Title: Formal versus informal L2 learning: How do individual differences and word-related variables influence French and English L2 vocabulary learning in Dutch-speaking children?

Author 1: Vanessa De Wilde
Affiliations Author 1: Department of Translation, Interpreting and Communication, Ghent University; Artevelde University of Applied Sciences Ghent
Author 2: Marc Brysbaert
Affiliation Author 2: Department of Experimental Psychology, Ghent University
Author 3: June Eyckmans
Affiliation Author 3: Department of Translation, Interpreting and Communication, Ghent University

Address for correspondence:
Vanessa De Wilde
Groot-Brittanniëlaan 45
9000 Ghent
Belgium
vanessa.dewilde@ugent.be

To be published in Studies in Second Language Acquisition

Acknowledgements: We thank Emma Boone and Elise Vanhoecke for help with testing. We thank the schools, the teachers, the pupils and their parents for participating in the project.

Formal versus informal L2 learning: How do individual differences and word-related variables influence French and English L2 vocabulary learning in Dutch-speaking children?

Abstract

A second language can be learned inside and outside the classroom. In this study we investigated the English and French vocabulary knowledge of 110 Dutch-speaking children (age 10-12), who received 100 hours of instruction in French, whereas their contact with English came from out-of-school exposure only. We examined the role of individual differences (out-of-school exposure and gender) and word-related variables (cognateness, frequency and language). The children completed a receptive vocabulary test in English and French and filled in a questionnaire. The results showed that the children had a larger vocabulary knowledge in English than in French, illustrating the power of contextual language learning. Word learning was influenced by the amount of exposure, word frequency and cognateness. Additionally, English words were easier to learn than French words for the participants we tested. Our results point to the need for out-of-school exposure to supplement language learning in the classroom.

Keywords: contextual word learning; out-of-school exposure; gender; cognates; frequency

Second language (L2) acquisition happens in two broad contexts: (1) Through repeated interactions with individuals speaking the target language, and (2) by means of L2 teaching in schools. Although the two contexts are sometimes seen in opposition, in practice they are largely complementary. Language learning (whether first or second language) involves both interaction with individuals talking the language, and storage/retrieval of information in the brain of the language learner.

Usage-based approaches to language learning assume that language learning emerges from the interaction of cognitive learning mechanisms with the linguistic input and stress the importance of interactions with individuals speaking the target language (Ellis,1998; Firth & Wagner, 1997). A term used for language learning as a by-product of communication is 'incidental language learning' (Hulstijn, 2003). A problem with this term, however, is that it gives the impression that language learning happens automatically, without any intention (or effort) to learn. Elgort and colleagues (2018) proposed the term 'contextual language learning' as a better alternative to refer to all types of language learning from context, independently of whether acquisition is intentional. We will use the term proposed by Elgort et al. throughout this article. A meta-analysis by Uchihara et al. (2019) showed that contextual language learning is ubiquitous, provided that learners are exposed repeatedly to the input.

A limitation of usage-based learning is that many L2 learners are not in close contact with individuals speaking the language they want to learn. In addition, L2 acquisition has to compete with L1 knowledge if the two languages use different expressions for the same idea. To tackle these issues, many countries include formal L2 teaching in the education curriculum. Children are brought in contact with a non-native language of interest, in the hope that this will be the stepping stone for future interaction and communication. It is assumed that languages can be learned in the same way as other skills: through explicit explanation, practice, testing and performance feedback (Doughty & Williams, 1998). Another aim of good (language) teaching

is to give children equal opportunities, independent of their background (Haycock, 2011). It is generally accepted that formal, classroom instruction is not enough for fluent language use (Bybee & Hopper, 2001; Ellis, 2002). As a result, language learners are encouraged to take part in activities exposing them to the L2 outside the classroom.

A question that to our knowledge has not yet been addressed, is how out-of-school contextual word learning compares to L2 learning in a formal classroom setting, particularly at the first stages of language acquisition. In the present study we profit from the Belgian context to start answering that question.

1. The Flemish Context

There are three official languages in Belgium: Dutch (spoken in the Northern half of the country), French (spoken in the Southern half), and German (spoken in a region on the border with Germany). In the Dutch-speaking region (Flanders), it has been decreed that learning French is a compulsory part of primary education. Children in primary education are expected to receive 120 hours of French classes in the final two years of primary school. Schools are allowed to start earlier with French classes and they can also introduce English and German through playful language activities at an earlier age but at the time of writing there are only few schools that structurally offer foreign language education in primary schools apart from the compulsory French classes. This means that foreign language classes in Flanders start relatively late compared to other European countries (Enever, 2011). Furthermore, the first foreign language to be taught is French. This also constitutes a difference with many other countries where English is the first foreign language to be taught.

When looking at the cultural context, however, we see that English is far more prominent in Flanders than French, for example in the media and in advertising. Radio stations play a lot of English music and television programs (mainly English-speaking) are subtitled rather than

dubbed. Children also have a lot of exposure to English through gaming, social media and online watching (e.g. YouTube) (De Wilde et al., 2020a; Puimège & Peters, 2019). As a result, children often have a lot of exposure to English before the start of formal English lessons. Because culture (television, radio, newspapers, ….) is regionalised, Belgian children in one language group have little contact with the other language, unless they actively seek it. There is a higher eagerness among Dutch-speaking children to learn English than French (Housen et al., 2001). [1]

A recent study by Peters et al. (2019) investigated French and English receptive vocabulary knowledge in Flemish adolescents in three age groups: 12-14-year-olds, 14-16-year-olds and 18-21-year-olds who had received a minimum of four, six and nine years of French classes and one, three and six years of English classes respectively. The study found that even though the learners received far more hours of instruction in French, they knew more English words than French words.

In the present study we compared vocabulary knowledge of English and French at an earlier moment when children have the opportunity of contextual English learning without any formal classroom instruction, and receive classroom teaching of French without (much) exposure to the language outside the classroom. If the English advantage observed in Peters et al. (2019) only emerges after formal education starts, this would indicate that contextual learning without supporting school instruction is largely inefficient. In contrast, if better English mastery is observed already before formal instruction starts, we have evidence that contextual learning of a non-dominant language is productive in the absence of explicit support.

Ideally, our assessment of language proficiency could involve all aspects of language use. Hulstijn (2011), for instance, defined language proficiency as the extent to which an individual possesses the linguistic cognition necessary to function in a given communicative situation in a given modality (listening, speaking, reading, writing). This involves knowledge of vocabulary

(including multiword expressions), grammar, and pragmatic considerations (which words can be used in in which contexts). A full assessment is difficult to realise, however, within the confines of most studies (including the present one). Luckily, vocabulary knowledge correlates highly with other aspects of language use, certainly at the early stages of L2 acquisition. For the population of children tested, we observed correlations of .7 between receptive auditory vocabulary knowledge and listening comprehension, reading and writing performance, and scores on a speaking test (De Wilde et al., 2020a). So, information about receptive auditory vocabulary mastery is a good first proxy of language proficiency in general.

2. Individual Differences: Out-of-school Exposure and Gender

2.1.Out-of-school exposure

There are several ways in which children can have out-of-school exposure to a foreign language, even when there is no teacher to guide activities building on classroom instruction (De Wilde et al., 2020a; Kuppens, 2010; Lindgren & Muñoz, 2013; Puimège & Peters, 2019; Sylvén & Sundqvist, 2012).

Traditional research on out-of-school learning has mainly looked at the effects of reading on word learning. Many studies have shown that reading is beneficial for word learning, both in L1 and in L2, as learners are able to pick up new words through reading (see Ford-Connors & Paratore, 2015 for a review of the evidence concerning L1 reading). A study by Elley and Mangubhai (1983) with young L2 learners showed that reading stories positively contributes to language learning. Horst et al. (1998) investigated L2 word learning (meaning recognition and word association) from extensive reading. The participants read and listened to a full-length novel. The results for the meaning recognition test and the word association test showed that word learning occurred through extensive reading/listening, with sizable learning gains after eight or more encounters with the word. Webb (2007) and Pellicer-Sánchez and Schmitt (2010)

investigated whether different aspects of vocabulary knowledge (e.g. receptive knowledge of meaning and form, productive knowledge of orthography, etc.) could be acquired through extensive reading. The studies again found that contextual L2 word learning through reading does occur but also found that some aspects (e.g. receptive orthographic knowledge) were easier to learn than others (e.g. productive knowledge of meaning and form). The results also showed that the number of encounters with the words influenced the test scores. Although reading is one of the most researched ways of contextual word learning, it does not apply to the Dutch-speaking primary school children we investigated. For them, reading in English is next to non-existent (De Wilde et al., 2020a; Peters, 2018).

Another type of exposure that has been investigated in relation to contextual language learning is watching television (with or without subtitles). A study by Lindgren and Muñoz (2013) compared language learning in children (age 10-11) from 38 schools in seven European countries and identified watching television as the most important out-of-school exposure variable for children's reading and listening skills. Two studies by Koolstra and Beentjes (1999) and d'Ydewalle and Van de Poel (1999) showed that children picked up new words from watching a short, subtitled film. Research by Kuppens (2010) also showed that watching television positively affects vocabulary knowledge (measured by Dutch-to-English and English-to-Dutch translation of frequent chunks). Puimège and Peters (2019) looked at the influence of watching television on young learners' vocabulary knowledge (meaning recognition and meaning recall). They found that watching television was an important predictor in the meaning recognition test. One of the arguments why TV watching contributes to vocabulary learning, is that this type of input provides large amounts of spoken, authentic input (Webb, 2015). Furthermore, TV watching is multimodal and offers visual support which may facilitate learning (Bisson et al., 2014).

A third type of out-of-school exposure is listening to music. This type of exposure contributed positively to listening and reading skills in the study by Lindgren and Muñoz (2013) but far less so than watching television. De Wilde et al. (2020a) and Peters (2018) reported that listening to English music did not contribute to contextual vocabulary learning. De Wilde et al. (2020a) even found a significant, though small, negative effect of listening to English music in a meaning recognition test. A possible explanation may be that listening to songs does not imply that the lyrics are understood. Furthermore, this type of input is unimodal. Pavia et al. (2019) found that contextual vocabulary learning did occur when listening to L2 songs. They found significant gains in spoken form recognition and collocation recognition but, again, not for form-meaning connection.

The study by De Wilde et al. (2020a) found that three other types of exposure were more beneficial for language learning than the types traditionally examined: gaming, using social media and speaking English. These types of exposure all offered opportunities for interaction. Several other studies also found that a foreign language (mainly English) can be picked up when learners are gaming. Sylvén and Sundqvist (2012) found that in a group of Swedish 11-12-year-olds, frequent gamers outperformed moderate gamers who in turn outperformed non-gamers on meaning recognition and form recall tests. Similar results were reported by Hannibal Jensen (2017) in a study investigating English receptive vocabulary knowledge in Danish children (8 and 10 years old). Puimège and Peters (2019) also found that gaming positively contributed to learners' vocabulary knowledge (meaning recognition and meaning recall). González Fernández and Schmitt (2015) found that using social media (together with other types of exposure) contributed to the learning of collocations through contextual language learning.

2.2.Gender

Studies on the role of gender in language learning have produced mixed results. Jaekel et al. (2017) ran a longitudinal study involving over 5000 German young learners of English. They looked into the role of gender and found that, in a formal context, girls performed significantly better on a test measuring receptive skills (listening and reading comprehension). Courtney et al. (2015) investigated French learning in over 250 English learners (age 10-12 years) and found that girls outperformed boys on two speaking tasks, a sentence repetition task and a picture-description task. Interestingly, the effect of gender disappeared when other factors such as L1 literacy were taken into account.

Several studies investigating contextual vocabulary learning through-out-of-school exposure also looked at the role of gender. Sylvén and Sundqvist (2012), who investigated vocabulary learning through gaming, found that boys knew more English words than girls. The authors stressed that this finding might not be due to gender but to the fact that boys were more frequent gamers, and thus had more exposure to English, which might lead to a larger vocabulary. Hannibal Jensen (2017) found that boys and girls performed equally well on a meaning recognition test but that there was an effect of gaming in the boys' scores and that this effect was absent for girls. Hannibal Jensen suggests that girls benefit more from input in the classroom (cf. findings by Courtney et al. (2015) and Jaekel et al. (2017) concerning L2 learning in the classroom) and boys compensated through gaming, resulting in similar vocabulary test scores for both groups.

Puimège and Peters (2019) found that boys have higher scores on meaning recognition and meaning recall tests than girls and they also attributed the advantage to a higher exposure in boys. Similar results were found for English receptive vocabulary knowledge by Peters et al. (2019). Their analysis showed a positive relation between gaming and vocabulary knowledge. It also showed that boys are more frequent gamers but there was no direct link between gender

and vocabulary knowledge. The boys' greater vocabulary knowledge was thus an effect of their gaming habits and not of their gender. For French, the 'gaming'-effect for boys was absent as the participants did not tend to game in French.

Overall, it seems that in contextual language learning conditions boys tend to outperform girls whereas in a classroom context it is often observed that girls outperform boys. Researchers have also found that these differences can often be explained by other factors. In the present study learners' receptive vocabulary knowledge in two foreign languages is investigated, one language is a school language, the other language is only encountered through out-of-school exposure. We want to look into the possible gender differences in these two learning contexts.

3. Word-related differences

3.1.Cognateness

A study by Van der Slik (2010) showed the importance of cognate linguistic distance in language learning. Cognate linguistic distance is a measure based on the number of cognates in two given languages (cognates are words with similar form and meaning in target languages, such as apple-appel in English and Dutch). Van der Slik's study, which investigated Dutch learning in nearly 6000 learners with different language backgrounds, found that cognate linguistic distance explained about 20% of the variation between learners with different mother tongues. Lindgren and Muñoz (2013) looked into the influence of out-of-school exposure, parents and cognate linguistic distance on young learners' reading and writing skills and found that the strongest predictor for reading and listening scores was cognate linguistic distance.

Research into receptive multilingualism, which is people's ability to communicate in their own language with people speaking a closely related but unknown language (Gooskens et al., 2018) has also shown the importance of cognates in language learning. One of the strategies speakers rely on, is cognate guessing, which is trying to figure out the meaning of a word in an

unknown language based on similarities with words in the native language (Vanhove & Berthele, 2015).

Studies investigating contextual vocabulary learning consistently showed that cognates are easier to learn. Already in 1985, Palmberg ran a study with young learners in Sweden, which showed the facilitative effect of cognates in word learning. Peters and Webb (2018) investigated the effect of cognateness on contextual vocabulary learning through viewing television. The authors found that the chances of learning a cognate were eight times higher for meaning recall and 2.5 times higher for meaning recognition. A similar observation was made by Vidal (2011), especially when listening to spoken input. Puimège and Peters (2019) and De Wilde et al. (2020b) found that young learners heavily rely on cognates when learning new words from out-of-school exposure. Finally, Muñoz et al. (2018) compared L2 English learning in young Spanish and Danish learners and found that the Danish learners profited from the fact that English and Danish share more cognates than English and Spanish. The Danish learners scored as high on a meaning recognition test of English as the Spanish learners, even though the Spanish learners had completed several years of instruction whereas the Danish learners had not received any instruction at the start of the study.

3.2.Word frequency

Another variable which has been widely studied and proven to influence word learning is frequency. People know more high-frequency words than low-frequency words (Brysbaert et al., 2018). Studies investigating young learners' contextual word learning through out-of-school exposure also showed an effect of frequency. Puimège and Peters (2019) found that more frequent words were known better, both in meaning recognition and in meaning recall. De Wilde et al. (2020b) did not find a main effect of frequency but the authors found a frequency effect for the more proficient learners.

4. Aims and research questions

In the present study we aim to investigate young learners' L2 vocabulary learning in two languages. One of the languages, English, is not instructed in school but is considered to be a desirable lingua franca. The other language, French, is one of the official languages in Belgium and is taught from primary school onwards. In this context, English word learning only happens through contextual language learning, whereas opportunities for learning French happen mainly in the classroom (although children can in principle easily search for out-of-school exposure). We will look into the receptive vocabulary knowledge (meaning recognition), the differences in out-of-school exposure and the influence of several individual differences and word-related variables on word learning in young learners for these two languages. The results will shed light on the role of contextual word learning in young learners' vocabulary development.

The research questions of this study are:

(1) How does receptive vocabulary knowledge of children at the end of primary school compare for a language they have been learning for 1.5 years in school (French) to a language they are interested in but which has not yet been instructed in school (English)?

(2) What are the differences in the children's out-of-school exposure to both languages?

(3) What is the role of individual differences (exposure and gender) and word-related variables (cognateness, frequency, and language) on children's L2 vocabulary learning in both languages?

(4) Which types of out-of-school exposure are most effective for vocabulary learning in both languages?

Regarding the first research question, one hypothesis is that learning English in an informal setting would lead to similar learning gains as two hours per week of formal instruction in French. This result would be in line with what was found by Muñoz et al. (2018) who found similar vocabulary knowledge of English in Spanish children after several years of instruction compared to Danish children at the start of instruction but with more out-of-school exposure to English. Alternatively, it could be that learning English in an informal setting leads to larger learning gains than 2 hours of formal instruction per week in French, as suggested by Peters et al. (2019) for children of an older age. Finally, it could be that out-of-school exposure does not have much effect, if not supported by formal instruction. In that case, we would expect the children to know more French after two years of teaching.

With regards to the second research question, we assume that children will have more out-of-school exposure to English than to French outside school for the reasons outlined above. Still, the exposure to French is unlikely to be fully absent and is encouraged by most teachers.

Concerning the third and fourth research questions, we expect that the amount of exposure will positively affect receptive vocabulary knowledge as was demonstrated in previous research (De Wilde et al., 2020a; Kuppens, 2010; Lindgren & Muñoz, 2013; Peters et al.; 2019; Puimège & Peters, 2019). We expect that gaming, using social media, speaking and reading English/French will positively impact meaning recognition. The role of listening to music and watching television is less clear as previous studies have shown mixed results. With regards to gender, we aim to investigate whether there is a difference in word learning between boys and girls since research on the matter has suggested that the gender differences are not gender differences per se but are due to other factors which are more or less present in different genders such as exposure through gaming or conscientiousness in school (Courtney et al., 2015; Peters et al., 2019). We also expect positive effects of cognateness and word frequency (De Wilde et al., 2020b; Lindgren & Muñoz, 2013; Puimège & Peters, 2019).

5. Method

5.1.Participants

The participants for this study were 128 children who were all in the last year of primary school in Flanders, Belgium. The children came from five schools (seven class groups). Although the sample may not be fully representative for the population, we made an effort to include schools with different profiles. Two of the schools were located in a rural area and three schools in an urban area. One of the three city schools had a large multilingual population (>60% of the pupils were from a multilingual background). 70% of the children reported they spoke Dutch at home (n = 91), which is also the language of instruction in all the schools. 30% of the children reported they spoke another language with at least one of their parents. 18 children reported that this language was either French or English. These 18 children were left out of the analyses, bringing the total number of participants to 110.[2] The majority of the children were 11 years old (n=80), 12 children were 10 and 18 children were 12 years old when the tests took place. 57 girls and 53 boys took part in this study. All the children in this study had received about 1.5 years of formal French instruction (approximately 100 hours) and no formal instruction in English. The national curriculum for French in Flemish primary schools focuses on communicative skills. Teachers report they mainly focus on vocabulary development, speaking, listening and reading skills and spend less time on writing and grammar (Vlaamse Overheid, 2018).

5.2.Instruments and procedure

5.2.1. Vocabulary test

The children's English vocabulary knowledge was measured with the Peabody Picture Vocabulary Test (PPVT 4, form A, Dunn & Dunn, 2007). In this meaning recognition test, the children hear a word and see four pictures. They have to match the word with the correct picture.

The children were tested on the first 120 items of the test, ten sets of 12 items each. The test was started from the beginning up to set 10 which is the starting set for children aged 11-12 years whose native language is English. To make sure the items were pronounced in the same way for all children, we used a recording. The items in the recording were pronounced by a highly proficient L2 speaker of English.

French receptive vocabulary knowledge was tested with the Echelle de Vocabulaire en Images Peabody (EVIP, Dunn et al., 1993). This is the French equivalent of the PPVT with comparable developmental scores when taken into account that the EVIP has fewer items than the PPVT (170 vs. 228). The participants heard the first 104 items (which is the starting item for children age 12 who are native speakers of French). The procedure is the same as for PPVT: after hearing the word, the participants choose the picture depicting the item. As with the PPVT, the children listened to a recording of the different test items, to ensure the same pronunciation. The items in the recording were pronounced by a highly proficient L2 speaker of French. 13 English and French items overlapped in meaning. Both tests had a high reliability for the group tested: PPVT (α = .92) and EVIP (α = .82).

The Peabody vocabulary tests were developed to test L1 vocabulary knowledge but they are often used to measure L2 vocabulary as well (e.g. De Wilde et al., 2020a; Hannibal Jensen, 2017; Muñoz et al., 2018). Goriot et al. (2018) studied whether the PPVT could be used to measure L2 vocabulary knowledge. The authors found this was possible but researchers should take L1 influences into account. Since the order of acquisition might be different for L1 and L2, we thus did not use the procedure suggested in the manual (which suggests stopping the test when the learner makes eight mistakes in a set of 12 for the English test, or eight consecutive mistakes for the French test) but we tested all the items both for the English and French vocabulary test. The test administration for both the English and the French test took about 30 minutes. The tests were administered in the classroom as previous studies have shown

this type of test is suitable for group administration in this age group (e.g. De Wilde et al., 2020a: Puimège & Peters, 2019). The French and English tests were done on different days to avoid fatigue. The participants were allowed to listen to the target items multiple times.

5.2.2. Questionnaires

A questionnaire was filled in by the children in the classroom. The questionnaire has been used in previous studies looking into out-of-school exposure and language learning in children (De Wilde et al., 2020a). The children filled in two identical questionnaires; one for English, the other for French. The questionnaires focused on daily exposure to different types of media (watching television with and without subtitles, listening to music, reading, using social media, gaming, and speaking) in English and French, the children's language background, their attitude towards the target language, etc. The questionnaires were completed in Dutch, the language of instruction. The questionnaire is available at https://osf.io/5y4x3/ and https://www.iris-database.org/.

5.2.3. Word-related variables

To determine cognateness, we measured orthographic and phonological similarity between the French and English vocabulary items and their Dutch translation equivalents. The similarity was operationalized by calculating Levenshtein distance between the item and its translation equivalent. The Levenshtein distance is the minimum number of insertions, deletions or substitutions necessary to change one word into another. We calculated Levenshtein distances using the Levenshtein.distance function from the vwr-package in R (Keuleers, 2013). In order to correct for word length, we normalized the Levenshtein distance as suggested by Schepens, Dijkstra and Grootjen (2012): normalized distance score = 1 – (distance/length), with distance representing the Levenshtein distance between the target item and its translation equivalent and

length representing the number of letters or sounds of the longest word (either the target or its Dutch translation equivalent). This results in a score between 0 and 1, with 0 representing no overlap between the word and 1 representing complete orthographic or phonological similarity.

Frequency measures were taken from databases based on subtitles: Subtlex-US (Brysbaert & New, 2009) for the English items, and Lexique (New et al., 2004) for the French items. The frequencies were expressed as Zipf-values (Van Heuven et al., 2014). The Zipf scale is a scale going from 1 (very low frequency words) to 6 (very high frequency content words) or 7 (a few function words, pronouns, and auxiliary verb forms like "have"). The calculation of Zipf values equals log10(frequency per million words) + 3.

5.3. Analysis.

Raw data and analysis files are available at https://osf.io/5y4x3/ and https://www.iris-database.org/. We calculated descriptive statistics for the English and French word recognition tests, the variables measuring individual differences, and the word-related variables.

We performed paired samples t-tests and calculated effect sizes to answer RQ1, which concerns the difference between the children's French and English vocabulary knowledge. As the exposure data were ordinal data, we used non-parametric tests. We performed Wilcoxon signed-rank tests to answer RQ2, which concerns the difference in children's out-of-school exposure to French and English, and we calculated Cliff's delta for repeated measures to estimate effect size. Cliff's delta is a measure of how often the values in one condition are larger than those in a second condition (with values between -1 and 1). We investigated possible differences in exposure between boys and girls with Wilcoxon rank-sum test and Cliff's delta effect size for between-groups designs.[3]

In order to determine which factors influence young learners' receptive L2 word learning (RQ 3), we analysed the data using a generalized linear mixed model (as the outcome

variable is binary, namely correct or wrong answer for the vocabulary items). In a linear mixed model it is possible to add both random and fixed effects which account for individual variation between items and participants. We constructed our model using the glmer-function from the R-package lme4 (Bates et al., 2015).

First, the continuous variables were centered on the mean. We then built a basic model with only random effects. We identified two random variables: items and participants.

We then added the fixed effects measuring individual differences: daily out-of-school exposure to the target language and gender (baseline = male). We also added the language-related fixed effects: language (baseline = English), phonological similarity, and frequency. To avoid a collinearity problem, we only added phonological similarity as phonological and orthographical similarity were highly correlated and learners were used to spoken input both through out-of-school exposure and in the classroom. We also added interactions between 'language' and the other fixed effects. Finally, we omitted interactions and variables until we identified the best fit. Model fit was assessed using anova(model 1, model 2), which gives a chi-square test of the relative fit of two embedded regression models (Brysbaert, 2020). We used the MuMin-package (Nakagawa & Schilezeth, 2013) to calculate marginal R^2, which measures the variance explained by the fixed effects only, and conditional R^2, which measures the variance explained by both the fixed effects and the random effects.

To answer RQ 4, we ran two separate multiple linear regression analyses, one for each language, with the total score of the vocabulary test as the dependent variable and the different types of input as the independent variables.

6. Results

6.1. Results for the English and French vocabulary tests

Because of the different numbers of items in PPVT and EVIP (120 vs. 104), the analyses are based on percentage correct, so that the numbers are comparable. Figure 1 shows that the distribution of percentage scores on both the English and the French receptive vocabulary tests were normally distributed. The minimum score for the English test was 33%, the mean was 64%, the median was 63% and the maximum was 97% (*SD* = 12.38). For the French test, the minimum score was 19%, the mean and the median were 41% and the maximum was 64% (*SD* = 9.32).

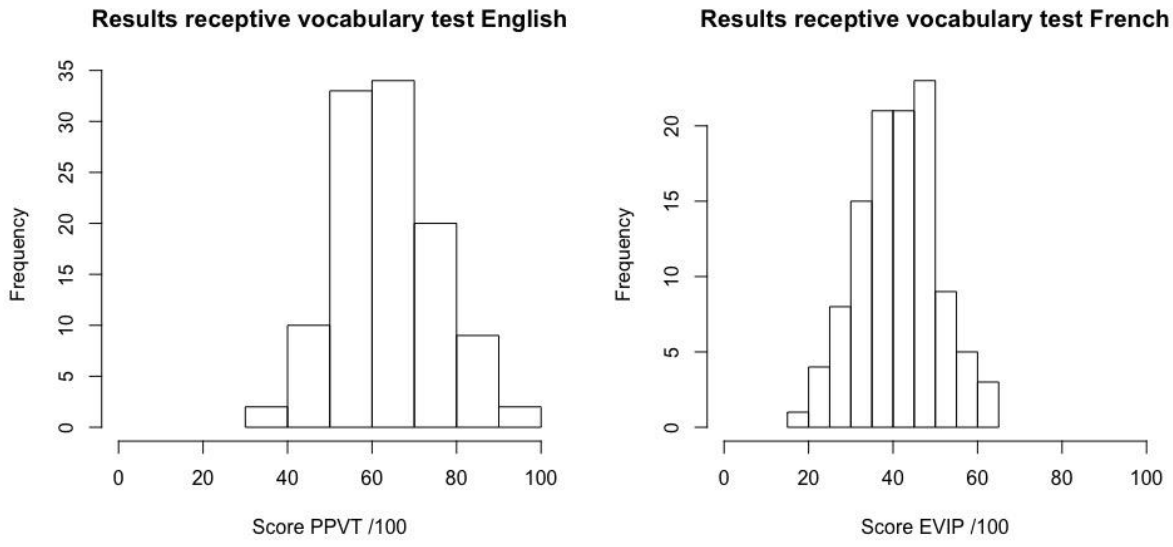


Figure 1. Distribution of the percentage scores on the PPVT-4 and the EVIP tests

A paired samples t-test revealed that children's overall score on the English vocabulary test was significantly higher than on the French vocabulary test with a large effect size ($t(109) = 22.21$, $p < 0.001$, $d = 2.09$). As previous research indicated that cognates are easier to learn than non-cognates, we additionally calculated the scores on the test when the cognates were left out.

We considered a word to be a cognate when the orthographic similarity score was .50 or higher (resulting in 54 Dutch-English cognates and 66 non-cognates in the PPVT and 23 Dutch-French cognates and 81 non-cognates in the EVIP). The minimum score on the English test without cognates was 17%, the mean was 51%, the median was 50% and the maximum was 97%. For French the minimum score was 7%, the mean and the median were 33% and the maximum was 59%. Again, a t-test revealed significantly higher scores for English than for French, although the effect size was substantially smaller ($t(109) = 13.17$, $p < 0.001$, $d = 1.31$).

6.2.Out-of-school exposure to English and French

Tables 1 and 2 show the different kinds of out-of-school exposure to English and French per day. The median for each type of input, results of Wilcoxon signed-rank tests comparing the amount of exposure in both languages and effect sizes can be found in table 3. For our analyses we assigned a score to the different categories: 0 = no exposure, 1 = 0-30 minutes, 2 = 30 minutes – 1 hour, 3 = 1 hour – 1 hour 30 minutes, 4 = 1 hour 30 minutes – 2 hours, 5 = more than 2 hours. The three types of watching television were all done significantly more in English than in French, but it can be observed that watching television without subtitles or with subtitles in the target language was rather uncommon in both languages for children this age. Three other types of exposure were also significantly more frequent in English than in French. These were listening to music, using social media and gaming. Only 69 children gave an estimate of gaming in French, the other children did not answer this question. From the responses we received, it is clear that gaming in French was highly infrequent among Flemish children. Therefore, we decided to also code the missing answers as 0. Two types of exposure were more common in French than in English: reading and speaking. A visual representation of the amount of exposure to both languages can be found in figure 2.

Table 1. *Percentage Frequency of Exposure to English per Day.*

	0 min	0-30 min	30 min-1h	1h-1h30min	1h30min-2h	>2hours
English spoken TV no subtitles	47	28	15	7	1	2
English spoken TV English subtitles	69	15	9	4	2	1
English spoken TV subtitles home language	19	16	30	20	8	6
Listening to English music	5	31	24	13	15	12
Reading in English	89	10	1	0	0	0
Playing English games	37	14	15	11	8	16
Using social media in English	27	31	19	10	3	10
Speaking English	56	37	4	0	2	1

Table 2. *Percentage Frequency of Exposure to French per Day.*

	0 min	0-30 min	30 min-1h	1h-1h30min	1h30min-2h	>2hours
French spoken TV no subtitles	85	7	5	2	0	0
French spoken TV French subtitles	93	7	0	0	0	0
French spoken TV subtitles home language	61	17	12	7	2	1
Listening to French music	49	42	8	1	0	0
Reading in French	69	10	21	0	0	0
Playing French games	94	3	2	1	0	0
Using social media in French	84	15	0	1	1	0
Speaking French	15	43	35	5	2	0

Table 3. *Results of Wilcoxon signed-rank tests and effect size for Different Types of Exposure in English and French.*

	English *Median*	French *Median*	*Cliff's Delta*	*Z-value Wilcoxon*	*p*
Watching tv without subtitles	1	0	.38	5.71	<.001 ***
Watching tv, subtitles in target language	0	0	.25	4.86	<.001 ***
Watching tv, subtitles in home language	2	0	.51	6.58	<.001 ***
Listening to music in target language	2	1	.71	8.26	<.001 ***
Reading in target language	0	0	-.22	-3.84	<.001 ***
Gaming in target language	1	0	.60	7.94	<.001 ***
Using social media in target language	1	0	.62	8.07	<.001 ***
Speaking in target language	0	1	-.53	-6.68	<.001 ***

* $p < .05$, ** $p < .01$, *** $p < .001$ (after Bonferroni correction for 8 comparisons)

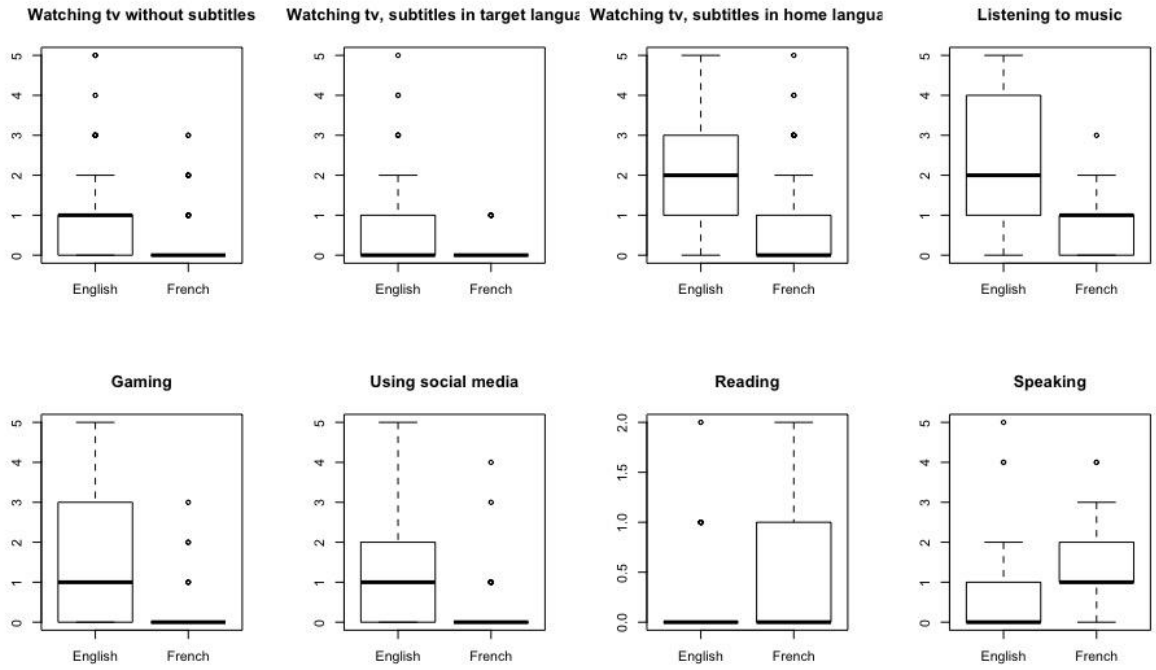


Figure 2. Boxplots showing the different types of out-of-school exposure in English and French.

The children were also asked two questions about speaking English or French. In the first question they were asked with whom they spoke the language. The results reflected the fact that French is a school language and English is only encountered outside the school. For English, 44 children reported speaking English with family or friends, eight children spoke English with strangers, two children with someone from the sports club and only one child reported speaking English with someone from school. For French, on the other hand, most children reported speaking French with people from school (*n*=70), 39 children spoke French with family or friends and four children spoke French to strangers. When asked about the situation in which they spoke the target language, there were also clear differences between the two languages. 17 children reported speaking English for fun, whereas only five children reported speaking French for fun. Nine children spoke English on holidays, only three spoke French. Seven children reported speaking English because it was necessary for communication purposes, whereas only one child thought French was necessary for communication purposes. The main reason for speaking French was practice for school (*n*=29) whereas only nine children reported speaking English to practice.

6.3. Gender

Our sample consisted of 53 boys and 57 girls. Wilcoxon rank-sum tests for exposure to English revealed that boys were more frequent gamers than girls. Wilcoxon rank-sum tests for exposure to French showed no significant differences between boys and girls. An overview of the results of the analyses can be found in table 4.

Table 4. *Results of Wilcoxon rank-sum tests and effect size for Different Types of Exposure and Gender.*

	Boys *Median*	Girls *Median*	*Cliff's delta*	*Z-value Wilcoxon*	*p*
Watching tv without subtitles EN	0	1	-.06	-0.56	n.s.
Watching tv, subtitles in target language EN	0	0	-.15	-1.70	n.s.
Watching tv, subtitles in home language EN	2	2	.12	1.14	n.s.
Listening to music in target language EN	2	2	-.08	-0.73	n.s.
Reading in target language EN	0	0	-.10	-1.64	n.s.
Gaming in target language EN	3	0	.74	6.84	<.001***
Using social media in target language EN	1	1	.13	1.19	n.s.
Speaking in target language EN	0	0	.14	1.45	n.s.
Watching tv without subtitles FR	0	0	-.06	-0.91	n.s.
Watching tv, subtitles in target language FR	0	0	.04	0.84	n.s.
Watching tv, subtitles in home language FR	0	0	-.14	-1.44	n.s.
Listening to music in target language FR	1	1	.01	0.06	n.s.
Reading in target language FR	0	0	.02	0.19	n.s.
Gaming in target language FR	0	0	.11	2.60	.168
Using social media in target language FR	0	0	.13	1.76	n.s.
Speaking in target language FR	1	1	-.13	-1.25	n.s.

* p < .05, ** p < .01, *** p < .001(after Bonferroni correction for 18 comparisons)

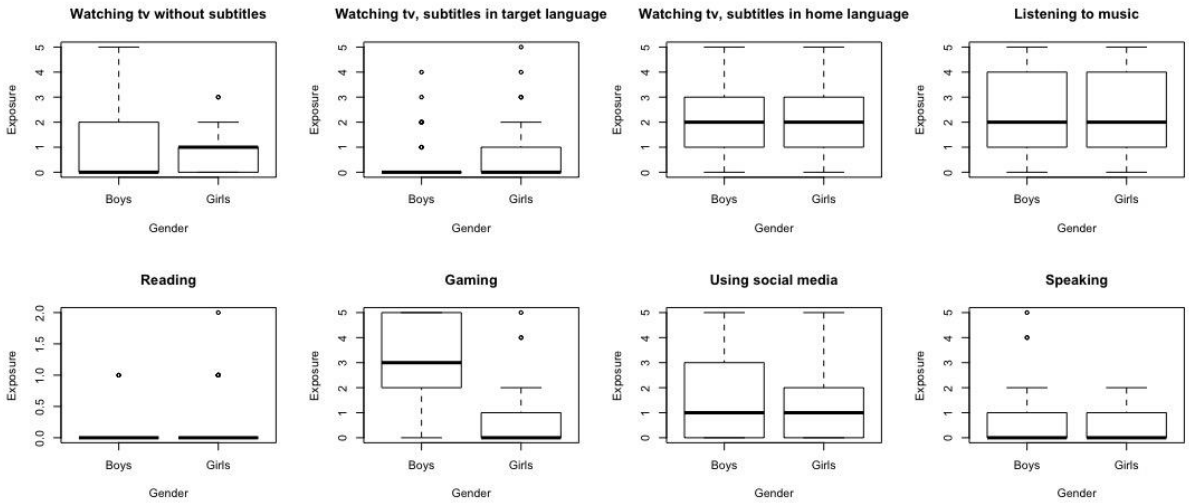


Figure 3. Boxplots showing the different types of out-of-school exposure for English split by gender.

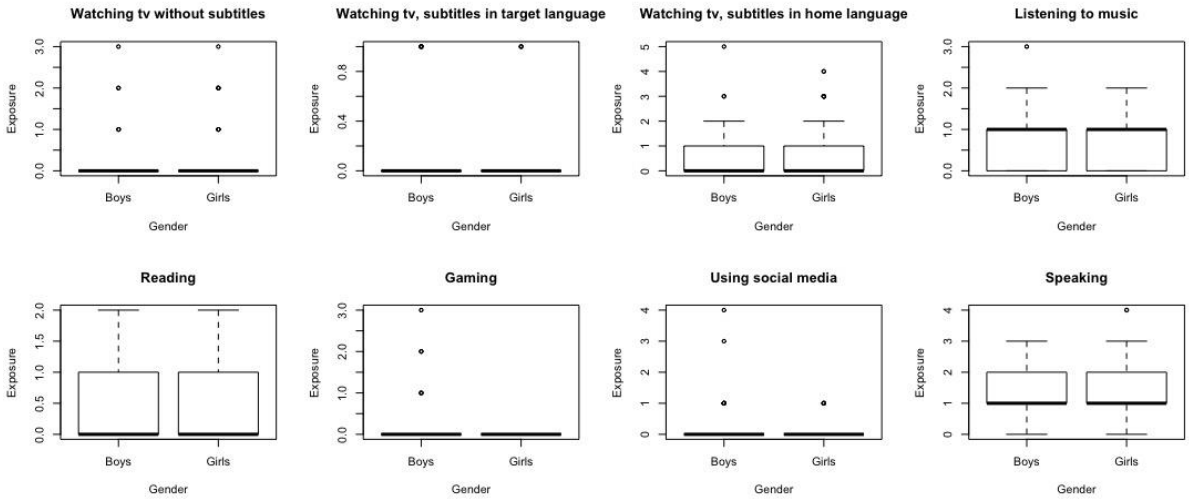


Figure 4. Boxplots showing the different types of out-of-school exposure for French split by gender.

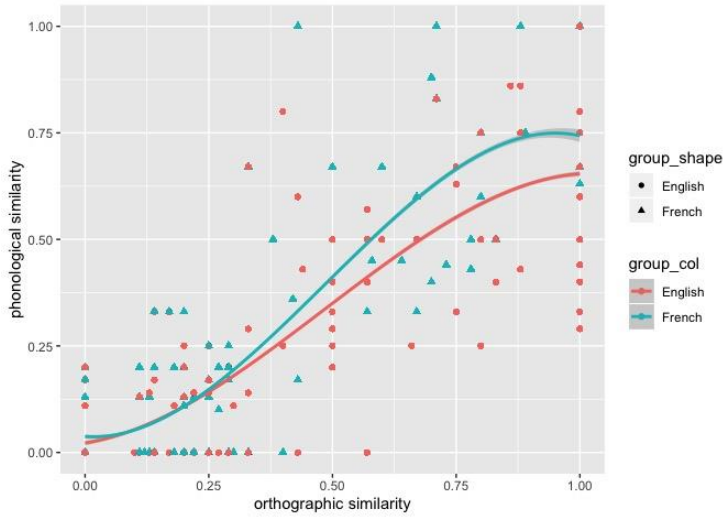
6.4.Word-related variables

The mean orthographic similarity for the English items was 0.43 (minimum similarity = 0.00, maximum similarity = 1.00, *SD* = 0.35), for the French items it was 0.28 (minimum similarity = 0.00, maximum similarity = 1.00, *SD* = 0.28). A t-test revealed that the orthographic similarity between the English items and their Dutch translation equivalents was significantly

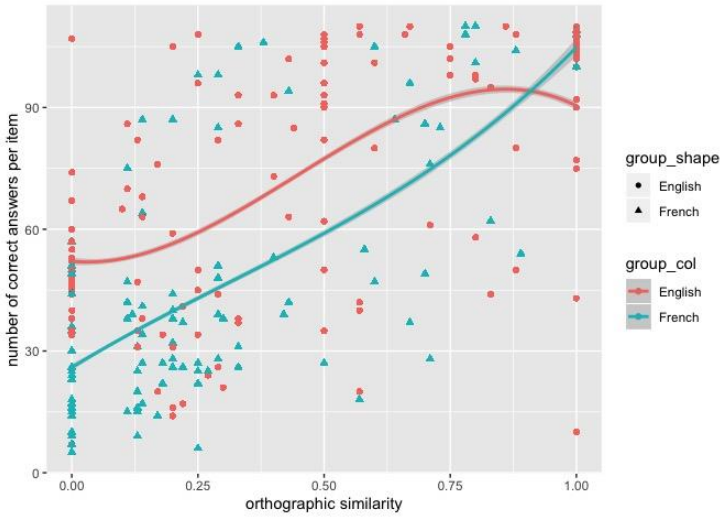
greater than the orthographic similarity between the French items and their Dutch translation equivalents. The effect size was medium ($t(221) = 3.65$, $p < 0.001$, $d = 0.48$). The mean phonological similarity for the English items was 0.30 (minimum similarity = 0.00, maximum similarity = 1.00, *SD* = 0.29), and for the French items it was 0.22 (minimum similarity = 0.00, maximum similarity = 1.00, *SD* = 0.28). A t-test revealed that the phonological similarity between the English items and the Dutch translation equivalent was significantly greater than the phonological similarity between the French items and their Dutch translation equivalent. The effect size was small, however ($t(219) = 2.01$, $p = 0.04$, $d = 0.27$).

The measures for phonological and orthographic similarity were strongly correlated ($r = .83$ for the English items and $r = .82$ for the French items) and so were the measures for orthographic similarity and the test results ($r = .52$ for English and $r = .66$ for French) and phonological similarity and the test results ($r = .56$ for English and $r = .66$ for French). The results are visualized in figures 5a-c.

a.



b.



c.

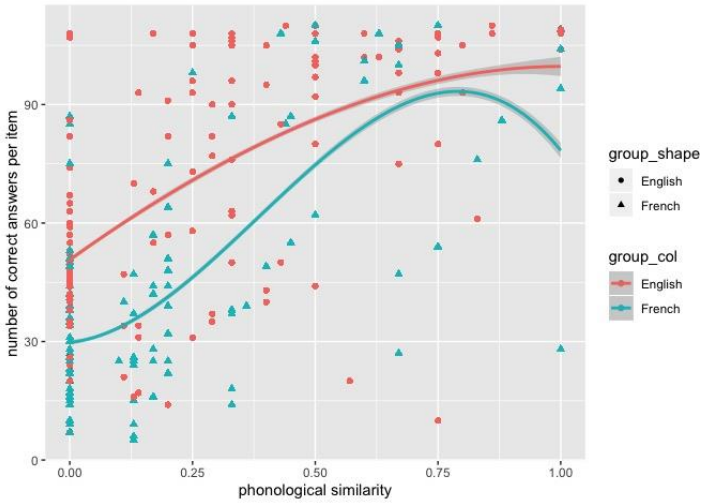


Figure 5a-c. Scatterplots showing the correlations between different measures of cognateness and between cognateness and the score per item split by language.

The mean item frequency for the English items was 3.98 (minimum Zipf score 2.30, maximum Zipf score 5.52, *SD* = 0.64) and the mean item frequency for the French items was 3.72 (minimum Zipf score 1.70, maximum Zipf score 6.01, *SD* = 0.80). The correlation between frequency and the test results was .20 for English and .11 for French.

6.5.How do child variables and word variables predict receptive vocabulary knowledge?

The basic generalized linear mixed effects model included only the random effects of participant and item: glmer(accuracy ~ 1|participant + 1|item). It showed that most of the variation was explained by the variable 'item' (variance = 3.95, *SD* = 1.99, ICC= .52); the variable 'participant' explained far less variation (variance = 0.39, *SD* = 0.62, ICC = .05).

The best model to answer RQ 3 contained random intercepts for item and participant and four significant fixed effects (exposure, language, phonological similarity and frequency), as shown by model comparisons using the procedure anova(model 1, model 2) from the R package lme4 (Bates et al., 2015). The results of the analysis can be found in table 5. There were significant positive main effects of exposure, phonological similarity and word frequency, and a significant negative main effect of language (French being more difficult than English). Gender was not a significant predictor of receptive word knowledge and did not improve the model. The model has a marginal R^2 of .26 and a conditional R^2 of .57. The fixed effects in the model thus explain 26 percent of the variance.

Table 5. *GLME Model Predicting Right or Wrong Answers on the Vocabulary Tests with total exposure.*

Random effects		Variance	*SD*
Item	(Intercept)	2.047	1.431
Participant	(Intercept)	0.356	0.596

Fixed effects	Estimate	*SE*	Z-value	p
(Intercept)	0.72	0.15	4.87	<.001 ***
Exposure	0.07	0.01	11.59	<.001 ***
Language	-0.60	0.20	-2.93	.003 **
Phonological similarity	4.01	0.35	11.40	<.001***
Frequency	0.41	0.14	3.01	.003 **

*** p<.001, **p<.01, *p< .05

To answer RQ4, we ran a linear regression for each language with the total score on the meaning recognition test as dependent variable and the eight types of input as independent variables. The results of the analyses can be found in tables 6a and 6b. Four variables were significant in the model for English: reading, listening to music, using social media and speaking. The model explained 29% of the variance. Only reading was significant in the model for French. The model for French explained less than 1% of the variance.

Table 6a. *Results of the regression model for English*

Predictors	B	SE	β	p
(Constant)	65.69	2.74		<.001 ***
Watching TV no subtitles	-1.08	1.19	-.08	.367
Watching TV subtitles target language	-0.93	1.20	-.06	.440
Watching TV subtitles home language	-0.02	0.94	.00	.984
Reading	13.92	3.62	.33	<.001 ***
Listening to music	2.25	0.95	.23	.019 *
Gaming	0.51	0.80	.06	.525
Using social media	2.07	0.97	.22	.035 *
Speaking	3.08	1.54	.18	.048 *
Model summary	Adjusted R-squared: .29, df 99			

*** p<.001, **p<.01, *p< .05

Table 6b. *Results of the regression model for French*

Predictors	B	SE	β	p
(Constant)	40.06	2.08		<.001 ***
Watching TV no subtitles	1.35	1.65	.09	.414
Watching TV subtitles target language	0.33	4.02	.01	.935
Watching TV subtitles home language	0.04	0.92	.00	.965
Reading	2.52	1.16	.21	.032 *
Listening to music	-0.27	1.45	-.01	.854
Gaming	-0.52	2.45	-.02	.832
Using social media	0.19	1.73	.01	.913
Speaking	1.10	1.10	.10	.320
Model summary	Adjusted R-squared: .001, df 101			

*** p<.001, **p<.01, *p< .05

7. Discussion

In the present paper we profited from the Belgian context to compare exposure-based and school-based language learning in the very first stages of L2 acquisition. Dutch-speaking primary school children must learn French at school (2 hours per week), but are more likely to need English for sought-after activities outside school. Our results show that even without any school-based support for English, vocabulary acquisition through contextual learning was larger for this language than for French. This is in line with previous research (Peters et al., 2019), but the present study shows that the advantage is already present at a stage when there is no formal support for learning English. This confirms what was found in previous studies concerning language learning through out-of-school exposure, namely that large language gains can be achieved without formal instruction (De Wilde et al., 2020a; Lefever, 2010; Puimège & Peters, 2019). A closer look at the results also reveals that the difference in the overall score for the English and French meaning recognition test cannot solely be attributed to the greater number of cognates in the English test. When analysing the test results without cognates, the scores for English were still significantly higher than for French.

The language difference was confirmed by the results from the questionnaire concerning out-of-school exposure to both languages. The average exposure to English was far higher than to French. Several activities were done in English daily by a majority of the children. These were: watching television, listening to music, gaming and using social media. Only two activities were more frequent in French: reading and speaking. These can be related to the fact that French is a school subject: children reported that the most frequent activity, speaking French, was often done to practise, whereas speaking English, even though less frequent, was done because it was necessary for communication purposes or because it was considered fun. Since there are a lot of opportunities to encounter English outside the classroom, the results for

English exposure showed a larger range, whereas out-of-school exposure to French is less frequent and translates itself into smaller individual differences between the children.

We additionally wanted to investigate the influence of a number of individual differences and word-related variables on receptive word learning. The children's gender did not predict whether they would know a word. This seems to suggest that gender in itself is not a predictor of linguistic performance but that the effects of gender which are sometimes found in studies looking at language acquisition should be attributed to confounded factors (such as exposure) (see also Courtney et al., 2015; Peters et al., 2019). [4]

Four variables were significant in the generalized linear mixed model and explained 26% of the variance: amount of exposure (participant variable), phonological similarity to the L1, frequency and language (stimulus variables). The significant effect of exposure indicates that, regardless of the language, the degree of exposure is one of the variables that helps build a learners' receptive vocabulary. This is in line with findings from previous research which already mentioned the importance of extra exposure to supplement classroom exposure (Bybee & Hopper, 2001; Ellis, 2002). The present study illustrates this need. It shows that contextual exposure leads to more vocabulary learning than explicit teaching and testing of words required for effective communication (the goal of L2 teaching in class).

A closer look at the influence of different types of out-of-school exposure shows that for English four types of input contribute to word learning: using social media, speaking, listening to music and reading. We hypothesized that the effect of reading, using social media and speaking would positively influence contextual L2 learning and this proved to be the case. We believe that these types of input contribute to L2 learning because there is no L1 support and learners need to figure out the meaning from context. This might seem to contradict findings in classroom contexts which have shown that using the L1 can benefit learning (e.g. Yanagisawa et al., 2020) but this need not be the case. In a classroom setting the main goal is learning,

whereas contextual learning happens when the learner is doing another activity. Some activities can be done without paying attention to the L2 (e.g. watching television with subtitles in the home language) but other activities require engagement with the L2 (e.g. using social media). The latter seem to lead to more learning in an out-of-school context.

In contrast with previous research, we did not find a significant effect of watching TV in the target language or of gaming on English L2 word acquisition, while we observed a positive effect of listening to English music. We hypothesize that the variable findings with these types of input are due to the fact that they do not 'force' learners to engage with the L2 input when encountered outside the classroom (learners can rely on L1 subtitles, ignore the lyrics in songs and avoid games with much language). The fact that learners can engage with these types of input in various ways may explain the mixed findings, although more research is needed to confirm this.

The model looking into word learning through out-of-school exposure in French hardly explained any variance (R^2 = .001). This is because there is little exposure to French outside the classroom and not much variability between the learners. Only one variable was significant: reading.

We also investigated three word-related variables. Cognateness proved to be important in L2 meaning recognition. Words in the target language that are alike in form and meaning to the Dutch translation equivalent are easier to learn. This is in line with findings concerning receptive multilingualism (Gooskens et al., 2018) and other studies with young L2 English learners (De Wilde et al., 2020b; Muñoz et al., 2018; Puimège and Peters, 2019). This is also in line with the parasitic hypothesis of word learning, which states that learners use their existing vocabularies to derive the meaning of new words, especially in the initial stages of word learning (Hall, 2002). For the participant group we tested, this gave an advantage to English, as there are more English cognates in Dutch than French cognates, both in spoken and

written mode. The study further confirms that more frequent words are known better than less frequent words, both for English and French. The final variable that was significant in the model was language. English words were known more often than French words. This variable showed that, even when exposure, cognateness and frequency were accounted for, English words were still better known than French words. There could be several reasons for this. As children were exposed more often to English than to French, they might be more familiar with the sounds of the language and they might find it easier to recognize words in English than in French. Further, the presence of many cognates in the input might also serve as contextual clues when learning new, non-cognate words. There might still be other reasons why English is easier than French. A look into motivation for example might give us more insight in the difference. Further research into this is warranted.

The present study also has some limitations. The most important is that our findings are limited to spoken vocabulary understanding. Other big goals of classroom teaching are L2 reading, L2 speaking and writing, and correct use of grammar. Readers may object that very different results will be found if one of these aspects is tested. This is a challenge for future research, although our prediction is that the pattern observed in the present article will generalize well to other aspects of language use, based on the correlation of .7 we found between spoken word understanding and other aspects of L2 proficiency in a previous study (De Wilde et al., 2020a). Secondly, our findings are limited to Dutch L1, which as a Germanic language is closer to English than to French. It would be interesting to repeat our study with a Romance L1. Would authors then still observe that the presence of English as a lingua franca outweighs two years of French L2 teaching? Finally, our study had considerably more power to find within-person differences (on stimulus-related variables) than between-group differences (on participant-related variables). Given the difficulty to test very large groups, we think the main avenue for research on between-group variables is for each study to report all

the participant-related effects independent of statistical significance and to aggregate the findings in a meta-analysis.

8. Conclusion

The results of this study show that even before the start of formal instruction in English, Flemish children's English receptive vocabulary knowledge is larger than their French vocabulary knowledge. It seems that English has a number of advantages over French for Dutch-speaking learners. The language is more closely related to Dutch, and, as it is the lingua franca, it is more prevalent in different contexts (such as watching television and using social media). The combination of these factors results in more language learning.

The learners' starting point for vocabulary learning is the L1, even in a language which shares fewer cognates such as French. The study further clearly shows the importance of contextual language learning in L2 word learning. If we aim for learners to have large L2 vocabularies, instruction time should be supplemented with out-of-school contact with the L2. The amount of exposure to the target language and the similarity between the foreign language and the L1 turn out to be the two main factors for young learners' successful language learning.

Notes

1 The Belgian context illustrates that in many countries multilingualism involves the acquisition of more than one L2 in parallel. We keep the term *L2 acquisition* for the sake of continuity with the literature.

2 A sample size of 110 allows researchers to find an effect size of d = .27 with 80% power in a pairwise comparison of two within-subjects conditions. It has 90% power to detect an effect

size of d = .31. It also has 80% power to find an interaction between two repeated measures with an effect size of d = .4 in one condition and none in the other (Brysbaert, 2019).

3 Researchers who need d-values for meta-analysis, can estimate them on the basis of the Wilcoxon test by means of the equations $d = z/\sqrt{N}$ for repeated measures and $d = 2 * z/\sqrt{N_1 + N_2}$ for between groups. So, the z-score of 5.71 for watching tv in English vs. French without subtitles in Table 3 translates to $d = 5.71/\sqrt{110} = .54$ and the z-score of 6.84 for gaming in English for boys versus girls in Table 4 translates to $d = 2 * 6.84/\sqrt{53 + 57} = 1.30$ (for further information, see Brysbaert, 2020).

4 At the same time, it should be taken into account that group sizes of N = 55 are rather low for between-group comparisons. To detect an effect size of d = .4 with 80% power, one needs two groups of 100 participants (Brysbaert, 2019). So, studies with larger numbers of participants are needed if the gender differences are the main variable of interest. Alternatively, a meta-analysis including the present findings together with those of other studies could also give a more definite answer.

References

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1).

Bisson, M.-J., van Heuven, W. J. B., Conklin, K., & Tunney, R. J. (2014). The Role of Repeated Exposure to Multimodal Input in Incidental Acquisition of Foreign Language Vocabulary: Repeated Exposure in Incidental FL Acquisition. *Language Learning*, *64*(4), 855–877.

Brysbaert, M. (2019). How many participants do we have to include in properly powered experiments? A tutorial of power analysis with reference tables. *Journal of Cognition, 2(1)*, 16.

Brysbaert, M. (2020). *Basic statistics for psychologists (2nd edition)*. Macmilan International.

Brysbaert, M., Mandera, P., & Keuleers, E. (2018). The Word Frequency Effect in Word Processing: An Updated Review. *Current Directions in Psychological Science*, *27*(1), 45–50.

Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, *41*(4), 977–990.

Bybee, J. L., & Hopper, P. J. (Eds.). (2001). *Frequency and the emergence of linguistic structure* (Vol. 45). John Benjamins.

Courtney, L., Graham, S., Tonkyn, A., & Marinis, T. (2017). Individual differences in early language learning: A study of English learners of French. *Applied Linguistics, 38*(6), 824-847.

d'Ydewalle, G., & Van de Poel, M. (1999). Incidental Foreign-Language Acquisition by Children Watching Subtitled Television Programs. *Journal of Psycholinguistic Research, 28*(3), 227–244.

De Wilde, V., Brysbaert, M., & Eyckmans, J. (2020a). Learning English through out-of-school exposure. Which levels of language proficiency are attained and which types of input are important? *Bilingualism: Language and Cognition,23*(1), 171–185.

De Wilde, V., Brysbaert, M., & Eyckmans, J. (2020b). Learning English through out-of-school exposure: how do word-related variables and proficiency influence receptive vocabulary learning? *Language Learning, 70*(2), 349-381.

Doughty, C., & Williams, J. (1998). *Focus on Form in Classroom Second Language Acquisition. The Cambridge Applied Linguistics Series*. Cambridge University Press

Dunn, L., & Dunn, L. (2007). *Peabody Picture Vocabulary Test (4th ed)*. American Guidance Service.

Dunn, L., Dunn, L, & Theriault-Whalen, C. (1993). *EVIP: échelle de vocabulaire en images Peabody.* Pearson Canada Assessment.

Elgort, I., Brysbaert, M., Stevens, M., & Van Assche, E. (2018). Contextual word learning during reading in a second language: an eye-movement study. *Studies in Second Language Acquisition*, *40*(2), 341–366.

Elley, W. B., & Mangubhai, F. (1983). The impact of reading on second language learning. *Reading research quarterly*, 53-67.

Ellis, N. C. (1998). Emergentism, connectionism and language learning. *Language learning*, *48*(4), 631-664.

Ellis, N. C. (2002). Frequency effects in language processing. *Studies in Second Language Acquisition*, *24*(02), 143-188.

Enever, J. (Ed.). (2011). *ELLiE - Early language learning in Europe*. British Council.

Firth, A., & Wagner, J. (1997). On discourse, communication, and (some) fundamental concepts in SLA research. *The Modern Language Journal*, *81*(3), 285-300.

Ford-Connors, E., & Paratore, J. R. (2015). Vocabulary Instruction in Fifth Grade and Beyond: Sources of Word Learning and Productive Contexts for Development. *Review of Educational Research*, *85*(1), 50–91.

González Fernández, B., & Schmitt, N. (2015). How much collocation knowledge do L2 learners have?: The effects of frequency and amount of exposure. *ITL - International Journal of Applied Linguistics*, *166*(1), 94–126.

Gooskens, C., van Heuven, V. J., Golubović, J., Schüppert, A., Swarte, F., & Voigt, S. (2018). Mutual intelligibility between closely related languages in Europe. *International Journal of*

Multilingualism, *15*(2), 169–193.

Goriot, C., van Hout, R., Broersma, M., Lobo, V., McQueen, J. M., & Unsworth, S. (2018). Using the peabody picture vocabulary test in L2 children and adolescents: Effects of L1. *International Journal of Bilingual Education and Bilingualism*, 1–23.

Hall, C. J. (2002). The automatic cognate form assumption: Evidence for the parasitic model of vocabulary development. *IRAL - International Review of Applied Linguistics in Language Teaching*, *40*(2), 69 – 88.

Hannibal Jensen, S. (2017). Gaming as an English Language Learning Resource among Young Children in Denmark. *CALICO Journal*, *34*(1), 1–19.

Haycock, K. (2001). Closing the achievement gap: Help-ing all students achieve. *Educational Leadership*, *58*,6–11.

Horst, M., Cobb, T., & Meara, P. (1998). Beyond A Clockwork Orange: Acquiring Second Language Vocabulary through Reading. *Reading in a Foreign Language*, *2*(11), 207–223.

Housen, A., Janssens, S., & Pierrard, M. (2001). *Frans en Engels als vreemde talen in Vlaamse scholen*. VUBPress.

Hulstijn, J. H. (2003). Incidental and Intentional Learning. In C. J. Doughty & M.H.Long(Eds.), *The Handbook of Second Language Acquisition* (pp. 349–381). Blackwell Publishing Ltd.

Hulstijn, J. H. (2011). Language proficiency in native and nonnative speakers: An agenda for research and suggestions for second-language assessment. *Language Assessment Quarterly, 8*(3), 229-249.

Jaekel, N., Schurig, M., Florian, M., & Ritter, M. (2017). From early starters to late finishers? A longitudinal study of early foreign language learning in school. *Language Learning, 67*(3), 631-664.

Keuleers, E. (2013), vwr: Useful functions for visual word recognition research. R-package version 0.3.0.

Koolstra, K. M., & Beentjes, J. W. J. (1999). Children's Vocabulary Acquisition in a Foreign Language Through Watching Subtitles Television Programs at Home. *Eductaion Technology Research and Development*, *47*(1), 51–60.

Kuppens, A. H. (2010). Incidental foreign language acquisition from media exposure. *Learning, Media and Technology*, *35*(1), 65–85.

Lefever, S. (2010). *English skills of young learners in Iceland:"I started talking English when I was 4 years old. It just bang… just fall into me.".* Presented at the Menntakvika Conference, Reijkjavik, Iceland.

Lindgren, E., & Muñoz, C. (2013). The influence of exposure, parents, and linguistic distance on young European learners' foreign language comprehension. *International Journal of Multilingualism*, *10*(1), 105–129.

Muñoz, C., Cadierno, T., & Casas, I. (2018). Different Starting Points for English Language Learning: A Comparative Study of Danish and Spanish Young Learners: Different Starting Points. *Language Learning*, *68*(4), 1076–1109.

Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R^2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, *4*(2), 133–142.

New, B., Pallier, C., Brysbaert, M., & Ferrand, L. (2004). Lexique 2: A new French lexical database. *Behavior Research Methods, Instruments, & Computers*, *36*(3), 516-524.

Palmberg, Rolf. (1985). How Much English Vocabulary Do Swedish-Speaking Primary-School Pupils Know before Starting to Learn English at School? In Ringbom, Hakan (Ed.), *Foreign Language Learning and Bilingualism* (Vol. 105, pp. 89–97). Publications of the Research Institute of the Åbo Akademi Foundation,.

Pavia, N., Webb, S., & Faez, F. (2019). Incidental vocabulary learning through listening to songs. *Studies in Second Language Acquisition*, *41*(04), 745–768.

Pellicer-Sanchez, A., & Schmitt, N. (2010). Incidental vocabulary acquisition from an authentic novel: Do Things Fall Apart. *Reading in a Foreign Language*, *1*(22), 31–55.

Peters, E. (2018). The effect of out-of-class exposure to English language media on learners' vocabulary knowledge. *ITL - International Journal of Applied Linguistics*, *169*(1), 142–168.

Peters, E., Noreillie, A., Heylen, K., Bulté, B., & Desmet, P. (2019). The Impact of Instruction and Out-of-School Exposure to Foreign Language Input on Learners' Vocabulary Knowledge in Two Languages. *Language Learning*, *69*(3), 747 – 782.

Peters, E., & Webb, S. (2018). Incidental vocabulary acquisition through viewing L2 television and factors that affect learning. *Studies in Second Language Acquisition*, 1–27.

Puimège, E., & Peters, E. (2019). Learners' English Vocabulary Knowledge Prior to Formal Instruction: The Role of Learner-Related and Word-Related Variables: Vocabulary Knowledge Prior to Formal Instruction. *Language Learning, 69*(4), 943 - 977.

Schepens, J., Dijkstra, T., & Grootjen, F. (2012). Distributions of cognates in Europe as based on Levenshtein distance. *Bilingualism: Language and Cognition*, *15*(01), 157–166.

Sylvén, L. K., & Sundqvist, P. (2012). Gaming as extramural English L2 learning and L2 proficiency among young learners. *ReCALL*, *24*(03), 302–321.

Uchihara, T., Webb, S., & Yanagisawa, A. (2019). The Effects of Repetition on Incidental Vocabulary Learning: A Meta-Analysis of Correlational Studies. *Language Learning*, *69*(3), 559–599.

Van der Slik, F. W. P. (2010). Acquisition of Dutch as a Second Language. *Studies in Second Language Acquisition*, *32*(03), 401–432.

van Heuven, W. J. B., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). Subtlex-UK: A New and Improved Word Frequency Database for British English. *Quarterly Journal of Experimental Psychology*, *67*(6), 1176–1190.

Vanhove, J., & Berthele, R. (2015). The lifespan development of cognate guessing skills in an unknown related language. *International Review of Applied Linguistics in Language Teaching*, *53*(1), 1–38.

Vidal, K. (2011). A Comparison of the Effects of Reading and Listening on Incidental Vocabulary Acquisition: Incidental Vocabulary Acquisition. *Language Learning*, *61*(1), 219–258.

Vlaamse Overheid (2018). *Peiling Frans in het basisonderwijs* [Survey into French in primary education]. Retrieved from http://eindtermen.vlaanderen.be/peilingen/basisonderwijs/peilingen/files/frans/Peiling_Frans2018_digitaal.pdf, June 2, 2020

Webb, S. (2007). The Effects of Repetition on Vocabulary Knowledge. *Applied Linguistics*, *28*(1), 46–65.

Webb, S. (2015). Extensive Viewing. *Language Learning Trough Watching Televion.–In: Nunan, David & Richards, Jack C.(Hrsg.). Language Learning Beyond the Classroom.*(pp. 159-168) New York, NY: Routledge.

Yanagisawa, A., Webb, S., & Uchihara, T. (2020). How do different forms of glossing contribute to L2 vocabulary leanring from reading?: A meta-regression analysis. *Studies in Second Language Acquisition*, *42*(2), 411–438.