figure 2. In English picture-ID at T1 (top left), as phonological similarity increases, accuracy stays roughly the same for Group 1 while it increases for Group 2. In English picture-ID at T2 (top right), both Groups' accuracy increases with increased phonological similarity, but the slope for Group 2 is steeper, showing a stronger cognate advantage (Group 1 OR=3.99 (95% CI 1.68-9.47); Group 2 OR=11.06 (95% CI 3.28-37.30)). In French picture-ID at T1 (bottom left), the effect of cognates is not significantly different across Groups. By T2 (bottom right), however, Group 2 shows a stronger increase in accuracy with increased phonological similarity compared to Group 1, whose accuracy does not change with increased phonological similarity.

[INSERT FIGURE 2]

RQ3: With respect to how the cognate advantage is influenced by relative language exposure, the moderating relationship was in the expected direction for all interactions – as English exposure increases, the advantage on picture-ID weakens (odds ratio less than 1) in English and strengthens in French – however, the interaction at T2 for French was not significant ($p=.11$).

To view the nature of the interaction, the effect of phonological similarity on accuracy at three levels of English exposure – one standard deviation below the mean, the mean, and one standard deviation above the mean – are plotted in Figure 3. At T1 (left), at low levels of English exposure (-1SD), children's accuracy increased steeply with phonological similarity on the English task (top) while at average and high levels of English exposure (+1SD), children showed relatively little change in accuracy as phonological similarity increases. The inverse pattern can be seen for French (bottom). At T2 (right), all children's accuracy increased with phonological similarity, but children with low English exposure again showed steeper gains than other children. On the French task, the pattern is the same as at T1 – the steepest increase in accuracy was seen at high English exposure levels – but the differences are non-significant.

[INSERT FIGURE 3]

5.3 Research Question 4

To investigate changes in the cognate advantage over time, we ran the same models described above but with an additional variable, time (T1 and T2), and with an interaction term for the effect of phonological similarity by time. As seen in Table 6, in the group as a whole, the cognate advantage strengthened over time for the BPVS, but did not significantly change for the EVIP. We also divided the participants into age groups and tested the interaction of time and phonological similarity in the same way to see if changes in the cognate advantage differed across ages. These results are also shown in Table 6. Again, the cognate advantage strengthened only in the BPVS, and Group 2 showed a bigger increase over time than Group 1.

[INSERT TABLE 6]

5.3.1 Controlling for differences in items taken at Time 1 and Time 2

In these analyses, the items included in the data for T1 and T2 overlap only partially because as children age, they move up to higher, more difficult test sets. While the effect of item difficulty is accounted for in our statistical analyses as a fixed effect, there may be other differences between items taken at T1 and T2 that could affect the cognate advantage (e.g., the prevalence of cognate items). To check if the increase in the cognate advantage seen in the interactions for the BPVS above was an artefact of differences between the items tested at T1 and T2, we ran the same interaction analysis with only those items that children responded to at both time points. The same pattern emerged: the interaction between phonological similarity and time was significant for the BPVS (OR=3.05, CI(95%)=1.81-5.15, $p<.001$) but not the EVIP (OR=1.02, CI(95%)=0.46-2.26, $p=.967$).

5.4 Results summary

Regarding RQ1, a cognate advantage was not present in Group 1 at either time point in French, nor in English at T1 (1st grade), but it was present at T2 (4th grade) in English. In contrast, in Group 2, a cognate advantage was present at both time points (5th and 8th grade) in both languages.

The difference in the cognate advantage across age groups was further explored to address RQ2. A moderation analysis revealed the cognate advantage was generally stronger for Group 2, but that in English this difference was marginally significant when children were in 1st and 5th grade, and significant in English and French when they were in 4th and 8th grades.

For RQ3, differently from age, the mediating role of relative exposure on the cognate advantage diminished in French over time, being significant at T1 for both tests but becoming non-significant at T2. In all cases, the advantage increased in strength as relative exposure to the language decreased.

Finally, regarding RQ4, we viewed the development of the cognate advantage longitudinally in all children (1st-4th and 5th-8th grades) and separated by age group. In all analyses, the cognate effect strengthened with time for English but not French. This pattern held even when items were held constant at both time points. In the analyses divided by age group, Group 2 showed a bigger increase in the cognate advantage (for English) than Group 1.

6. Discussion

In this study, we set out to view the development of the cognate advantage in picture-ID in bilingual children of a broad age range spanning early elementary to middle school years. We predicted that children would show a cognate advantage that would increase in strength as they entered the years where significant gains in skills supporting cognate recognition have been reported. Our results partially confirmed these predictions. The cognate advantage in English but not French increased with age, especially in late elementary and middle school years. We also predicted that children's relative exposure to the languages would be inversely related to the strength of the cognate advantage, and our results confirmed this prediction, although differences were non-significant at T2 for French. Below we discuss implications for these findings in more detail.

6.1 The development of the cognate advantage in elementary to middle school years

In our participants, the cognate advantage on picture-ID emerged in early elementary school years and showed the strongest growth from late elementary to middle school years. This late and long development of the cognate advantage may seem surprising given that children as young as 1st grade (age 6) acquiring English and French have shown cognate awareness (Hipfner-Boucher et al., 2016) and even younger children have shown a cognate advantage in picture-naming (Sheng et al., 2016) and picture-ID (Goriot et al., 2021). However, there are several reasons why these children may develop a cognate advantage later than has been reported elsewhere.

First, differences in the tasks used may have influenced results: children in Hipfner-Boucher and colleagues' (2016) study were given brief cognate instruction and then asked to determine if word pairs were cognates while children in this study received no cognate instruction and were asked to identify the meaning of single words. Second, the cognate advantage may vary in nature across language modalities; for example, the skills supporting the cognate advantage in picture-ID, a receptive task, may develop later than for picture-naming, an expressive task. This is because in picture-ID, the cognate advantage can arise when a child uses a known cognate to guess the meaning of its TE. During the study, some children even commented on using such a strategy to infer the meaning of new words. This skill, i.e., connecting known and unknown words via their similarity in form, may develop only after children acquire cognate awareness. However, this type of guessing strategy may be less common in picture-naming: children may be less likely to produce a word by guessing its form based on cognate knowledge. The cognate advantage in picture-naming may depend more on the accessibility of existing lexical entries for cognates vs. non-cognates, which does not depend on cognate awareness. Future research should examine the cognate advantage using

receptive and expressive tasks to explore this possibility, and ideally, using receptive tasks with an online component – e.g., reaction time in a word comprehension task – to explore cross-modality differences while controlling for the effect of a guessing strategy.

The development of the cognate advantage may also differ across language pairs. Studies reporting a cognate advantage in younger children's picture-ID performance have involved children acquiring language pairs with different cognate distributions. For example, in closely related languages such as Catalan and Spanish, Frisian and Dutch and Swedish and German, there is a high number of cognates in words across frequency levels (Bosma et al., 2019; Lindgren & Bohnacker, 2020) and a high number of direct cognates (Bosch & Ramon-Casas, 2014). French and English are estimated to have over 20,000 cognate pairs owing to their shared Latin and Greek origins, roughly 30% of entries in general purpose French dictionaries (Frunza & Inkpen, 2009). However, most high frequency words in English are of German/Scandinavian origin and the vast majority of academic words, acquired only in adolescence, are of Latin/Greek origin (Corson, 1997). Cognates are therefore skewed in their distribution in English, tending to be lower frequency, later acquired words. French words of Latin/Greek origin, however, are more evenly distributed across frequency levels. Confirming this pattern, the cognate status of items on the BPVS showed a significant positive correlation with item difficulty (i.e., a negative correlation with word frequency) while this relationship was non-significant for the EVIP.

Thus, properties of the languages being acquired may affect the development of the cognate advantage. In language pairs that share a large portion of their vocabularies across all difficulty levels, regularities across cognates may be more readily noticeable at young ages. However, cognate recognition skills may continue to develop in older children acquiring language pairs like French and English due to their distribution, which skews towards low-frequency words in English.

6.2 The cognate advantage in French and English

The asymmetrical distribution of French-English cognates may also account for the unexpected finding that the cognate advantage grew exclusively and was generally stronger in English, in all children regardless of relative exposure, in particular in the older age group. This may be because the older children are beginning to acquire the lower frequency words where cognates in English are concentrated. Another possibility is that this pattern derives from the dominance of French at school and in the community. In this bilingual program, only English language and literature are taught in English in elementary school, while History/Geography is added in middle school. Thus, children are exposed solely to French academic language in the remaining subjects, accounting for three quarters of curriculum time. If children's academic vocabulary is stronger in French, they would be more likely to use their cognate knowledge in French to boost their (cognate) performance in English. Future studies might view the cognate effect in French and English in school-aged bilinguals whose school time is spent primarily in English, rather than French, to see if the asymmetry remains. The stronger effect in English may also be due to unmeasured exposure to French in the wider community of these children living in France. However, overall, these children's standard scores in French and English were nearly equivalent at both time points (see Table 1), suggesting overall balanced lexical skills in the two languages. In future studies, these factors could be explored by manipulating the language of the community and schooling, e.g. comparing children in communities where English is also widely spoken, like French-speaking

Canada, with children in France, and comparing children who receive schooling predominantly in English to those schooled predominantly in French.

6.3 Limitations and future directions

There are some limitations to this study. First, because the data were taken from a larger study, cognate status was not manipulated, and therefore could be confounded with other variables such as word frequency. While we address the role of word frequency in our statistical analysis, using mixed effects modeling that includes the random effect of item and the fixed effect of difficulty (based on word frequency), a more robust test of the independent contribution of phonological similarity to children's accuracy would be to match items for frequency and other factors that also may influence accuracy, such as part of speech and context of use. Furthermore, our small sample size decreased our statistical power, and not controlling for relative exposure could have unduly influenced our results. For example, it is possible that a few individuals with highly imbalanced exposure could influence the relative strength of the cognate advantage across languages at the group level. For example, if our participants had on average lower exposure to English than French, this could account for a stronger effect of cognates in English. However, as seen in Figure 1, our participants covered a wide range of relative exposure to the languages and the mean exposure was 54% (SD=23) English at Time 1 and 51% (SD=21) at Time 2; thus, the finding of a stronger cognate effect in English is unlikely to be a result of the specific characteristics of our sample. Furthermore, despite the small sample size and lack of controls, this study replicated findings from larger studies of the effect of relative exposure on the cognate advantage (e.g., Bosma et al., 2019; Robinson Anthony, et al., 2020). We would encourage replicating the study with more participants to provide more robust results. However, collecting longitudinal data for a larger group would take many years in our study context for two reasons. First, bilingual programs of this type are rare in France and second, because there is a high degree of mobility in this population, so many participants would leave the School before the end of the study. Finally, it is possible that our operationalization of language dominance did not capture the relevant variation in skill in these children. Preliminary findings on data from the INEXDEB project (Cohen, 2015) suggest that the relationship between cumulative exposure and language proficiency diminishes as children get older, as they attain a critical mass of data in each language (Gathercole, 2002). Having rich qualitative exposure, through reading for example, may become more relevant than exposure quantity in shaping proficiency. This may account for the diminishing role of exposure in modulating the cognate advantage over time, in particular in French. In other words, it may be that variation in children's language experience at these ages is better captured with qualitative measures and measures of language use. Future research should take a multi-faceted approach to assessing language experience to test the role of various aspects of experience on the cognate advantage at these ages.

7. Conclusions

Our findings taken together point to growth in the cognate facilitation effect on picture-ID from elementary into middle school years, with the fastest growth occurring in late elementary to middle school years. We also found a moderating effect of exposure on the cognate advantage in picture-ID: the lower the exposure to a language, the stronger the cognate advantage in that language. In addition to the moderating role of individual's relative exposure, at the group level, we found a moderating role of language – the effect of cognates was overall stronger and showed more growth in English than French. This asymmetry may

arise due to the nature of these children's language exposure at school, which is predominantly in French, the predominance of French in the wider community, or the distribution of cognates in French and English. Our results do not distinguish these possibilities, but rather highlight the need for future research that takes a systematic approach to comparing cognate effects across language pairs and contexts. Our results also point to the possibility that cognate effects differ across receptive and expressive modalities, a point of variation in cognate effects that is not yet well-understood. These results have implications for dual language education. The curriculum at the bilingual program that these children attend does not include cognate instruction in any formal way, and furthermore, there is a strict separation of languages by context. Our results have shown that even in the absence of formal instruction, children learn to use cognate knowledge to comprehend new words, in particular as they encounter new, primarily academic language in late elementary to middle school years. As children gain experience, perhaps particularly with written and academic language, they can increasingly make the connections needed to bootstrap their comprehension in one language based on their knowledge of the other. This may be particularly useful as children develop reading comprehension skills and become increasingly immersed in the community language, in this case French.

Thus, dual language programs should actively support the development of this skill with explicit instruction and practice across academic areas. Raising children's awareness of cognates should enhance cognate recognition and accelerate cognate acquisition (Nagy et al., 1993). As Rubin observed (1987, p. 16), "once the student's attention is drawn to the relationship, the same student may learn several hundred words in a very short time." Intrinsic to this is the need for classrooms where children are encouraged to draw on their multiple linguistic resources, enabling them to make links between their languages, rather than keeping them apart (García & Wei, 2014; Hélot, 2008).

References

- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1–7. 2014.
- Bosch, L., & Ramon-Casas, M. (2014). First translation equivalents in bilingual toddlers' expressive vocabulary: Does form similarity matter? *International Journal of Behavioral Development*, 38(4), 317-322.
- Bosma, E., Blom, E., Hoekstra, E., & Versloot, A. (2019). A longitudinal study on the gradual cognate facilitation effect in bilingual children's Frisian receptive vocabulary. *International Journal of Bilingual Education and Bilingualism*, 22(4), 371-385.
- Cohen, C. (2015). *INEXDEB (INput et EXpérience dans le DEveloppement Bilingue)*. Lyon, France: Laboratoire Interactions, Corpus, Apprentissages et Représentations.
- Cohen, C., & Mazur-Palandre, A. (2018). Le langage oral en production chez les enfants bilingues : quels liens avec l'exposition ? SHS Web of Conferences 46, 10004. *Congrès Mondial de Linguistique Française - CMLF 2018*. <https://doi.org/10.1051/shsconf/20184610004>.
- Corson, D. (1997). The learning and use of academic English words. *Language Learning*, 47(4), 671-718.
- Costa, A., Caramazza, A., & Sebastian-Galles, N. (2000). The cognate facilitation effect: implications for models of lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26(5), 1283.
- Dunn, L. M., Dunn, L. M., Whetton, C., & Burley, J. (1987). *The British Picture Vocabulary Scale* (2nd Edition.). Windsor: NFER-Nelson.
- Dunn, L. M., Thérault-Whalen, C. M., & Dunn, L. M. (1993). *Echelle de vocabulaire en images Peabody*. Toronto: Psycan.
- Floccia, C., Sambrook, T. D., Luche, C. D., Kwok, R., Goslin, J., White, L., ... Plunkett, K. (2018). Vocabulary of 2-year-olds learning English and an additional language: norms and effects of linguistic distance. *Monographs of the Society for Research in Child Development*, 83(1), 1–135. <https://doi.org/10.1111/mono.12348>
- Frunza, O., & Inkpen, D. (2009). Identification and disambiguation of cognates, false friends, and partial cognates using machine learning techniques. *International Journal of Linguistics*, 1, 1–37.
- García, O., & Wei, L. (2014). *Translanguaging: Language, bilingualism, and education*. London: Palgrave MacMillan.
- Gathercole, V. C. M. (2002). Grammatical gender in bilingual and monolingual children: A Spanish morphosyntactic distinction. In D. K. Oller & R. E. Eilers (Eds.), *Language and literacy in bilingual children* (pp. 207-219). Clevedon: Multilingual Matters.
- Goriot, C., Van Hout, R., Broersma, M., Lobo, V., McQueen, J. M., & Unsworth, S. (2021). Using the Peabody picture vocabulary test in L2 children and adolescents: effects of L1. *International Journal of Bilingual Education and Bilingualism*, 24(4), 546-568.
- Hancin-Bhatt, B., & Nagy, W. (1994). Lexical transfer and second language morphological development. *Applied Psycholinguistics*, 15(3), 289-310.
- Hélot, C. (2008). Bilingual education in France: School policies versus home practices. In C. Hélot & A.- M. de Mejía (Eds.), *Forging multilingual spaces: integrated perspectives on majority and minority bilingual education* (pp. 203-227). Bristol, UK: Multilingual Matters.
- Hélot, C. (2011). Children's literature in the multilingual classroom: developing multilingual literacy acquisition. In C. Hélot & M. Ó Laoire (Eds.), *Language policy for the*

multilingual classroom: Pedagogy of the possible (pp. 42-64). Bristol, UK: Multilingual Matters.

- Hipfner-Boucher, K., Pasquarella, A., Chen, X., & Deacon, S. H. (2016). Cognate Awareness in French Immersion Students: Contributions to Grade 2 Reading Comprehension. *Scientific Studies of Reading*, 20(5), 389-400.
- Kelley, A., & Kohnert, K. (2012). Is there a cognate advantage for typically developing Spanish speaking English-language learners? *Language, Speech, and Hearing Services in Schools*, 43, 191-204.
- Koo, T.K., & Li, M.Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, 15(2), 155-163.
- Lindgren, J., & Bohnacker, U. (2020). Vocabulary development in closely-related languages: Age, word type and cognate facilitation effects in bilingual Swedish-German preschool children. *Linguistic Approaches to Bilingualism*, 10(5), 587-622.
- Malabonga, V., Kenyon, D. M., Carlo, M., August, D., & Louguit, M. (2008). Development of a cognate awareness measure for Spanish-speaking English language learners. *Language Testing*, 25(4), 495-519.
- Nagy, W. E., García, G. E., Durgunoğlu, A. Y., & Hancin-Bhatt, B. (1993). Spanish-English bilingual students' use of cognates in English reading. *Journal of Reading Behavior*, 25(3), 241-259.
- Pérez, A. M., Peña, E. D., & Bedore, L. M. (2010). Cognates facilitate word recognition in young Spanish-English bilinguals' test performance. *Early Childhood Services (San Diego, Calif.)*, 4(1), 55.
- 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.
- Potapova, I., Blumenfeld, H. K., & Pruitt-Lord, S. (2016). Cognate identification methods: Impacts on the cognate advantage in adult and child Spanish-English bilinguals. *International Journal of Bilingualism*, 20(6), 714-731.
- Proctor, C. P., & Mo, E. (2009). The relationship between cognate awareness and English comprehension among Spanish-English bilingual fourth grade students. *TESOL Quarterly*, 43(1), 126-136.
- Quirk, E. (2020). *Beyond Exposure: Markers of English Proficiency in School-aged French-English Bilinguals* (Doctoral dissertation, City University of New York).
- Robinson Anthony, J. J. D., Blumenfeld, H. K., Potapova, I., & Pruitt-Lord, S. L. (2020). Language dominance predicts cognate effects and metalinguistic awareness in preschool bilinguals. *International Journal of Bilingual Education and Bilingualism*, 1-20.
- Rosselli, M., Ardila, A., Jurado, M. B., & Salvatierra, J. L. (2014). Cognate facilitation effect in balanced and non-balanced Spanish-English bilinguals using the Boston Naming Test. *International Journal of Bilingualism*, 18(6), 649-662.
- Rubin, J. (1987). Learner strategies: Theoretical assumptions, research history and typology. In A. Wenden & J. Rubin (Eds.), *Learner strategies in language learning* (pp. 15-30). Englewood Cliffs, NJ: Prentice Hall.
- Schepens, J., Van der Slik, F., & Van Hout, R. (2013). The effect of linguistic distance across Indo-European mother tongues on learning Dutch as a second language. In L. Borin & A. Saxena (Eds.), *Approaches to measuring linguistic differences* (pp. 199-230). De Gruyter Mouton.

- Sheng, L., Lam, B. P. W., Cruz, D., & Fulton, A. (2016). A robust demonstration of the cognate facilitation effect in first-language and second-language naming. *Journal of Experimental Child Psychology*, 141, 229-238.
- Stadthagen-González, H., Gathercole, V. C. M., Pérez-Tattam, R., & Yavas, F. (2013). Vocabulary Assessment of Bilingual Adults: To Cognate or Not to Cognate. In V. C. M. Gathercole (Ed.), *Solutions for the assessment of bilinguals* (pp. 125-145). Bristol: Multilingual Matters.
- Stoet, G. (2017). PsyToolkit: A novel web-based method for running online questionnaires and reaction-time experiments. *Teaching of Psychology*, 44(1), 24-31.
- Umbel, V. M., Pearson, B. Z., Fernández, M. C., & Oller, D. K. (1992). Measuring bilingual children's receptive vocabularies. *Child Development*, 63(4), 1012-1020.
- Unsworth, S. (2013). Assessing the role of current and cumulative exposure in simultaneous bilingual acquisition: The case of Dutch gender. *Bilingualism: Language and Cognition*, 16, 86–110.
- Unsworth, S., Chondrogianni, V., & Skarabela, B. (2018). Experiential measures can be used as a proxy for language dominance in bilingual language acquisition research. *Frontiers in psychology*, 9, 1809.

Figure 1: Proportion of cumulative English exposure at T1 and T2 for all children.

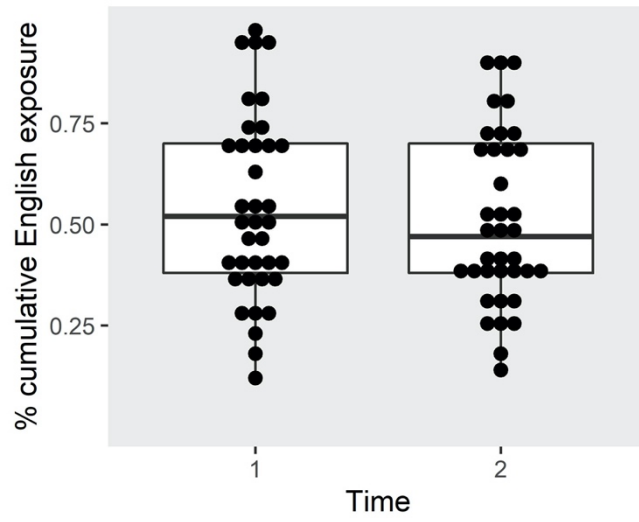


Table 1: BPVS and EVIP picture-ID scores ($M_{\text{test}}=100$, $SD_{\text{test}}=15$) for all ($n=37$) children by age group and time. Age-standardized scores are shown with raw scores in parentheses.

		BPVS T1	EVIP T1	BPVS T2	EVIP T2
Group 1 ($n=19$)	Mean	104 (70)	109 (77)	115 (110)	118 (119)
	SD	11 (12)	25 (28)	10 (11)	15 (16)
	Range	86-124 (50-92)	49-154 (11-125)	92-133 (84-129)	83-138 (80-142)
Group 2 ($n=18$)	Mean	113 (114)	111 (128)	120 (135)	117 (145)
	SD	15 (16)	22 (15)	16 (11)	13 (12)
	Range	77-135 (73-136)	40-136 (97-150)	93-160 (118-155)	90-139 (118-167)

Table 2: Generalized Mixed-Effects Model testing the effect of phonological similarity of TEs on English picture-ID (BPVS)

<i>Coefficient</i>	<i>Odds Ratio</i>	<i>Conf. Int</i> (95%)	<i>p-value</i>	<i>Odds Ratio</i>	<i>Conf. Int</i> (95%)	<i>p-value</i>
<i>Group 1</i>						
	<i>T1</i>				<i>T2</i>	
Intercept	0.32	0.16 – 0.64	0.001	5.03	3.34 – 7.58	<0.001
Phon. Sim.	1.66	0.72 – 3.82	0.233	3.99	1.68 – 9.47	0.002
Item diff.	0.00	0.00 – 0.00	<0.001	0.00	0.00 – 0.00	<0.001
Rel. Exp.	0.82	0.23 – 2.95	0.766	0.27	0.06 – 1.28	0.099
<i>Group 2</i>						
	<i>T1</i>				<i>T2</i>	
Intercept	5.01	3.05 – 8.23	<0.001	26.24	12.21 – 56.40	<0.001
Phon. Sim.	3.06	1.20 – 7.79	0.018	11.06	3.28 – 37.30	<0.001
Item diff.	0.00	0.00 – 0.00	<0.001	0.00	0.00 – 0.00	<0.001
Rel. Exp.	1.25	0.28 – 5.56	0.769	1.09	0.31 – 3.79	0.891

Notes: T1 - G1 - Marginal R^2 / Conditional $R^2=0.445/0.61$; G1 - Marginal R^2 / Conditional $R^2=0.304/0.578$; T2- G1- Marginal R^2 / Conditional $R^2=0.357/0.534$; G1 - Marginal R^2 / Conditional $R^2=0.356/0.628$

Table 3: Generalized Mixed-Effects Model testing the effect of phonological similarity of TEs on French picture-ID (EVIP)

<i>Coefficient</i>	<i>Odds Ratio</i>	<i>Conf. Int (95%)</i>	<i>p-value</i>	<i>Odds Ratio</i>	<i>Conf. Int (95%)</i>	<i>p-value</i>
<i>Group 1</i>						
	<i>T1</i>			<i>T2</i>		
Intercept	1.35	0.86 – 2.11	0.195	7.45	4.61 – 12.02	<0.001
Phon. Sim.	1.96	0.79 – 4.84	0.146	1.08	0.43 – 2.68	0.875
Item diff.	0.00	0.00 – 0.00	<0.001	0.00	0.00 – 0.00	<0.001
Rel. Exp.	0.01	0.00 – 0.07	<0.001	0.10	0.03 – 0.35	<0.001
<i>Group 2</i>						
	<i>T1</i>			<i>T2</i>		
Intercept	7.72	4.45 – 13.37	<0.001	71.52	22.72 – 225.12	<0.001
Phon. Sim.	2.30	1.09 – 4.85	0.028	3.19	1.20 – 8.49	0.020
Item diff.	0.00	0.00 – 0.00	<0.001	0.00	0.00 – 0.00	<0.001
Rel. Exp.	0.24	0.07 – 0.82	0.023	0.50	0.06 – 2.13	0.259

Notes: T1 – G1: Marginal R^2 / Conditional $R^2=0.205/0.511$; G1: Marginal R^2 / Conditional $R^2=0.254/0.413$; T2 – G1: Marginal R^2 / Conditional $R^2=0.193/0.417$; G1: Marginal R^2 / Conditional $R^2=0.273/0.497$

Table 4: Generalized Mixed-Effects Model testing the effect of phonological similarity of TEs on English picture-ID with age groups as a predictor and interactions between exposure, age and phonological similarity

<i>Coefficient</i>	<i>Odds</i>	<i>Conf. Int</i>	<i>p-value</i>	<i>Odds</i>	<i>Conf. Int</i>	<i>p-value</i>
	<i>Ratio</i>	<i>(95%)</i>		<i>Ratio</i>	<i>(95%)</i>	
	<i>T1</i>			<i>T2</i>		
Intercept	0.74	0.49 – 1.11	0.147	5.88	3.67 – 9.41	<0.001
Phon. Sim.	1.78	0.84 – 3.78	0.133	5.12	2.18 – 12.06	<0.001
Age group	6.51	3.63 – 11.68	<0.001	2.78	1.61 – 4.78	<0.001
Rel. Exp.	1.32	0.48 – 3.66	0.587	0.66	0.25 – 1.74	0.399
Item diff.	0.00	0.00 – 0.00	<0.001	0.00	0.00 – 0.00	<0.001
Phon. Sim. x	3.35	1.55 – 7.25	0.002	2.94	1.30 – 6.66	0.010
Age group						
Phon. Sim. x	0.15	0.04 – 0.61	0.008	0.19	0.05 – 0.73	0.016
Rel. Exp.						

Notes: T1 - Marginal R² / Conditional R²=0.336/0.575; T2 - Marginal R² / Conditional R²=0.366/0.594

Table 5: Generalized Mixed-Effects Model testing the effect of phonological similarity of TEs on French picture-ID with age groups as a predictor and interactions between exposure, age and phonological similarity

<i>Coefficient</i>	<i>Odds</i>	<i>Conf. Int</i>	<i>p-value</i>	<i>Odds</i>	<i>Conf. Int</i>	<i>p-value</i>
	<i>Ratio</i>	<i>(95%)</i>		<i>Ratio</i>	<i>(95%)</i>	
	<i>T1</i>			<i>T2</i>		
Intercept	1.53	0.98 – 2.40	0.063	9.29	5.70 – 15.13	<0.001
Phon. Sim.	2.28	1.19 – 4.39	0.014	1.80	0.77 – 4.18	0.174
Age group	3.97	2.19 – 7.19	<0.001	3.07	1.80 – 5.23	<0.001
Rel. Exp.	0.06	0.02-0.19	<0.001	0.12	0.04-0.34	<0.001
Item diff.	0.00	0.00 – 0.00	<0.001	0.00	0.00 – 0.00	<0.001
Phon. Sim. x	2.19	0.88 – 5.41	0.091	4.30	1.74 – 10.61	0.002
Age group						
Phon. Sim. x	14.11	2.72 – 73.34	0.002	3.41	0.76 – 15.25	0.108
Rel. Exp.						

Notes: T1 - Marginal R^2 / Conditional $R^2=0.223/0.498$; T2 - Marginal R^2 / Conditional $R^2=0.235/0.488$

Table 6: Interaction of the effects of phonological similarity and time on accuracy in English and French picture-ID

Age group	Odds Ratio	Conf. Int (95%)	p-value	Odds Ratio	Conf. Int (95%)	p-value
BPVS				EVIP		
Both	2.79	1.66 – 4.70	<0.001	1.16	0.63 – 2.12	0.627
Group 1	3.09	1.45 – 6.59	0.003	0.96	0.41 – 2.23	0.918
Group 2	4.42	2.07 – 9.41	<0.001	1.31	0.55 – 3.12	0.536

Notes: BPVS- Both: Marginal R^2 / Conditional $R^2=0.326/0.544$; G1: Marginal R^2 / Conditional $R^2=0.372/0.536$; G1: Marginal R^2 / Conditional $R^2=0.287/0.559$; EVIP – Both: Marginal R^2 / Conditional $R^2=0.173/0.431$; G1: Marginal R^2 / Conditional $R^2=0.191/0.447$; G1 - Marginal R^2 / Conditional $R^2=0.170/0.373$