Do Dutch/English cognates and word frequency affect English vocabulary testing in secondary bilingual education?

Evidence from the Peabody Picture Vocabulary Test

Alicia Abad Perales
S4743431
July, 2017
Master in Linguistics
Radboud Universiteit, Nijmegen

First supervisor: Mirjam Broersma & Roeland van Hout
Second supervisor: Claire Goriot

Table of contents

I. Acknowledgments

I would like to express my gratitude to the school where the data for this thesis project was gathered from for its assistance and support in recruiting the pupils. Moreover, many thanks to professor Vanessa Lobo for generously arranging everything and patiently answer all my questions. I would also like to thank the pupils who participated in this study.

I would also like to thank Mirjam Broersma for providing me with insightful comments of my first drafts. Finally, I gratefully acknowledge the help and advice of Claire Goriot and Roeland van Hout on the statistical analysis.

1. Introduction

This thesis aims to examine to what extent Dutch first language (L1) vocabulary interferes with English second language (L2) vocabulary testing. This topic pertains to the broader field of bilingual education, which has become an increasingly significant phenomenon in primary and secondary schools in Europe. The expansion of this type of education has grown along with the concern to gain a detailed understanding of its effects in comparison with traditional monolingual education. For this reason, measuring the impact of bilingual education has become an interesting issue to tackle by researchers studying bilingualism.

One of the most prominent fields of bilingual education is the boost it provides to the acquisition of L2 vocabulary. This study pursues that line and, as many others, focuses its efforts in the effect cognates have over L2 vocabulary acquisition. It has been studied that cognates, lexical items which are similar in terms of their phonology and orthography across two languages, play an important role when learning words in the second language (Schepens et al., 2013). In order for researchers to address this issue, they have created specific stimuli containing cognates to investigate the performance of bilingual participants. However, this study goes further and works with an English standardized test, the Peabody Picture Vocabulary Test (PPVT). By having a standardized L1 test, it is ensured that (1) the stimuli were not created with the purpose of testing cognates but instead were part of a more extended L1 (English) vocabulary testing purpose, and (2) provides a representative sample of the overall vocabulary in which cognates might appear or not, which adds ecological validity to the study.

A detailed examination of the PPVT observes that "it has an appealing nature, it requires no written or oral response, no reading is needed, there is minimal risk for stress or perceived failure, and it is applicable to a wide age range" (Dunn & Dunn, 2007). As a result, many researchers have administered the test not only to L1 speakers, but also to L2 learners in order to measure their L2 vocabulary. However, this instrument was first designed to test L1 vocabulary and, as a consequence, it raises the question whether the PPVT is reliable for weighing non-native vocabulary as well. This study aims to give an answer to this question by investigating the effect L1 Dutch cognates have on the outcomes of the vocabulary test.

The present study is organized in the following way:

Section 1.1 first describes the role of bilingualism in society and introduces the importance of the topic of bilingual education at the European level. Its first subsection, 1.1.1, focuses on the Content and Language Integrated Learning program (CLIL), an initiative that seeks to integrate the second or foreign language in the content of the classroom. Subsection 1.1.2 provides an up-to-date account of studies addressing the advantages bilingual education yields over monolingual education. Then, in 1.1.3, the literature that focuses on the cognitive benefits that arise from bilingualism is discussed. Section 1.2 describes Dutch secondary bilingual education, explaining its structure and how the bilingual curriculum is embedded. Section 1.3 addresses the effects that the mother tongue has on the acquisition of vocabulary in a second language. Recent studies focused on the cognate advantage effects as well as the L1 word frequency effect affecting the acquisition of L2 vocabulary will be discussed. Finally, section 1.4 concludes with an overview of the present study and the research questions under investigation.

Chapter 2 discusses the methodology followed in this study. Section 2.1 explains the profile of the participants engaged in this study and the structure and organization of their secondary school regarding its curricula in bilingual teaching. Section 2.2 introduces the materials and the procedure used in this project and it consists out of several subsections. Subsection 2.2.1 and 2.2.2 explain the system used for choosing the translations and their phonetic transcriptions, respectively. The next subsection, 2.2.3, explains how the cognate distance was operationalized. To end the methodology chapter, section 2.3 describes the analysis performed on the data.

Chapters 3 displays the results and chapter 4 presents a discussion of the results. The discussion section will provide an interpretation of the results.

Chapter 5, the conclusion, summarizes the findings and closes the thesis, providing some ideas for future research. Finally, an appendix will be provided.

1.1 Bilingual education

In today's globalized world, the ability to accomplish communication in more than one language is increasingly important. Due to this, mastering a second language has become a central issue for many governments. Bilingualism is a great asset to achieve as it entails many advantages: it may increase citizens' employability, facilitate access to services and rights, and contributes to solidarity through enhanced intercultural dialogue and social cohesion

(European Commision, 2008). Moreover, bilingualism is not only a skill commended by multinational associations such as the European Union but it is also something that citizens themselves recognize as a valuable expertise to accomplish. According to a survey report published by Eurobarometer (2012), there is a broad consensus among Europeans that everyone in the EU should be able to speak at least one foreign language. Bilingual education plays a critical role in the maintenance and development of bilingualism in the European community. Furthermore, bilingualism has been said to enhance cultural awareness, both of the culture of origin and the culture of the additional language (García, 2009). Baetens Beardsmore (1994) identified the cultural competence as containing four major parts: (1) knowledge of language, history, traditions and institutions of the target culture; (2) feelings, referring to attitudes towards a culture; (3) behavior, related to the capacity to act in a culturally appropriate way; and (4) metacultural awareness, related to the understanding of the distinctiveness of cultures. As Toyoda (2017) points out, learning a foreign language exposes students to alternative world views and enables them to absorb target cultural information directly from original sources. However, many times it is not bilingualism itself but the teacher, note van Rossum & Vismans (2006), who plays the main role in the acquisition of cultural knowledge and intercultural skills.

1.1.1 Content and Language Integrated Learning (CLIL)

The priorities assigned to foreign language learning have led many countries to adopt initiatives pursuing bilingual education (Eurobarometer, 2012). Enever (2011) notes that, over the last years, efforts have been made to advance the introduction of foreign language learning in the educative systems. “Content and Language Integrated Learning” (CLIL) has gained ground, especially in secondary education. The CLIL program consists of the use of the foreign language as the language of instruction in one or more subjects (Eurydice, 2001), where the main focus is to rely on the content rather than on the foreign language in which it is taught. The assumption behind this approach is that using the foreign language for content learning will enhance implicit learning of that language, leading to higher levels of proficiency than can normally be achieved in monolingual education (Admiraal, Westhoff & de Bot, 2005). It has been said that CLIL is about learning by construction, rather than learning by instruction (Wolf, 2002).

Unlike typical content-based instruction (CBI), which tends to focus on accuracy, CLIL programs foster fluency by increasing the communication between teachers and learners (García, 2009). Concerning this point, the study of Muñoz (2002) claims that one of the main advantages for teachers to use the second language as the medium of instruction and communication is that it makes the teacher more aware of the learners' linguistic needs, and that triggers tuned-in strategic language behavior, e.g. teachers' strategies to make the input comprehensible and context-embedded.

Another crucial point that distinguishes CLIL programs from other CBI teaching methodologies is that the former is more flexible and balanced about the potential of the L1 in CLIL classes. Although there is a growing body of literature that endorses the use of the L1 in bilingual education, monolingual L2 immersion education is still dominant in some countries (Lin, 2015), that is, using the second language as the only language of instruction. One of the hypotheses underlying monolingual L2 education is the maximum input hypothesis (Krashen, 1982) from second language acquisition studies. This hypothesis supports the idea that the student will achieve bilingualism faster and easier with the maximum amount of L2 input possible, rather than combining L2 classes with L1 classes and, thus, exposing the student to quantitatively less input. Alternatively, the idea of being educated through immersion classes avoiding any contact with the L1 in the class suggests a complete exclusion of the L1 and also the possible benefits of the L1 in L2 learning. Lin (2015) points out that this denigrating approach towards the use of the L1 usually comes from and is enforced by official top-down policies that try to legislate language use.

On the other hand, Lin (2013) provides a list of the functions that the L1 might have in bilingual education to help students improve their L2. These possibilities include using their L1 in functions related to unpacking the field of the study, that is, to translate, elaborate or exemplify some L2 content through the L1; textual functions, which include highlighting topic shifts and transitions; and finally interpersonal functions such as negotiating shifts in the methodology or narrowing the social distance with the students.

As CLIL education involves a wide range of opportunities and approaches to learn language through content (e.g. students exchanges, local projects, partial immersion, total immersion, etc., Mehisto et al., 2008), it is difficult to exactly delimitate the characteristics of this type of education. That is, the possible forms CLIL can take make it difficult to provide a definition that covers all of them and, thus, define with precision what CLIL education is. Cenoz et al. (2014) also point out that such an inclusive conceptualization of CLIL education makes it lack the identification of pedagogical tools or theoretical constructs which apply to most of the CLIL practices. Nevertheless, due to its broad approach, CLIL has also been claimed to increase learners' motivation and confidence in both the language and the subject; it diversifies methods and forms of classroom practice; it allows learners more contact with the target language and it complements other subjects rather than competes with them (Nuffic, 2016). Many European countries have included this type of practice in which instruction is provided in the L2 in their educational system, such as Germany (Gebauer, Zaunbauer, & Möller, 2013), Spain (Jiménez Catalán & Terrazas Gallego, 2005) or Finland (Merisuo-Storm, 2007) but also non-European countries such as Canada (Jared, Cormier, Levy, & Wade-Woolley, 2011) or China (Cheng, Li, Kirby, Qiang, & Wade-Woolley, 2010). CLIL programs may differ with respect to a number of characteristics, i.e. the starting point, the amount of instruction given in the L2 and the ratio of L1 instruction to L2 instruction at different stages within the program (Bergstrom, Klatte, Steinbrink & Lachmann, 2016). The increasing number of countries whose schools are implementing CLIL programs in their agendas has led to a renewed interest in its efficacy and measurement.

1.1.2 Performance advantages bilingual education yields

One of the biggest concerns of bilingual education was the possible unfavorable effect of the L2 on the L1. Several studies suggest that learning a second language is not detrimental to the development of the first one (Genesee, 2004, 2015). Goorhuis-Brouwer and de Bot (2010) ran a longitudinal study whose aim was to address, among other questions, whether early English language teaching (EELT) has a negative effect on L1. The participants of this study were different cohorts of children in several schools in Rotterdam and Groningen. The authors measured English and Dutch proficiency by means of the Reynell test (Edwards, Fletcher, Garman, Hughes, Letts, & Sinka, 1997), which is a standardized proficiency test. While the English version of the Reynell test includes both comprehension and production, the Dutch version concerns only language comprehension. After gathering data over a period of five years, the authors concluded that the results did not point to a negative effect of EELT on the Dutch language development, not even for children whose mother tongue was not Dutch. These findings were found consistently for different cohorts of children over time, from different schools in different regions and with different researchers and testers.

According to Bergstrom (2016), the aim of immersion programs is to achieve an additive bilingualism, where the L2 is learned while maintaining the L1 and that the fact that the L1 is not degraded is presumably due to students' daily contact with their L1 outside of school (e.g., at home, through media, keeping in touch with friends).

In addition, bilingual education not only is not harmful for the development of the first language but extensive research has focused on the linguistic skills bilingual education yields. Van der Leij, Bekebrede & Kotterink (2010) tested bilingual Dutch-English (N=27) and monolingual Dutch (N=25) primary school pupils at two different points in time within an interval of one year. At the time of the first measurement, the bilingual pupils were acquainted with English as a spoken language for about 3.5 years, and as a written language for about 1.5 years. The Dutch monolingual pupils, on the other hand, did not receive any formal instruction in English. The goal of this longitudinal study referred to the effect of bilingual instruction on L1 Dutch and L2 English reading and vocabulary proficiency in comparison to single language Dutch instruction. For that, the authors ran a variety of tests among the two different groups of pupils. Tests measuring Dutch proficiency concerned vocabulary via a standardized test developed by the Central Institute for Test Development (CITO in Dutch), phoneme awareness via a computerized task (Bekebrede, van der Leij, & Share, 2009), word reading fluency both within one minute (Brus & Voeten, 1997) and three minutes (Verhoeven, 1995), orthographic knowledge via an orthographic decision task developed by van der Leij & Morfidi (2006), orthographic awareness via an orthographic decision task (Siegel, Share, & Geva, 1995), and reading comprehension via a standardized test (Verhoeven, 1993). English proficiency was measured by almost the same tests: vocabulary via Peabody Picture Vocabulary Test (Dunn & Dunn, 1981), word reading fluency via one minute (Fawcett & Nicolson, 1996), loanword reading fluency (Schijf, 2006), orthographic knowledge via an orthographic decision task (Olson et al., 1994), and orthographic awareness via an orthographic decision task (Siegel et al., 1995). The results of this thorough study support the idea that bilingual instruction has a positive influence on L2 English reading and vocabulary performance compared with single language instruction. This positive effect was found on all variables with the exception of orthographic knowledge. The authors concluded that bilingual instruction contributed significantly to L2 vocabulary and L2 orthographic awareness.

Another study by Verspoor, De Bot & Van Rein (2011) measured the L2 proficiency in terms of better writing skills of those students immersed in secondary school bilingual programs compared to their peers in the monolingual program. Specifically, the authors were interested in the role out-of-school social media play in proficiency development. Data was gathered from four secondary schools in the Netherlands and was part of a cross-sectional longitudinal project aimed at the assessment of the effectiveness of bilingual education. Participants were in the first grade of secondary school (*N*=240) and the third grade of secondary school (*N*=316). In order to assess the starting proficiency level, a background questionnaire concerning the amount of contact with English and motivation and attitudes towards the language (Berns et al., 2007), a standardized reading test (only administered to the pupils from the first year, as it was too easy for third year pupils), a self-assessment of English proficiency (administered via can-do scales from the Common European Framework of Reference for Languages set) and CITO scores were collected. In order to measure L2 proficiency, the authors administered several tests both at the beginning of the school year and twice more during the year. That is, pupils from the selected schools were tested a total of three times a year. L2 proficiency was measured by a vocabulary test (Meara & Buxton, 1987) and six writing tasks (on personal topics and increased in difficulty over the years) during the project. The findings of this study revealed that the lack of input from the out-of-school media had a long term effect on the development of English vocabulary and writing skills. However, pupils immersed in a bilingual program who did not have contact with social media generally scored higher than the pupils in the monolingual program who had contact with social media. In other words, pupils in the bilingual program, although lacking English out-of-school input, were able to overtake monolingual pupils who had regular contact with social media outside the school. The authors finally observed that the relationship between the role of out-of-school and in-school input and developing proficiency is complex in the sense that language input is not a stable variable.

Admiraal et al. (2005) conducted a longitudinal study looking for the development of receptive vocabulary (by means of the English as a Foreign Language Vocabulary Test, developed by Meara (1992)), oral proficiency (through the Cito Oral Proficiency Test for English) and reading comprehension (via the national final examinations for English for intermediate general secondary education, MAVO) in pupils immersed in Dutch-English bilingual education. The development of four cohorts of students (*N*=1305) during 6 years was analyzed and compared to the results of students in the regular monolingual school program. The results for the receptive vocabulary tests revealed a significant higher score for those students immersed in bilingual education. However, the results seemed to point to the fact that, in general, students immersed in the bilingual program did not acquire English words at a faster rate than students of the control group. Regarding oral proficiency, bilingual students scored better than their peers in the control group for both general oral proficiency and pronunciation. Finally, the results for reading comprehension also showed an advantage for bilingual students as they scored significantly better than monolingual students. The initial score on the EFL Vocabulary Test appeared to have a significant effect on the reading comprehension test score.

1.1.3 Cognitive advantages bilingual education yields

Not only bilingual education has proven to be beneficial for the quality of the output in the L2 but also bilingualism has proven to grant cognitive advantages. These cognitive advantages in bilinguals have often been operationalized by a greater metalinguistic awareness compared to that of monolinguals. Metalinguistic awareness has been described as the ability to treat language as an object of thought (García, 2009). Bialystok (2006) points out that having two linguistic systems may trigger the structural patterns as a more noticeable feature and foster the attention of children to their systematic features.

Metalinguistic concepts should not be seen as a monolith and, thus, Bialystok (2001, 2004) divides metalinguistic awareness into word awareness, phonological awareness (awareness of the sound system) and syntactic awareness (awareness of the word order). Regarding word awareness, it seems like bilingual children are able to make a distinction between word and meaning before monolingual children. This abstraction goes through two steps, according to Bialystok (2006). The first one is segmentation of the speech into meaningful units, that is, words, which means that children have word awareness; the second one is the referential arbitrariness, in which children disentangle the conventional relation between a word and its referent. Bilingual children have been found to be better at this skill than monolingual children.

Concerning phonological awareness, a study by Eviatar and Ibrahim (2000) compared two groups of bilingual children and monolingual children in several phonological awareness tasks which included initial phoneme detection, final phoneme detection and phoneme-syllable deletion. Bilingual children outperformed the monolingual children in these tasks. This result arose despite the fact that bilingual children scored worse in the vocabulary test, which might indicate that vocabulary knowledge is not correlated with phoneme awareness and that it would constitute a poor predictor of phoneme awareness. However, there is a correlation between phonological awareness and orthography. A study of Cheung et al. (2005) worked with three groups of participants: a group speaking Cantonese and, therefore, reading characters, a group speaking Cantonese but reading both characters and an alphabetical system for written Cantonese, and a third group reading and speaking English. Results from this study revealed that orthography played an important role in phonological awareness as the highest scores pertain to the English speaking group and the lowest score to the group that didn't have an alphabetic written system, that is, the Cantonese group.

As far as syntactic awareness is concerned, the most widely used task to measure it is the judgment of acceptability of a sentence. Not only do these tasks work with bilingual children at the grammatical level, but they also perform better when the sentence is semantically anomalous. In a study by Moreno et al. (2010), they performed an EEG task on bilinguals who ended up showing a smaller cognitive conflict judging anomalous sentences than the monolingual group. In summary, bilingual children have been found to hold a more analytic approach on several features of the language in comparison to the capacities of monolingual ones.

A recent trend in research, pinpoints Baker (2001), is that studies nowadays do not focus their investigations on the cognitive performance of bilinguals compared to monolinguals but instead on the process of thinking. That is, how their logic operates as they learn two or more languages. Concerning this line of reasoning, a study by Gajo and Serra (2002) investigated how mathematics was taught in a bilingual class and compared their results with those of the monolingual class. As expected, they found that bilingual students scored slightly better than the monolingual students but that there were other differences. Both groups developed different ways of processing subject knowledge. On the one hand, the monolingual group's strength was informational knowledge, that is, the capacity to memorize knowledge. On the other hand, the strength of the bilingual group relied in operational knowledge, in other words, they were better at transferring and applying knowledge to new situations. Sorace (2005) specifies that a higher exposure to foreign languages from an early age is likely to affect L2 processing abilities, too. This is because, according to the author, people exposed to foreign languages from early ages have greater opportunities to integrate syntax and other knowledge of interpretation and production. It is not unexpected, then, that students in immersion classes obtain a higher L2 proficiency compared to students with conventional language-driven L2 instruction (Genesee, 2004). However, it is important to note, as Bialystok endorses, the idea that bilingualism itself is not determinant to reach academic achievement. Instead, the abilities yielded by bilingualism make it easier to approach the learning process.

So far, this study has focused on the attributes of general bilingual education. The section that follows will describe the functioning of the bilingual education in the Netherlands.

1.2 Dutch bilingual education

English plays a strong role in Dutch society nowadays. In terms of official education, English lessons were made compulsory in Dutch secondary schools in 1968 and, in 1986, in primary schools as well (Eurydice, 2001). Moreover, since the past decades, all levels of education in the Netherlands have been characterized by an increase in teaching English as Foreign Language (EFL) and also as English embedded in bilingual education. The first secondary school started to use English as the medium of instruction in September 1989; while in September 2005 around 70 schools for secondary education were using English as the medium of instruction in a large part of their curriculum, with more than 5000 students (Admiraal et al., 2005). This should not be a surprise as a Eurostat release (2014) reveals that many countries in the EU coincide in having English as, by far, the most common foreign language studied at lower secondary level.

The Dutch economy has been considered as one of the most important factors regarding English language development in the country, since a large part of its economical transactions are international (Edwards, 2016) and so English serves as the language of business, or as a lingua franca. However, Dutch culture is also strongly linked to English. Due to the new media (i.e., social media) and videogames, the use and presence of English has expanded among young people that now have a higher proficiency in English than ever before (Ammon & McConnell, 2002). According to Berns, De Bot and Hasebrink (2007), this trend of English development through social media among young users is expected to increase.

Turning now in detail to the structure of the Netherlands' secondary education, it is intended for children in the age group 12 to 16, 17 or 18, depending on the school type they choose: VMBO (V*oorbereidend Middelbaar Beroepsonderwijs*), pre-vocational secondary education, which lasts 4 years; HAVO (*Hoger Algemeen Voortgezet Onderwijs*), senior general secondary education, which lasts 5 years; VWO (*Voorbereidend Wetenschappelijk Onderwijs*), pre-university education which lasts 6 years.

The VWO stream, which prepares more academically able students for university, is itself divided into two streams (Nuffic, 2015): *atheneum*, that is, the regular VWO schools, and *gymnasium* (which includes Greek and Latin). These two streams can be taught following a monolingual approach or a bilingual approach. In the vast majority of the cases, the second language in the bilingual program is English (Edwards, 2016). A quarter of the VWO schools in the Netherlands are now reported to be bilingual schools (Dronkers, 2013). Likewise, bilingual secondary education has been increasingly implemented in the two streams left, HAVO and VMBO, although to a lesser extent (Nuffic, 2016). The successes of bilingual VMBO have proved that this education is not only suitable for brighter pupils.

Although with certain restrictions, secondary education Dutch schools are relatively free to design their bilingual curriculum. It is usual that lower secondary education offers a curriculum of 50% or more of the subjects taught in English, whilst in upper secondary education this ratio lowers and Dutch becomes, in most cases, the medium of instruction of the curricula again. The reason for this is that the final school-leaving examinations are in Dutch, so schools and parents tend to prefer this language over English in order to fully prepare students for those exams (Edwards, 2016).

The qualification of the faculty represents an important aspect to bilingual education curricula, as well. Schools that offer bilingual education typically rely on their regular (Dutch) staff, offering additional training to teach their subject matter in English (Eurydice, 2004). Dutch secondary school teachers teaching in bilingual programs are required to hold, at least, the Cambridge Advanced diploma for lower secondary education and the Cambridge Proficiency diploma as a must for upper secondary education.

The following section will discuss the influence the L1 has on the L2, especially in the vocabulary acquisition.

1.3 L1 effects on L2 vocabulary testing

A large and growing body of literature has investigated the numerous ways the vocabulary of the first language can affect the vocabulary learning process of the second language. Much of the previous research on this topic has fixed its attention on cognates. Cognates are words in two or more languages which are very similar to each other in terms of meaning, orthography, and/or phonetics. Moreover, there are other kinds of words that share orthographic and/or phonetic features but not meaning, which are the so called interlingual homographs, also known as “false friends” (Dijkstra, Grainger & van Heuven, 1999).

Schepens, Dijkstra, Grootjen & van Heuven (2013) compared frequency and orthographic and phonetic similarity of cognates of six languages (Dutch, English, German, French, Spanish and Italian) in order to shed light on the understanding of cognate processing. Their findings suggest that orthographic and phonetic similarity is higher between two typologically related languages than between two languages typologically unrelated. Further examination revealed that cognate frequency also correlates highly with language relatedness. In other words, the cognate frequency is larger between languages that share typology than between languages that don’t share typology.

These results are similar to those of Van der Slik (2010), who analyzed the data from 5763 people with 11 different West European languages as their first language on the acquisition of Dutch. Van der Slik worked with two types of measures –a cognate linguistic distance measure and a genetic linguistic distance measure– and concluded that the cognate linguistic distance has a larger explanative power than that of the genetic distance measure. That is, the L2 proficiency of learners with different mother tongues relied, to a great extent, on whether their first language shares more or fewer cognates with the target language. In the same line, Lindgren & Muñoz (2013) also identified young learners’ receptive language skills to be influenced by the cognate linguistic distance.

Since there are several methods to identify a cognate, Potatova (2016) tackled the differences between objective and subjective cognate identification approaches and determined which approach would impact findings of cognate effects. The author points out that the comparison of approaches is particularly important for bilingual children, as they show attenuated and inconsistent cognate effects and differences in methodology may impact whether a cognate advantage is found or not. In order to address this question, Potatova compared an objective criterion based on phonological overlap, two subjective criteria based on a translation elicitation task and a fourth hybrid criterion integrating objective and subjective standards. Cognate identifications were made in the Peabody Picture Vocabulary Test-III in Spanish and English. The findings of this study suggest that, as hypothesized, each method selects different subsets of test items as cognates and that the phonological overlap criterion identified significantly more cognates than any other approach.

These results support the methodology adopted for my study in that a phonological overlap (i.e. cognate distance) seems to be the best approach.

Taken together, these studies support the notion that cognates represent an important and very influential variable when it comes to L2 vocabulary learning. Likewise, it is also worth keeping the cognate advantage status in mind when carrying out cross-linguistic research. Furthermore, cognate understanding has been shown to facilitate word recognition (Hoshino & Kroll, 2007) as well as oral and written comprehension (August et al., 2005; Dessler et al., 2011; in Simpson Baird, 2016).

In addition, the frequency of the words in both languages also has been proven to affect their recognition. The PPVT assumes that certain words are easier to learn than others and that this fact is presumably due to its frequency (Dunn & Dunn, 2007). Therefore, it is thought that learners would acquire the most frequent words first rather than the most infrequent ones. However, as explained earlier, the first purpose of the English PPVT is to assess L1 vocabulary growth and the word frequency for English as the first language doesn't necessarily have to match with the frequency of those words of a native speaker of other language. In other words, the frequency of the words in English as an L1 might not be the same as the frequency of the words in English as an L2. This is the reason why the English PPVT administered to Dutch pupils, as this study does, might yield less reliable conclusions.

In summary, L2 vocabulary learning, and therefore testing, should be addressed carefully as, at least, two variables such as cognates and word frequency can play an important role.

1.4 This study

The aim of this study is to examine to what extent cognates in L1 Dutch vocabulary interfere with L2 English vocabulary testing in secondary education. This was operationalized by the effect of English-Dutch cognates displayed in the PPVT-IV. In addition, besides addressing the influence of cognates in vocabulary testing, this study also tackles the influence that L1 word frequency has on L2 word learning. It has been proposed that children are more likely to have knowledge of an L2 word that is highly frequent in their mother tongue than of a word that is less frequent in their L1 (Lobo, 2013). As indicated previously, the PPVT's first goal was to assess vocabulary in the native language. Nevertheless, as it is being administered to foreign speakers, the frequency of the words in both languages might not match and, thus, compromise the outcomes of the test. Moreover, it is questionable that the frequency of words in English native speakers coincide with that of learners of English. As was discussed in the section above, English in the EU represents a second language to a larger extent than a native language. Consequently, the words learned in the foreign language education or via social media do not have to be as frequent as in a L1 English context (Lindgren & Muñoz, 2013).

The overall research question that this study tries to answer is how the L2 vocabulary scores are affected by cognates and word frequency effects.

The first research question is whether Dutch words that show an orthographic and phonetic overlap with English pairs affect English vocabulary testing scores in secondary bilingual education. It is hypothesized that, indeed, cognates in Dutch and English will have a positive influence in the L2 vocabulary score.

The second research question is whether L2 English vocabulary testing is influenced by word frequencies. It is hypothesized that L1 word frequency will have a positive influence on the L2 vocabulary.

The third and final research question concerns whether the pupils in the bilingual *gymnasium* stream (that included Greek and Latin, although taught in Dutch) perform better at the English PPVT than pupils in the bilingual *atheneum* stream (which does not include Greek and Latin). As some of the PPVT items have a Hellenic or Romanic roots, it is hypothesized that there will be a positive effect of these subjects on the English vocabulary scores.

2. Method

2.1 Participants and school

The participants (*N*= 64; 35 women) for this study were recruited among the pupils of a secondary school in the south of the Netherlands. This secondary school offers two streams for the VWO students to choose: a monolingual one and a bilingual one, the latter one divided at the same time into the *atheneum* and *gymnasium* streams. The participants in this study are students of the bilingual *atheneum* and *gymnasium* streams plus a control group formed by the students in the monolingual stream who followed all of their subjects in Dutch. In this secondary school, the subjects taught in English are: Math, Geography, Physical Education, Biology, Chemistry, History, Religious Studies, Economics, and European and International

Orientation. These subjects are all taught in English in the first three years and comprise 60% of the instruction time. The Dutch language is reserved for Tutor group lessons, Latin and Greek (only offered in the *gymnasium* stream) and the modern foreign languages (Spanish, French and German), although in these latter subjects the target language as the language of instruction is encouraged. From the fourth year onwards, the total time spent instructing in English drops to 30%. The reason for this is because the pupils must take the Dutch school-leaving exams at the end of their school career, as explained earlier (Edwards, 2016).

For this study, data from the development of English vocabulary in bilingual and monolingual pupils were gathered during three years. As they were tested in the first three years of their secondary education, the ratio of English-based instruction is, as mentioned, still at its highest, representing 60% of the time. Finally, parents of participating children gave informed consent for their participation.

2.2 Materials and procedure

The test used in this study is the Peabody Picture Vocabulary Test (PPVT). The PPVT is a widely spread test aimed to measure hearing passive vocabulary (Dunn & Dunn, 2007). Although this test was first designed as a tool to evaluate L1 vocabulary proficiency, it is very commonly used to test L2 learners in their L2 vocabulary acquisition. This evolution on the administration of the PPVT raises questions as to whether it is appropriate to apply a tool dedicated to L1 vocabulary to L2 learners, and what results it yields in this context. Especially when such an important variable as the cognate advantage in L2 learning has not been studied and implemented in the results of the PPVT of L2 participants. As previously explained, this thesis attempts to answer that question.

The PPVT consists of several training items and 228 test items arranged in 19 sets of 12 items each sorted by increasing difficulty. According to the technical specifications of the PPVT, items were reviewed and empirically analyzed for difficulty, validity and freedom from bias with respect to sex, ethnicity, geographic region and socioeconomic status (SES). Each item is made of four colored pictures displayed together with an audio of an English word pronounced by a native speaker. Target items can be verbs, nouns or adjectives. The task of the participant is choosing the picture that best describes the word uttered in each item.

The PPVT is suitable for infants of 2;6 years upwards and it has established starting points according to the age of the participant. The set in which the participant commits no more than one mistake represents the basal set. When the participant makes eight or more mistakes in one set, the participant has reached the ceiling set and the test is concluded. Although this is the standard administration procedure, the school in which the data was collected followed a slightly different method which was not based on the participants' age in order to establish the starting point. Instead, despite their age, all participants started the test from item 1, so as to assess all their vocabulary thoroughly.

The modality employed for gathering data for the third year was the Form A of the computer-based Peabody Picture Vocabulary Test, 4th edition (PPVT-IV). However, for the first and second year the data was gathered by a teacher of the school using the paper-based Peabody Picture Vocabulary Test, 4th edition. Whilst in the computer version the items are presented on a screen and the uttered word is an audio displayed together with the item, the paper version consists of a book whose pages represent each item and the target word is uttered by the tester, in this case also a native speaker of English. Pupils were individually tested in a quiet room in their school.

2.2.1 Translations

In order to establish cognates in English and Dutch, translations needed to be done. Dutch translations from item 1 to item 168 were taken from Goriot et al. (in prep). They translated those items comparing the meaning of the word in the *Longman Dictionary of Contemporary English for Advanced Learners* (Longman, 2012) to the meaning described in the online version of the *Van Dale Dutch-English translation dictionary* (Albers, 2015). If the description of a translation matched the description of the original English word, that word was selected. Items from 169 on were translated using the *Van Dale English-Dutch translation dictionary*, following the explained methodology. Translations were double checked by a Dutch native speaker. In addition, verbs in the PPVT-IV are presented in their –ing form such as *incarcerating* or *replenishing*. Since this kind of verbs that express ongoing actions are very uncommon in Dutch (in this case the translations would be *opsluiten* and *aanvullen*, respectively), it was decided to take the stem of the verb rather than their full form. Thus, we worked with *incarcerate* or *replenish* and, as a result, *opsluit* and *aanvul*. Following this method, the translation of items was unproblematic until the last set. Set 19 presents a great number of low frequency words that have no direct translation into Dutch. That is the reason that some of the last items present no translation.

Several words had more than one possible translation so they were individually studied and collated by a bilingual English-Dutch native speaker, indicating which one should be more accurate based on their meaning. Besides this criteria for selecting the right translation, the frequency of words in both languages was also taken into account. This means that when a word had two possible semantically alike translations, the translation whose frequency better matched the original word frequency was chosen. For example, for the English word *submerge* two Dutch translations were originally made. One of them was *onderwater* and the other one was *onderdompelen*. A second Dutch native speaker closer inspected the meaning of the two words and decided that *onderdompelen* represented a more accurate translation of the English word *submerge* than *onderwater*. In this case, the meaning was the determinant factor for choosing the right translation. However, frequency also played a role. For the English word *angler*, whose frequency was 0.37, two Dutch words were selected: *visser* and *hengelaar.* Both words have the same meaning as *angler* but *visser* had a frequency of 4.2 per million and *hengelaar* had a frequency of 0.02 per million. As both words matched in meaning but *hengelaar* was closer in frequency, it was decided that the chosen translation would be *hengelaar*.

In order to determine the frequency of words in both languages, two corpora were used. The SUBTLEX-US corpus (Van Heuven, Mandera, Keuleers, & Brysbaert, 2014) was used for English words and the SUBTLEX-NL (Keuleers, Brysbaert, & New, 2010) for Dutch words. The advantage of these corpora over others is that they are based on film and television subtitles instead of edited text, therefore, based on spoken language instead of written texts. Studies from Brysbaert, Buchmeier, et al. (2011) and Brysbaert, Keuleers & New (2011) suggested that word frequencies based on film and television subtitles are better predictors of word processing times than those of written texts such as books. As a result, these corpora appeared to be more appropriate for this project. However, in a similar way that translations raised some problems in the last set, the establishment of the frequency was also problematic for the last sets. As stated above, these corpora are based on film and television subtitles, and, thus, entails some impediments in the highly specialized words: the corpora do not register their frequency because they are barely used in spoken language. Another possible explanation for the missing frequencies is that the words in the higher sets were of such a low frequency that they were unlikely to occur in any corpus at all, including written text corpora. This being so, some items from Set 17 onwards (although a single item in Set 15 too) showed no frequency, for example the word *weerhaak* in Set 17 or *achteroverliggend* and *bloemkelk* in Set 19. These words didn't appear in the SUBTLEX-NL corpus and, therefore, were assigned with a frequency of 0. Nevertheless, whereas these translations didn't appear in the SUBTLEX-NL corpus, their original English word did appear in the SUBTLEX-US corpus. That is *barb*, *supine* and *calyx*, respectively, although with an extremely low frequency.

In summary, Dutch translations were chosen according to two main criteria: the corresponding meaning of the Dutch word (double checked by at least two native Dutch speakers) and the similarity of the frequency between both of them, when available.

2.2.2 Phonetic transcriptions

In order to determine the cognate distance, phonetic transcriptions were also needed. English transcriptions as well as Dutch ones up to item 168 were taken from Goriot et al. (in prep). The rest of the phonetic transcription of Dutch translations were taken from *Uitspraakwoordenboek* (Heemskerk & Zonneveld, 2000). The *Longman Pronunciation Dictionary* (Wells, 2008) was used for the English transcriptions. Similarly as with the translations and the frequency, not all the phonetic transcriptions could be found in the *Uitspraakwoordenboek* and the *Van Dale Dutch dictionary* was also used for the missing phonetic transcriptions. Despite the two sources consulted, some words from Set 16 onwards do not have any transcription because neither source provide it. This resulted in a Levenshtein distance of 0, as it will be explained in the next section.

2.2.3 Cognate distance

It has been demonstrated that bilingual children use sound similarities to identify cognates (Kieffer & Lesaux, 2008). In the study conducted by Dressler et al. (2011), Spanish pupils in an English bilingual program were asked to infer the meaning of unknown English words. Students often referred to the sound similarities in cognate pairs and, the authors added, pronounce the English word with the Spanish phonology. This finding suggests that sound is a source of information that serves as a cue for cognate identification.

The Levenshtein distance, which is a distance based on phonetical and orthographical resemblance, seemed appropriate to apply in this study. The Levenshtein distance was calculated and used as discrimination measure to determine whether Dutch and English words were cognates or not. The Levenshtein distance is an algorithm that calculates the minimum number of insertions, deletions, and substitutions that are needed to edit one string to become another (Schepens et al., 2013). Recent studies have successfully used this distance as in Heeringa (2004), Kondrak & Sherif (2006) or Yarkoni, Balota & Yap (2008). The Levenshtein distance algorithm is expressed as LD = 1- $\frac{\text{distance}}{\text{length}}$. *Distance* stands for the total number of insertions, deletions and substitutions and *length* stands for the total number of letters the longest word in the pair has. The resulting value ranges between 0 (no overlap at all) and 1 (completely alike). It was applied for both the phonetic and the orthographic strings. For example, the orthographic LD of the English-Dutch pair *bus-bus* resulted in 1 as both orthographic forms are the same. However, its phonetic distance is 0.67 as they are not pronounce the same and the differences in their pronunciation, /bʌs/ in English and /bʏs/ in Dutch, respectively, are operationalized as $1 - \frac{1}{3} = 0.67$. For this study, following Schepes et al. (2013), if the result of that logarithm is equal or higher than 0.5 it was considered a cognate.

2.3 Analysis

Analyses were computed with the statistical program SPSS version 23. In the first place, the P-values of all the items according to year (first, second and third) and type of education (*gymnasium*, *atheneum* and the control group) were calculated. The P-value is an indication of how difficult an item is. Expectations are that items in the higher sets are more difficult items, as the PPVT is arranged that way and that the overall difficulty of the test is perceived as lower over the years. Afterwards, in order to examine whether cognates influence their performance on L2 English vocabulary testing, it was calculated how many cognates the PPVT-IV contains in each set and put it in relation with the P-values, which will reveal if cognates were perceived easier or difficult than average items. Thirdly, the relation English and Dutch word frequency was investigated and a plotted figure was created to illustrate it. Similarly to the cognates and P-values, frequencies and P-values were correlated to know their relationship. Finally, a linear regression was run in order to determine which variables (orthographic cognates, phonetic cognates, English frequency, Dutch frequency) influence the P-value of the items.

3. Results

3.1 P-value

In order to assess the overall difficulty of the PPVT-IV, the P-value was calculated. The P-value ranges from 0 (very difficult) to 1 (very easy). Figure 1 shows an overview of the average P-value per set.

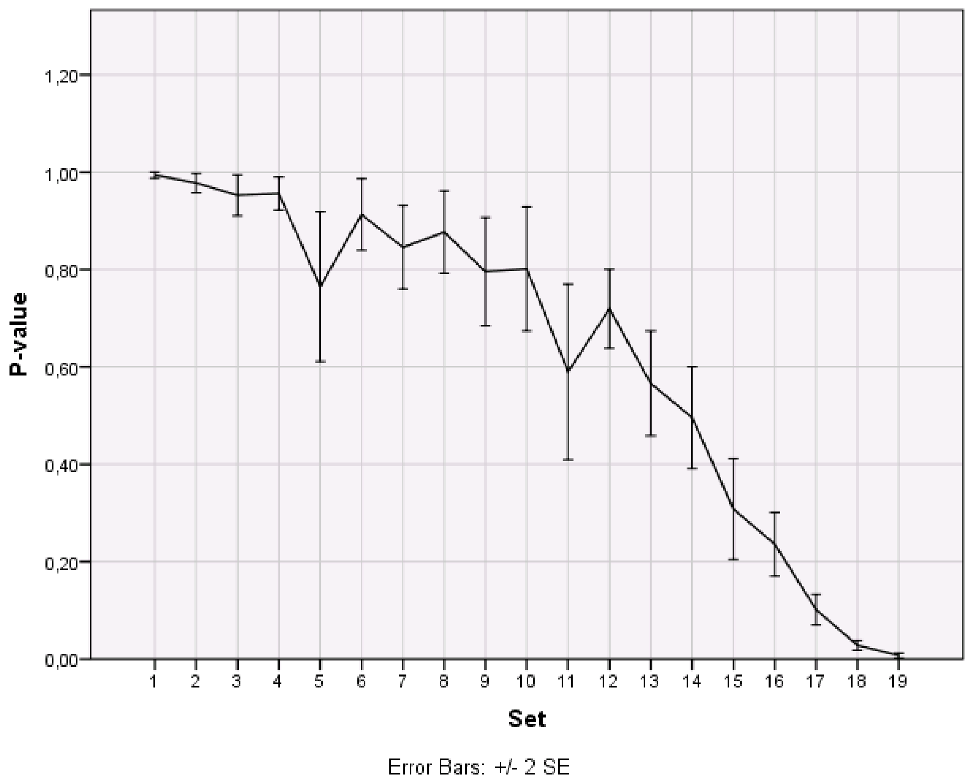

Figure 1. *Average P-value per set.*

As expected, the perceived difficulty drops along the sets, beginning with an almost perfect performance (1) and ending with almost complete incomprehensibility (0). Although the P-values follow a general decreasing line, it is noticeable that it is not consistent and has some clear deviations. For example, Set 5 is noticeably more difficult than Set 4 but also more difficult than Set 6, when the test is meant to display progressive difficulty. This results in Set 4 and Set 6 having a similar difficulty. One of the reasons Set 5 dramatically slumps might be because it includes the English word *vest*, which refers to a sleeveless garment whereas in Dutch, *vest* refers to a cardigan. That is, this word represents a false friend. The item in which this word appears displayed four images among which a vest and a cardigan were included and pupils often chose the image depicting the cardigan. This means that the fact that a false friend is included in Set 5 might have increase its difficulty. The word *envelope* also appears in Set 5. Although its English frequency is relatively high (2.40), its Dutch frequency is null. It seems possible that Dutch pupils didn't know the word *envelope* due to its low Dutch frequency. Another word that is low frequent in Dutch is the English word *buckle*. This word has an English frequency of 1.8 but its translation, *gesp*, has a frequency of only .67. The incorporation of the false friend plus the low Dutch frequency of these two words might have contributed in the perceived difficulty of Set 5.

Set 11 represents another deviation as it is more difficult than Set 10 but concurrently also more difficult than Set 12. Closer inspection to this particular set revealed many words with low Dutch frequency. For example, the word *inflated* translated into Dutch as *opgepompt* has a Dutch frequency of .17. Another low frequency word that appear in Set 11 is *tusk* whose Dutch translation *slagtand* has a frequency of .19. These are two of the lowest Dutch frequency words that appear in Set 11 and, just as *envelop* and *gesp* in Set 5, they might have contributed to its perceived difficulty.

Figure 2 shows the average P-value per set but it distinguishes between years. It is expected that pupils in year 1 find the PPVT more difficult than in the next two years. However, as the following graph illustrates, this difficulty drops and remains almost the same during year 2 and 3. This fact suggests that there is a vocabulary boost between year 1, and years 2 and 3. That is, while it seems to be a boost in the English vocabulary of the Dutch pupil in year 1, the vocabulary growth seems to be less clear between years 2 and 3.

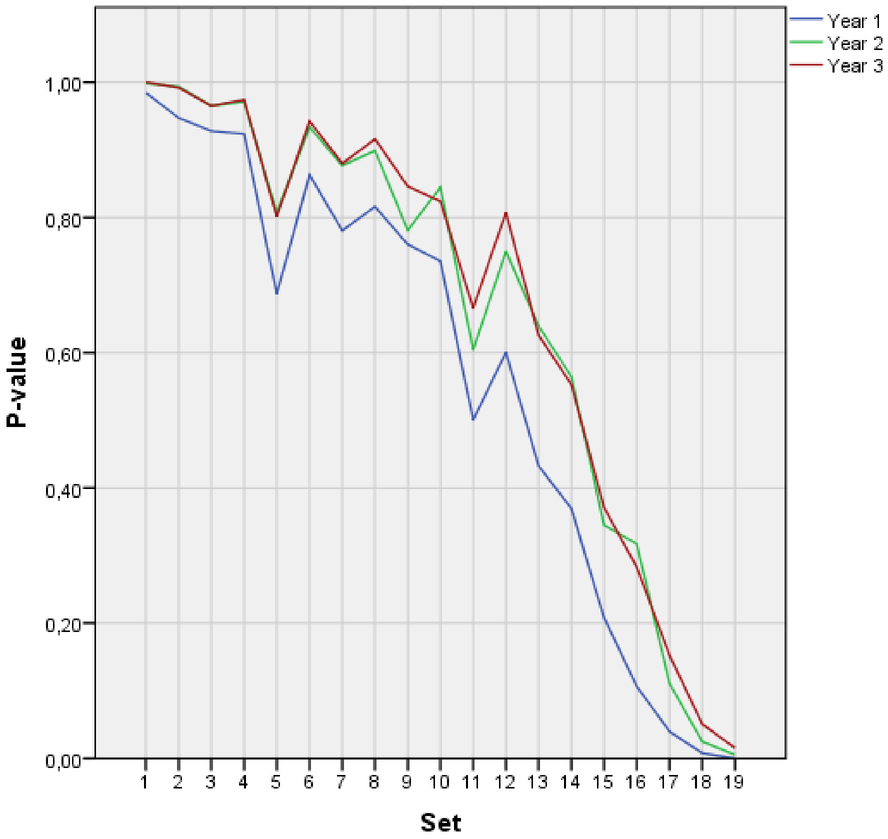

Figure 2. *Average P-value per set and per year.*

Table 1 is useful for understanding more accurately the vocabulary acquisition progress pupils go through in year 1. Whereas in year 1 a pupil following the *atheneum* stream, for example, scored on average a P-value of .53, in year 2 and 3 the same pupil score a P-value mean of .61 and .66, respectively. This means that the growth is larger from year 1 to 2 (.08) than from 2 to 3 (.05).

Table 1. *Descriptive statistics of the P-value per stream and year.*

	N	*M*	*SD*
Atheneum 1	13	,53	,38
Gymnasium 1	7	,63	,39
Control 1	15	,51	,38
Atheneum 2	31	,61	,38
Gymnasium 2	20	,69	,35
Control 2	2	,64	,41
Atheneum 3	28	,66	,37
Gymnasium 3	19	,72	,33
Control 3	20	,60	,38

Table 1 shows a small predicament in that the control group in year 2 has a higher mean that the bilingual *atheneum* stream in the same year. This can be due to the fact that the number of pupils in the control group in year 2 is only 2 and with a large standard deviation. These two facts combined can explain the incongruence.

Closer inspection of Table 1 reveals that inside every year, scores are better for those pupils in the *gymnasium* stream followed by the *atheneum* stream and, finally, those pupils in the monolingual stream —the control group— are the ones that scored lower in the test. Table 1 shows this trend. That is, *gymnasium* pupils perceived the test as less difficult than *atheneum* pupils and, as expected, than the pupils embedded in a monolingual program.

These results are in line with my third research question, for which it was hypothesized that students following the bilingual *gymnasium* program would score better than those in the bilingual *atheneum* program. The only academic difference between these two streams is that the former one includes the subjects of Greek and Latin, although in Dutch, and the latter does not. Arguably, the reason for this slightly better scoring for *gymnasium* students is that the knowledge of Hellenic and Romance words represents an advantage when it comes to English vocabulary recognition. As English and Dutch share typology, both are Germanic languages, Germanic English words can be easier for Dutch students to recognize. However, the Hellenic and Romance English words present in the PPVT might be easier for students that are in contact with other languages sharing that typology, that is Greek and Latin from the *gymnasium* stream. The clearest example that illustrates best this fact is the word *pentagon*. This word has a Hellenic root (*penta-* “five” and *–gonos* “angle”). Whereas this item has a mean P-value of .68 in the *gymnasium* stream, the mean P-value in the *atheneum* stream is .25. The control group score a mean P-value of .09 in this word. Similar, although not as high, contrast between P-values per stream are the words *transparent* (*gymnasium* mean P-value of .95 and *atheneum mean* P-value of .60), *dissect* (*gymnasium* mean P-value of .95 and *atheneum* mean P-value of .71), *archeologist* (*gymnasium* mean P-value of .98 and *atheneum* mean P-value of .76) or the word *quintet* (*gymnasium* mean P-value of .48 and *atheneum* mean P-value of .28). All these words share a common Hellenic or Romanic root. It is remarkable that the pupils from the *atheneum* stream didn’t use some of the orthographic and phonetic similarities for some of these words as *archeologist* or *quintet*, whose orthographic LD with their Dutch translations are high (.67 and .71, respectively). Figure 3 shows the performance per stream in a more visual way.

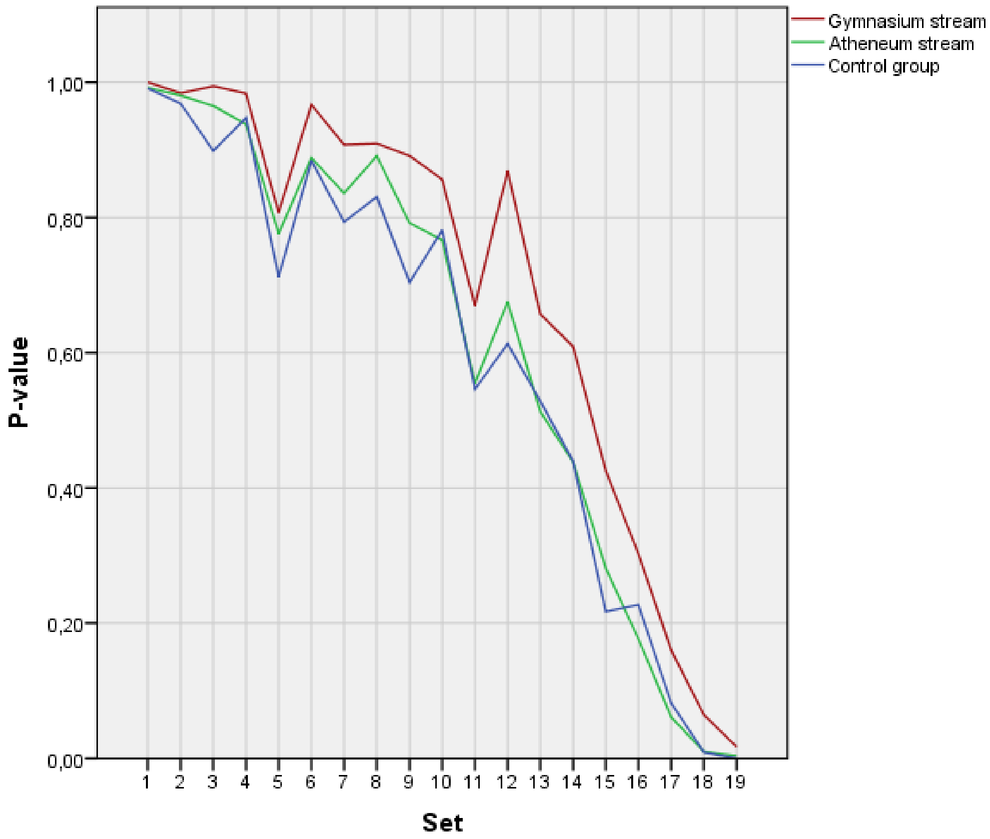

Figure 3. *Average P-value per set per stream.*

3.2 Cognates

My first research question concerns the role that cognates play in the performance of English vocabulary recognition. For that, I calculated both the orthographic and phonetic overlap that English-Dutch words have in the PPVT. Table 2 represents an overview of the descriptive statistics of the orthographic and phonetic Levenshtein Distance.

Table 2. *Descriptive statistics of the orthographic and phonetic Levenshtein Distance (LD).*

	N	*Minimum*	*Maximum*	*M*	*SD*
Orthographic LD	228	,00	1,00	,37	,33
Phonetic LD	228	,00	,80	,23	,23

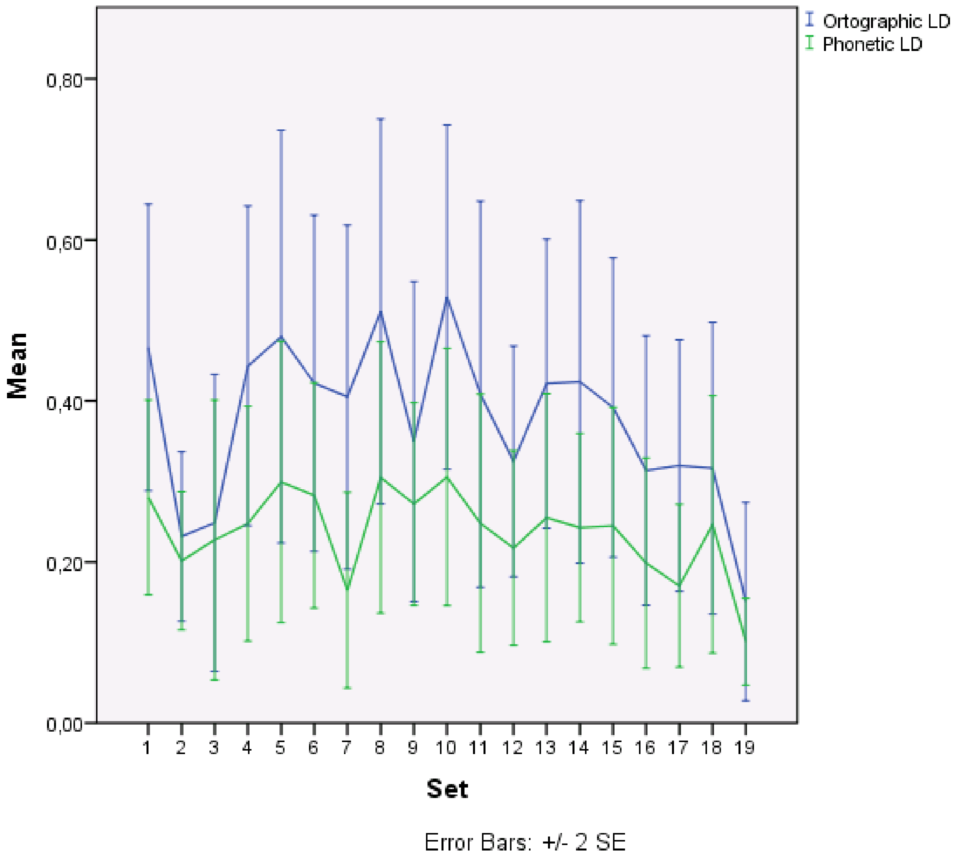

Figure 4. *Mean of the orthographic and phonetic overlap between English-Dutch words per set.*

Figure 4 shows that English and Dutch words in the PPVT present a stronger similarity in their orthography and, to a lesser extent, in their phonology. Although the supremacy of the orthographic over the phonetic resemblance remains the same, the difference between these two variables is not the same along the test. Whereas in Set 2 and Set 3 the difference is very little, in Sets 7, 8 or 10 this difference is remarkably bigger.

Some examples of words that are very similar in their orthography but not so much in their phonology are the pair *digital-digitaal* whose orthography LD is .88 but phonetic LD (/ˈdɪdʒɪtl/ and /diɣit'al/, respectively) is .29; the pair *aquarium-aquarium* with an orthographic LD of 1 and phonetic LD of .44 (/əˈkweəriəm/ and /akw'arijʏm/, respectively); or the pair *archeologist-archeoloog* with an orthographic LD of .67 and phonetic LD (/ɑːkiˈɒlədʒɪst/ and / ɑrxejol'ox/) of .08. However, as can be expected from the graph in Figure 5, there are not so many words whose phonetic distance is higher that its orthography distance. For example, the pair *dance-dans* has an orthography LD of .60 and a phonetic LD of .80 (/dɑːns/ and /dɑns/);

or the pair *cupola-koepel* with an orthography LD of only .17 but a much higher phonetic LD of 0.63 (/ˈkjuːpələ/ and /ˈkupəl/). Nevertheless, the difference of distances between these kinds of words is not as big as the difference between words with higher orthographic LD and lower phonetic LD.

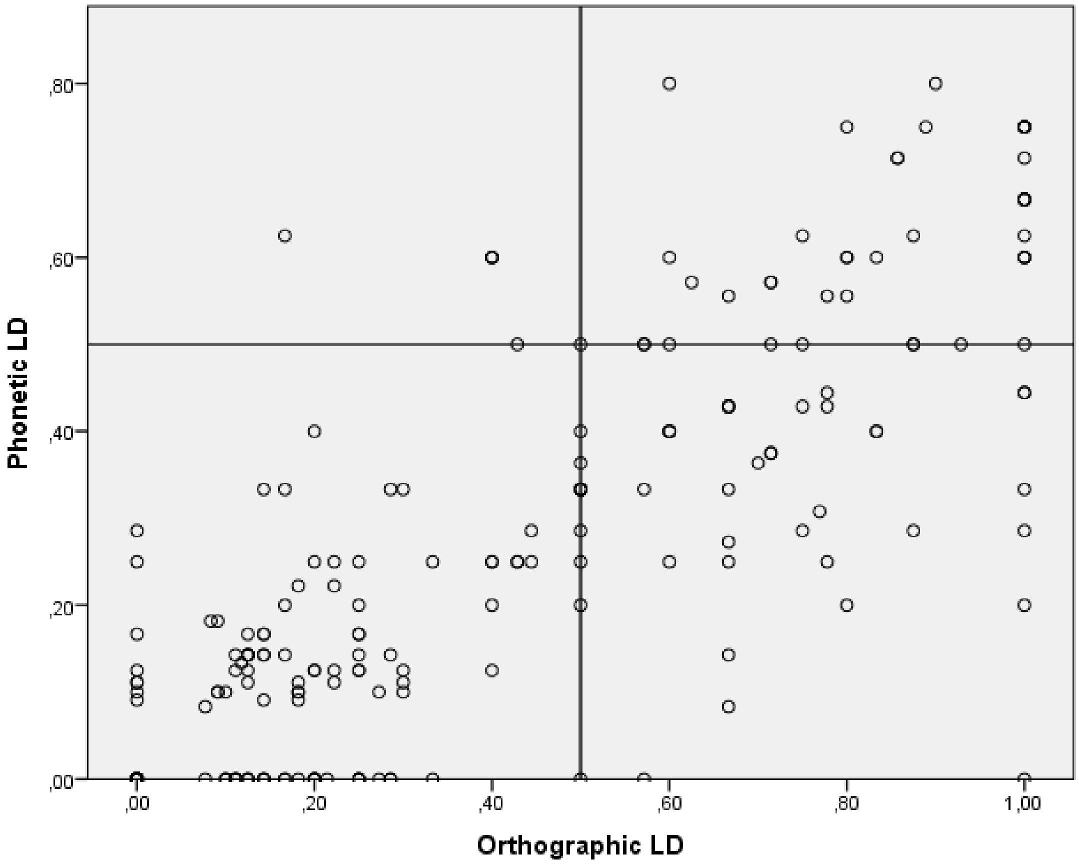

Figure 5. *Correlation between the orthographic and phonetic Levenshtein Distance.*

Table 3 represents an overview of the number of orthographic and phonetic cognates per set. In total, there are 85 orthographic and 47 phonetic English-Dutch cognates in the PPVT-IV. The cognate variability is rather high as in Set 2, for example, there are no phonetic cognates, whereas in Set 5 there are 5. Regarding orthographic cognates, the maximum number of cognates per set is 7, as identified in Set 1 and Set 8 whereas the lowest number of cognates per set is 1, in Set 19.

Table 3. *Number of orthographic and phonetic cognates per set.*

	Set 1	*Set 2*	*Set 3*	*Set 4*	*Set 5*	*Set 6*	*Set 7*	*Set 8*	*Set 9*	*Set 10*	*Set 11*	*Set 12*	*Set 13*	*Set 14*	*Set 15*	*Set 16*	*Set 17*	*Set 18*	*Set 19*
Orthographic cognates (N)	7	3	3	6	6	5	5	7	5	6	5	3	6	5	3	3	3	3	1
Phonetic cognates (N)	2	0	3	2	5	4	2	4	2	4	2	2	2	3	3	3	1	3	0

In order to investigate whether the pupils perceive cognate items as more easy than non-cognate items, therefore resulting in a higher P-value, it is necessary to know the relationship between cognates and P-values. Table 4 shows the correlation between the two types of cognate as the independent variable and the overall and year-by-year P-value as the dependent variable.

Table 4. *Correlation between Levenshtein Distances and P-values.*

	Orthographic LD	*Phonetic LD*
Orthographic LD	1	
Phonetic LD	,814**	1
P-value year 1	,254**	,242**
P-value year 2	,233**	,201**
P-value year 3	,256**	,229**
Overall P-value	,252**	,227**

** $p < .01$

Results from Table 4 reveal that there is a positive relationship between LDs and P-values. In other words, the higher the LD, and therefore the higher the similarity between the English and Dutch word, the higher the P-value, that is, the easier the item is perceived by the pupil. The next two figures, Figure 6 and Figure 7, show the correlation between orthographic LD and the average P-value, and phonetic LD and the average P-value, respectively.

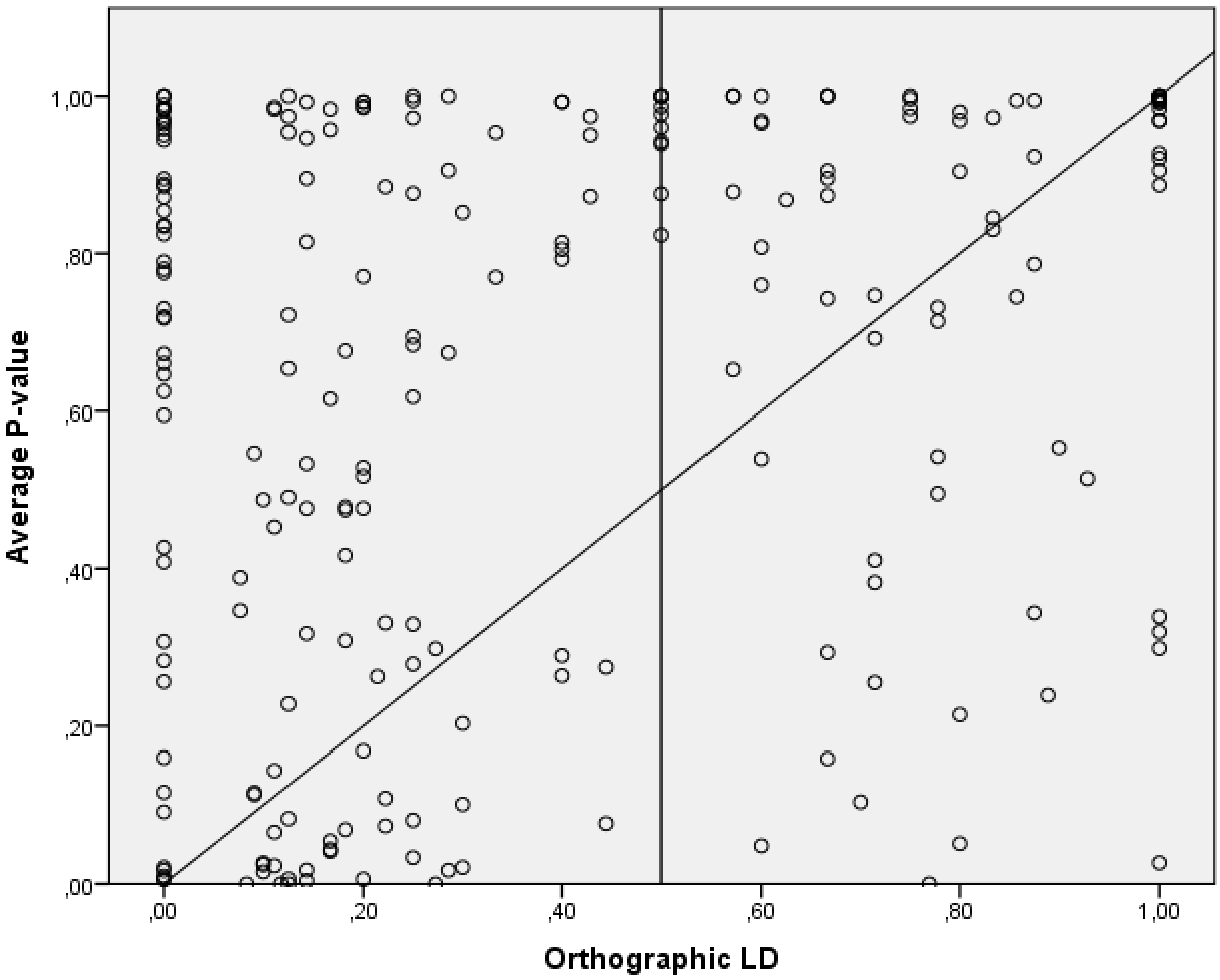

Figure 6. *Correlation between the orthographic Levenshtein Distance and the average P-value of the items.*

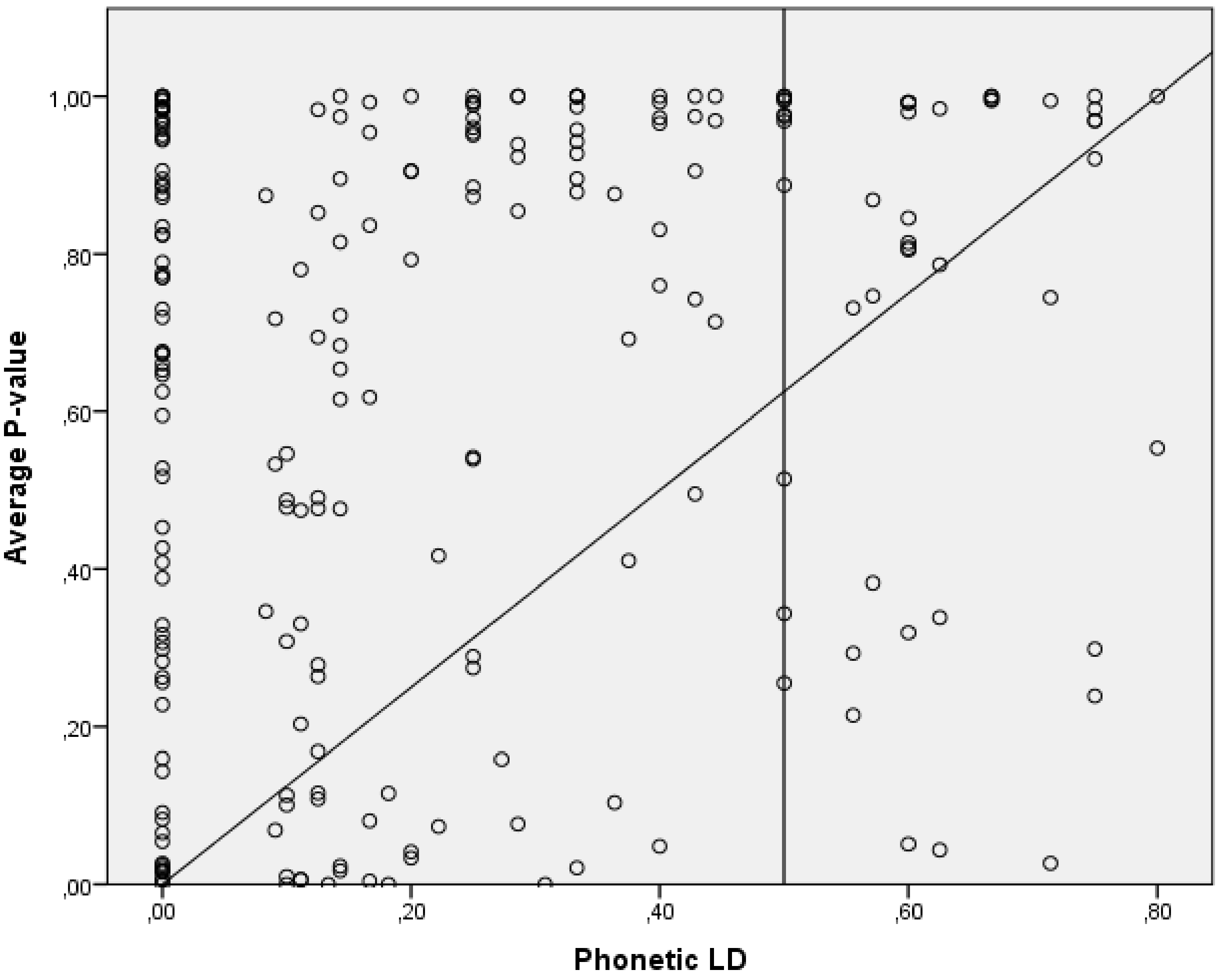

Figure 7. *Correlations between the phonetic Levenshtein Distance and the average P-value of the items.*

The comparison between the last two scatterplots reveals that there are more orthographic cognates (a high P-value and a high LD) than phonetic ones, which is consistent with results in Table 3. It also shows that false friends (a low P-value and a high LD) are more infrequent than cognates for both orthography and phonology. For example, the upper right dots in Figure 6 represent words such as *kiwi, aquarium, hyena, flamingo, cactus, tunnel* or *panda*, which are written exactly the same in English as in Dutch, and that present an orthographic LD of 1 and also an average P-value of 1. In these cases the cognate effect arises clearly. However, there are less numerous cases of orthographic false friends in the lower right part of the scatterplot. The most eminent example of an orthographic false friend is the word *vest*. As previously explained, this string of letters has different meanings in English and Dutch. Whereas in English represents a sleeveless garment, the Dutch word *vest* refers to a cardigan. This is the reason why the item *vest* has a contrast so high in its orthographic LD (1) and its average P-value (.30). Other words that are not false friends but have a high orthographic LD and a low P-value are *convex* (LD of 1 and P-value of .03), *sedan* (LD of 1 and P-value of .32), *dromedary* (LD of .89 and P-value of .24), *primate* (LD of .80 and P=value of .05)*, terpsichorean* (LD of .77 and P-value of 0) or *cerebral* (LD of .71 and P-value of .26). Although these strings of words have the same meaning in Dutch, and therefore not represent a false friend per se, pupils in the three years found them difficult. The reason for this inconsistency is not clear but it can be that their frequency in Dutch is low and, therefore, impede them for knowing their translation in English. *Convex* (*convex* in Dutch) has an English frequency of .18 and a Dutch frequency of 0, *sedan* (*sedan* in Dutch) has an English frequency of 1.02 and Dutch frequency of .86, *dromedary* (*dromedaris* in Dutch) not only has a low Dutch frequency of .13 but an English frequency even lower (.08), *primate* (*primaat* in Dutch) has a Dutch frequency of .26 and an English frequency of .52, *terpsichorean* (*terpsichore* in Dutch) has an extremely low frequency in English (.02) and a null frequency in Dutch, and finally *cerebral* (*cerebraal* in Dutch) has also a very low frequency in Dutch of .16. although higher in English (2.02).

Regarding phonetic cognates, words such as *dance, lamp* or *penguin*, besides a high orthographic LD, they also have a high phonetic LD. The English-Dutch pair *dance-dans* has a phonetic LD of .80 and a P-value of 1, the pair *lamp-lamp* have a phonetic LD of .75 and a P-value of 1, and the pair *penguin-pinguin* have a phonetic LD of .71 and a P-value of .99. Although scarce, there are also items with a high phonetic LD but low P-value. Examples of these words are *convex* (phonetic LD of .75 and P-value of .03), *reprimand* (phonetic LD of .80 and P-value of .55), *cerebral* (phonetic LD of .75 and P-value of .24), *vest* (phonetic LD of .75 and P-value of .30) or *cupola* (phonetic LD of .63 and P-value of .04). As the orthographic low frequency words, these words might show a high contrast between phonetic LD and P-value because their frequency is too low. *Convex* has a Dutch frequency of 0 and an English frequency of 0.18, *reprimand* has a Dutch frequency of .26 and an English frequency of .43, whereas *cupola* has an English frequency of 0 whereas its Dutch frequency is somehow higher, 2.19.

As explained, the cognate factor might not be the only reason why the pupils perceive the English word as easier. It also can be possible that, even not being a cognate and not having any phonetic or orthographic resemblance, they know the English word because it is very frequent. The opposite case is also possible: a cognate might not be recognized if its frequency is too low (as happened with *primate, dromedary* or *cerebral*, among other words). The following section describes the role frequency plays in L2 vocabulary testing.

3.3 Frequencies

As the PPVT is a test whose difficulty increases over the course of time of the test, it is expected that this difficulty comes along with lower frequency words. Figure 5 compares the average frequency per set of both English and Dutch words.

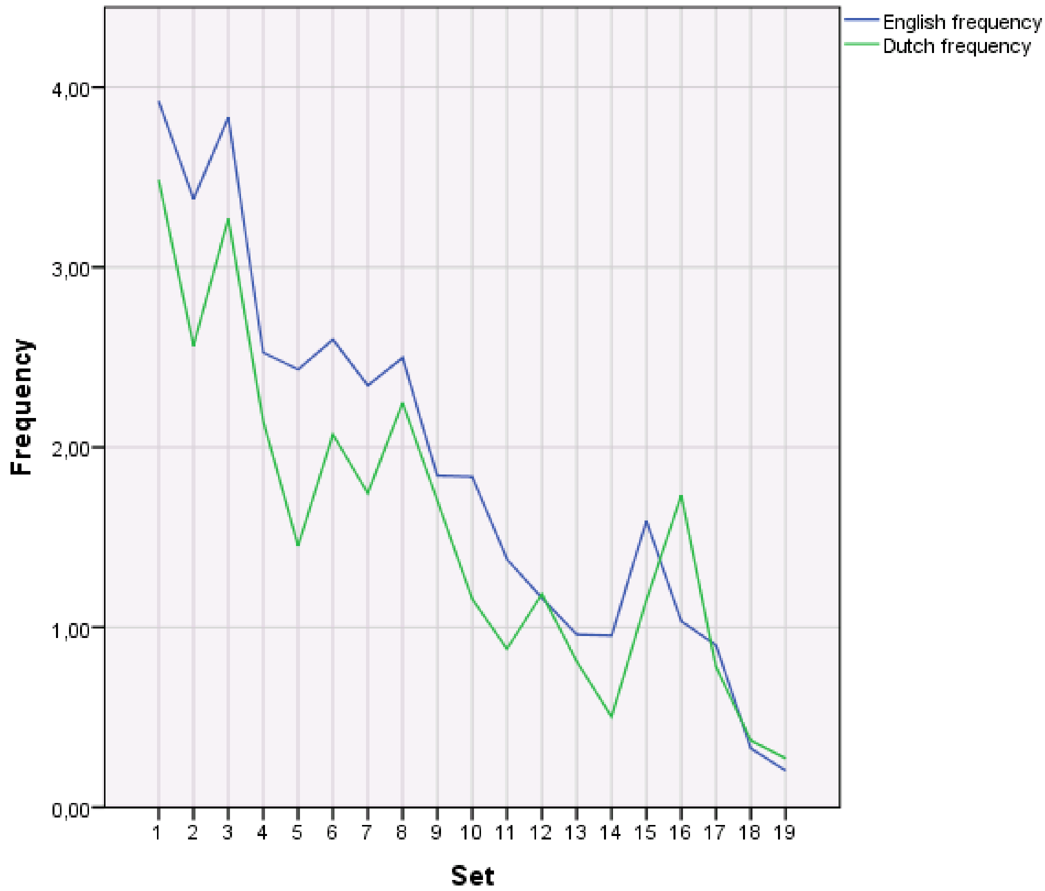

Figure 8. *Average frequency per set.*

As expected, items in the PPVT decrease in frequency in the higher sets, which makes it increasingly difficult. Nevertheless, the descent is not constant nor gradual. Whereas the frequency from Set 1 to Set 2 drops, it rises again from Set 2 to Set 3 and then dramatically slumps until Set 5. Overall, these frequency peaks are present along all the test. In addition, although the English word frequency is higher than the Dutch word frequency in almost all the sets, the final ones, from Set 16 onwards, Dutch frequency seems to be slightly higher than English frequency.

The table below illustrates the correlation between frequencies as the independent variable and P-values as the dependent variable.

Table 5. *Correlations between frequency and P-values.*

	English frequency	*Dutch frequency*
English frequency	1	
Dutch frequency	,618**	1
P-value year 1	,650**	,447**
P-value year 2	,617**	,395**
P-value year 3	,606**	,402**
Overall P-value	,634**	,421**

** $p < .01$

From the data in Table 5, it can be seen that frequencies in both English and Dutch significantly correlate with the overall P-value and the P-values in each year. This means that the more frequent a word is in English, the more easy it is for the pupil to recognize it. At the same time, the more frequent a word is in Dutch, the greater the probability of the pupil knowing the word in English. As previously mentioned in Figure 7, it also appears that low frequently cognates, even if they share a high resemblance with their Dutch translations, are not recognized as such by pupils. Earlier explained words such as *convex, sedan* or *dromedary* were not recognized as cognates because of their low frequency in their L1 Dutch. It can be hypothesized, then, than word frequency can affect the identifications of words as cognates or non-cognates. These results are in line with my second research question, which hypothesized that the frequency of words would have a positive effect on the knowledge of English words.

Although correlations between frequencies seems high, in order to closely inspect this relationship, Figure 9 represents a scatter plotted correlation comparing both frequencies.

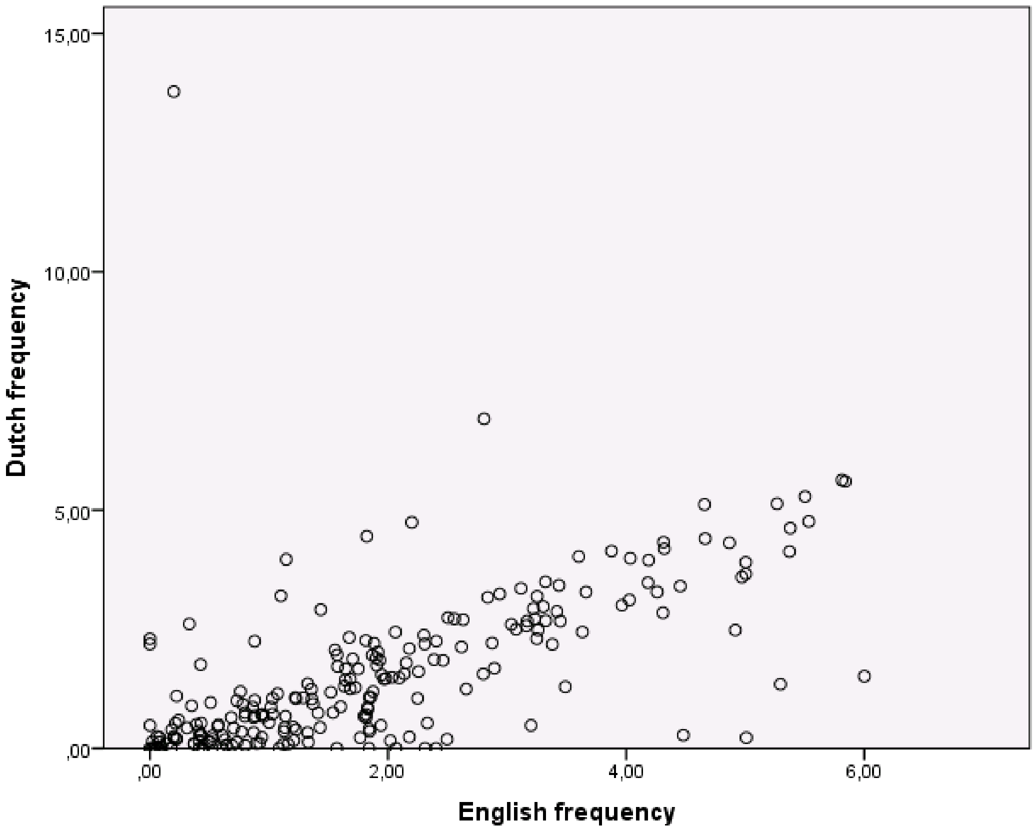

Figure 9. *Correlation between English word frequency and their translation in Dutch.*

As can be seen, in general Dutch and English frequencies correlate positively between each other, which suggests that overall the low frequency words in English correlate with the Dutch translations chosen. However, there is one prominent exception. This is the word *incarcerate*, in Set 16, whose chosen Dutch translation was *opsluit*. However, while *incarcerate* has a frequency of .20, *opsluit* has a frequency of 13.78, much higher. This difference could have been avoided if the Dutch translation would have been *inkerkeren*. This word has a much lower frequency, just as *incarcerate*, and would not have been considered an outlier. It can be expected that this extremely high frequency of the word *opsluit* also affected the average Dutch frequency in Set 16 in Figure 8.

There are some other exceptions as the English word *saw* with a frequency of 6 whose Dutch translation *zaag*, nevertheless, has a frequency of 1.51. Another example would be the English word *sort* with a frequency of 5.01 while *sorteer*, its Dutch translation, has a frequency of only .22. Or the English word *lever* and its translation *hefboom*, both with a frequency of 3.20 and .48, respectively.

The scatterplot also shows that the contrast can also happens vice versa, that is, there are some words whose frequency in Dutch is higher than in English. The most representative example of this phenomenon is the word *net*: while in English it has a frequency of 2.81, in Dutch its frequency rises up to 6.91. This can be due to the fact that the word *net* in English only has one meaning, an openwork fabric, while the same word in Dutch has multiple meanings, most of them also highly frequent. This means that all those meanings convert in one word and, thus, bloat its frequency. Another word that has a relatively high contrast between frequencies is the word *detonation*, *ontploffing* in Dutch, whose English and Dutch frequencies are 1.82 and 4.45, respectively.

3.4. Regression analyses

Regression analyses were made in order to determine which variables influence the P-value of the items. Table 6 provides the results of the linear regression with P-values as the dependent variable, the overall P-value as well as in each year, and word frequency and orthographic and phonetic distances as predictors.

Table 6. *Linear regression with the log P-values as the dependent variable and frequency and LDs as predictors.*

	Overall P-value		*P-value year 1*		*P-value year 2*		*P-value year 3*	
	B	SE	B	SE	B	SE	B	SE
Orthographic LD	,998**	,293	1,024**	,301	,994**	,315	,998**	,289
Phonetic LD	-,196	,419	-,234	,430	-,367	,451	-,224	,414
English frequency	,503**	,049	,510**	,050	,539**	,052	,478**	,048
Dutch frequency	,063	,044	,066	,045	,042	,048	,056	,044
R Square	,521		,514		,491		,501	

As Table 6 indicates, only the orthographic LD and the English frequency are significant predictors of the P-values. This fact suggests that pupils from the first three years of secondary school rely more on the similarities between the English-Dutch orthography than on the similarities on their phonology. These results might be surprising since the PPVT is a test administered orally and the only cue they have in order to score right is the uttered word, not a written version of it. However, pupils in secondary education already have knowledge of the English orthography and, thus, rely more on this feature since it shares more characteristics with Dutch orthography than the Dutch phonology with the English phonology.

Table 7. *Linear regression with the significant predictors with the Stepwise method.*

	Overall P-value		*P-value year 1*		*P-value year 2*		*P-value year 3*	
	B	SE	B	SE	B	SE	B	SE
Orthographic LD	,879**	,169	,883**	,174	,780**	,182	,864**	,167
English frequency	,545**	,038	,553**	,039	,564**	,041	,514**	,038
R Square	,516		,509		,488		,497	

Table 7 shows a linear regression carried out with the stepwise method. This is a method that is constantly being reassessed each time a predictor is added to the equation in order to remove any non-significant predictor. Using this method, only the individual contribution of the predictors that explain best the distribution remain. The results of the regression analysis show that more than 50% of the variation in the overall P-value can be explained by the orthographic LD and the English frequency.

4. Discussion

The present study was designed to determine the effects of English-Dutch cognates and word frequency in L2 English vocabulary testing in English-Dutch bilingual pupils during their first three years of secondary education. The Peabody Picture Vocabulary Test (4th edition) was administered to pupils in year 1, 2 and 3 of their secondary education.

This project aimed to answer three research questions. The first one concerned the effect English-Dutch cognates have on the L2 English vocabulary acquisition. It was hypothesized that cognates would have a positive effect on L2 vocabulary acquisition. For this question, the orthographic and phonetic Levenshtein Distance between the English words and their Dutch translation was calculated. The Levenshtein Distance is a distance based on the orthographic and phonetic resemblance between two words. The closer this distance is to 1, the more similar they are. The hypothesis of a cognate effect advantage is confirmed by results showed in Table 4, where a positive correlation between orthographic and phonetic similarities and P-values is displayed. These results suggest that the closer orthographically and/or phonetically an item of the test is, the easier this item is perceived to be by the pupil. This fact stands for the first three years of secondary education that this study has worked with. These findings are in line with those of Van der Slik (2010) and Lindgren and Muñoz (2013) whose studies suggest that receptive language skills are affected by the cognate effect. Closer examination, as displayed in Table 6, revealed that pupils in the three years relied more on the orthographic similarities than in the phonetic similarities between English-Dutch word pairs. This finding might result striking since the PPVT is an oral-administered test and, therefore, could exist the possibility that the uttered words in the test serve as a more reliable cue than its orthography, a feature that does not appear in the test. However, pupils in secondary education are already familiarized with written English and it can be possible that the pupils relied more on the English orthography since it shares more characteristics with Dutch orthography. As the correlation between the orthographic LD and the phonetic LD is very high (r = .814), it is expected that the reliance on the orthographic LD has a relationship with the reliance on the phonetic LD. Therefore, although the phonetic distance didn't arise as a significant predictor according to the analysis made, it is still possible that it played an important role as a predictor of the P-value.

The second research question concerned what effect L1 English word frequency would have over the L2 Dutch word vocabulary recognition. It was hypothesized that the more frequent a word is, the more easy the word would be for the pupil. In order to establish the word frequency in those languages, English and Dutch frequencies per million were extracted from the corpus SUBTLEX-US (Van Heuven, Mandera, Keuleers, & Brysbaert, 2014) and SUBTLEX-NL (Keuleers, Brysbaert, & New, 2010), respectively. The results displayed in Table 5 support the second hypothesis. There, it can be seen that English and Dutch word frequencies correlates positively with the overall P-value as well as the individual P-value per year. This finding indicates that the more frequent a word in English is, the more easy it is for the pupil to know this English word. In addition, it suggests that the more frequent a Dutch word is, the greater the probability of the pupil knowing its English translation. These results are consistent with data observed in Schepes et al. (2013). As Table 6 indicates, further examination revealed that English frequency, and not Dutch frequency, is a significant predictor of the P-value. A possible explanation for this might be that, as mentioned previously, social media plays an important role in Dutch society nowadays (Edwards, 2016). It can be possible that these participants benefitted from this fact and that the amount of English in social media could serve as a reliable source of L1 English input for Dutch pupils. There are, however, other possible explanation. Since the Netherlands is one of the countries in which children are more exposed to English via subtitled television and films (Edwards, 2016), and since the corpus SUBTLEXT-US is based on American movies and TV programs, it seems possible that participants benefitted from this match and that, therefore, the frequency of L1 English words appear as a good predictor of their L2 English vocabulary. On the other hand, analyses revealed that, although English and Dutch frequencies are correlated (r = .618), Dutch frequency is not a significant variable in order to predict the P-value of the items, as Table 7 displays.

Taken together, these two first research questions and respective answers point to a clear effect of cognate and frequency variables in L2 vocabulary testing. Although the orthographic distance arose as a significant predictor for the difficulty of the items, there are also possibilities that the phonetic distance plays a role in its perceived difficulty as well. On the other hand, English frequency arose as a significant predictor of the difficulty of the items.

Finally, the third and last research question asked whether the pupils in the bilingual stream *gymnasium*, that includes Greek and Latin in its curriculum, perform better at the English PPVT than the pupils following the bilingual *atheneum*, that does not include Greek and Latin. Results from this study point to a better performance of the former group of students compared to the latter group. Table 1 and Figure 3 support the idea that this particular stream find the items in the PPVT-IV easier, and thus perform better than the pupils in the *atheneum* stream. According to this data, an implication that can be inferred is that subjects such as Greek and Latin provide Hellenic and Romance typological knowledge, respectively, that pupils are able to apply to English vocabulary testing. The PPVT-IV contains several words that do not share a Germanic etymology but do share a Romance or Hellenic one, such as *pentagon*, *transparent* or *archeologist*. Therefore, it is likely that pupils with knowledge of different typological languages use this cue and apply it in order to recognize L2 words with Hellenic and Romance roots and perform better than pupils that do not have a different typological background (as Dutch and English are both Germanic languages). However, there might be another possible explanation. Since *gymnasium* students scored better at the CITO test, it cannot be ruled out the possibility of interference of this variable. It would be interesting to assess other differences between the *atheneum* and *gymnasium* stream. It would be useful if further studies administer a language background questionnaire to pupils in both streams in order to determine whether there are differences in the English use of pupils of both streams and, if so, investigate whether it is a significant characteristic that has an effect on their L2 English vocabulary proficiency.

In summary, the results of this study strengthens the idea that there is a relationship between L2 English vocabulary acquisition and cognates and frequency. On the one hand, cognates have been shown to facilitate new vocabulary learning due to the phonetic and orthographic similarities between the Dutch and English words. This result further support the idea of many other studies such as those of Schepens et al. (2013), Van der Slik (2010) or Lindgren & Muñoz (2013), as previously reviewed. On the other hand, frequency results suggest that words that are more frequent in both languages are likely to be learned first.

As mentioned in the literature reviewed, the PPVT is a test that measures the receptive vocabulary. As such, it cannot predict performance of productive vocabulary. However, there are studies suggesting that the development of receptive vocabulary usually precedes productive vocabulary in language acquisition (Webb, 2008). Bearing in mind this fact, and since the comprehension and the production of a language are usually correlated (Zink & Lejaegere, 2002; Bornstein & Hendricks, 2012), it is expected that the English production scores of the pupils immersed in the bilingual program also outperform those of the students in the monolingual program.

Further research is required to investigate into a deeper level the reasons why *gymnasium* pupils perform better than *atheneum* pupils. In future investigations, it might be advisable to assess pupils from both streams via a language background questionnaire about their English use outside the school and contrast their answers by stream. The findings from the third research question raise intriguing questions regarding the nature and extent of Hellenic and Romance roots serving as an advantage in English or when does this effect start to arise. A study taking these questions into account would help to shed light on the phenomenon of transferring knowledge from one language to another.

5. Conclusions

The purpose of the present study was to examine the effect of L1 Dutch on L2 English vocabulary testing in a secondary education school in the Netherlands. This influence was operationalized as (1) the cognate effect between English-Dutch and (2) the frequency of words in both languages. This study has found that these two parameters affect L2 vocabulary testing through the PPVT-IV. On the one hand, the more an L2 word resemblance to their translation in the L1 in terms of phonetics and orthography the easier for the participants to correct identify them. On the other hand, L2 words that were frequent in their L1 were easier to identify than those words that were less frequent. This study has also shown that those words that are infrequent in their L1, regardless their orthographic and phonetic resemblance, will not be identified as cognates due to its low frequency, as Figure 6 and Figure 7 showed.

The evidence from this study suggests that the outcome from the PPVT in bilingual English-Dutch participants might be compromised and the standardize test might be biased due to English-Dutch cognate and frequency effects that other pairs of languages do not share. In addition, this study has raised important questions concerning the nature of the vocabulary tests and suggesting that the PPVT seems to be less appropriate to perform cross-linguistics studies, as the influence L1 Dutch on L2 English vocabulary testing may be not the same as the influence other mother tongue has over L2 English vocabulary testing. Although the PPVT is still an appropriate tool for measuring vocabulary proficiency if the mother tongue keeps the same, further cross-linguistic research investigating vocabulary growth should bear in mind these results.

6. References

Admiraal, W., Westhoff, G., & de Bot, K. (2006). Evaluation of bilingual secondary education in the Netherlands: Students' language proficiency in English. *Educational Research & Evaluation, 12*(1), 75-93.

Albers, M. (2015). *Van Dale middelgroot woordenboek* (Second edition). VBK Media.

Ammon, U., & McConnell, G. D. (2002). *English as an academic language in Europe*. Frankfurt: Peter Lang.

August, D., Carlo, M., Dressler, C., & Snow, C. (2005). The critical role of vocabulary development for English language learners. *Learning Disabilities Research & Practice*, *20*, 50–57. doi:10.1111/j.1540-5826.2005.00120.x

Baetens Beardsmore, H. (1994). "Language Shift and Cultural Implications in Singapore." In S. Gopinathan, Ho Wah Kam, A. Pakir, and V. Saravanan (eds.). *Language, Education and Society in Singapore: Issues and Trends*. Times Academic Press, Singapore, pp. 47-64.

Baker, C. (2001). *Foundations of Bilingual Education and Bilingualism*. Third edn. Multilingual Matters, Clevedon.

Bekebrede, J., van der Leij, A., & Share, D. L. (2009). Dutch dyslexic adolescents: Phonological core variable orthographic differences. *Reading and Writing: An Interdisciplinary Journal, 22*, 133–165.

Bergström, K., Klatte, M., Steinbrink, C., & Lachmann, T. (2016). First and second language acquisition in German children attending a kindergarten immersion program: A combined longitudinal and cross-Sectional study. *Language Learning, 66*(2), 386-418.

Berns, M., de Bot, K., & Hasebrink, U. (2007). *In the Presence of English. Media and European Youth*. Berlin: Springer.

Bialystok, E. (2001). *Bilingualism in Development: Language, Literacy, and Cognition*. Cambridge University Press, New York.

Bialystok, E. (2004). "Language and Literacy Development." In T.K. Bhatia and W.C. Ritchie (eds.). *The Handbook of Bilingualism*. Blackwell, Malde, MA, pp. 577-601.

Bornstein, M. H., & Hendricks, C. (2012). Basic language comprehension and production in >100,000 young children from sixteen developing nations. *Journal of Child Language, 39*(4), 899–918. doi: 10.1017/S0305000911000407

Brus, B. T., & Voeten, M. J. M. (1973). *Een-Minuut-Test [One-Minute-Test]*. Nijmegen, The Netherlands: Berkhout.

Brysbaert, M., Buchmeier, M., Conrad, M., Jacobs, A. M., Bölte, J., & Böhl, A. (2011). The word frequency effect: A review of recent developments and implications for the choice of frequency estimates in German. *Experimental Psychology, 58*, 412–424.

Brysbaert, M., Keuleers, E., & New, B. (2011). Assessing the usefulness of google books' word frequencies for psycholinguistic research on word processing. *Frontiers in Psychology, 2*, 27.

Cenoz, J., Genesee, F., & Gorter, D. (2014). Critical analysis of CLIL: Taking stock and looking forward. *Applied Linguistics, 35*, 243–262. doi:10.1093/applin/amt011

Cheng, L., Li, M., Kirby, J.R., Qiang, H., & Wade-Woolley, L. (2010). English language immersion and students' academic achievement in English, Chinese and mathematics. *Evaluation & Research in Education*, *23*(3), 151-169.

Cheung, H., Chen, H. C., Lai, C. Y., Wong, O. C., & Hills, M. (2001). The development of phonological awareness: Effects of spoken language experience and orthography. *Cognition, 81*, pp. 227-41.

Dijkstra, T., Grainger, J., & van Heuven, W. J. B. (1999). Recognition of Cognates and Interlingual Homographs: The Neglected Role of Phonology. *Journal Of Memory And Language*, *41*(4), 496-518. doi:10.1006/jmla.1999.2654

Dressler, C., Carlo, M. S., Snow, C. S., August, D., & White, C. E. (2011). Spanish-speaking students’ use of cognate knowledge to infer the meaning of English words. *Bilingualism: Language and Cognition*, *14*, 243–255. doi:10.1017/S1366728910000519

Dronkers, J. (2013, July 17). Sluipenderwijs wordt Nederland tweetalig. De Volkskrant. Retrieved from www.volkskrant.nl/vk/nl/3184/opinie/article/detail/3476812/2013/07/17/Sluipenderwijswordt-Nederland-tweetalig.dhtml

Dunn, L. M., & Dunn, D. M. (2007). *PPVT-4 manual*. Bloomington: NCS Pearson, Inc.

Dunn, L. M., & Dunn, L. (1981). Peabody picture vocabulary test—revised. Circle Pines, MN: American Guidance Service.

Edwards, A. (2016). *English in the netherlands : Functions, forms and attitude*. Amsterdam: John Benjamins Publishing Company.

Edwards, P., Fletcher, P., Garman, M., Hughes, A., Letts, C., & Sinka, I. (1997). *The Reynell developmental language scales*. Windsor: NFER-Nelson.

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

Eurobarometer. (2012). Europeans and their Languages. *European Commission*.

European Commission. (2008). Multilingualism: An Asset for Europe and a Shared Commitment: Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions.

European Council. (2002). *Barcelona European Council 15-16 March 2002: Presidency Conclusions.*

Eurostat. (2014). English, French and German still most common foreign languages studied at lower secondary level in the EU28 in 2012… but Spanish learning has increased more. *European Commission*.

Eurostat. (2016). 60% of lower secondary level pupils studied more than one foreign language in 2014. *European Commission*.

Eurydice. (2001). Foreign language teaching in schools in Europe. *European Commission*.

Eurydice. (2004). Content and Language Integrated Learning (CLIL): At school in Europe. The Netherlands: National Description 2004/05. *European Commission.*

Eurydice. (2012). Key data on teaching languages at schools in Europe. *European Commission*.

Eviatar, Z. & Ibrahim, R. (2000). Bilingualism is as bilingual does: Metalinguistic abilities of Arabic - speaking children. *Applied Psycholinguistics, 21*, pp. 451-71.

Fawcett, A. J., & Nicolson, R. I. (1996). *Dyslexia screening test*. London: The Psychological Corporation.

Gajo, L., & Serra, C. (2002). "Bilingual Teaching: Conecting Language and Concepts in Mathematics." In D. So & G. Jones (eds.). *Education and Society in Plurilingual Contexts*. VUB Brussels University Press, Brussels, pp 75-95.

García, O. (2009). *Bilingual education in the 21st century: A global perspective*. Malden, MA: Wiley-Blackwell.

Gebauer, S. K., Zaunbauer, A. C. M., & Möller, J. (2013). Cross-language transfer in English immersion programs in Germany: Reading comprehension and reading fluency. *Contemporary Educational Psychology*, *38*, 64–74. doi:10.1016/j.cedpsych.2012.09.002

Genesee, F. (2004). What do we know about bilingual education for majority-language students? In T. K. Bhathia & W. C. Ritchie (Eds.), *The handbook of bilingualism and multiculturalism* (pp. 547–576). Malden, MA: Blackwell.

Genesee, F. (2015). Myths about early childhood bilingualism. *Canadian Psychology/psychologie Canadienne, 56*(1), 6-15. doi:10.1037/a0038599

González, J. A., & Barbero, J. (2013). Building bridges between different levels of education: Methodological proposals for CLIL at university. *Language Value, 5*(1), 1–23.

Goorhuis-Brouwer, S., & de Bot, K. (2010). Impact of early English language teaching on L1 and L2 development in children in Dutch schools. *International Journal of Bilingualism, 14*(3), 289-302.

Goriot, C., van Hout, R., Broersma, M., Unsworth, S., & McQueen, J. M. (in prep). L1-effects in an L2 receptive vocabulary task: the PPVT-4.

Heemskerk, J. S., & Zonneveld, W. (2000). *Uitspraakwoordenboek*. Het Spectrum.

Heeringa, W. (2004). *Measuring dialect pronunciation differences using Levenshtein distance* (Groningen dissertations in linguistics, 46; Groningen dissertations in linguistics (Grodil), 46).

Hoshino, N., & Kroll, J. F. (2007). Cognate effects in picture naming: Does cross-language activation survive a change of script? *Cognition*, *106*, 501–511. doi:10.1016/j.cognition.2007.02.001

Jared, D., Cormier, P., Levy, B. A., & Wade-Woolley, L. (2011). Early predictors of biliteracy development in children in French immersion: A 4-year longitudinal study. *Journal of Educational Psychology*, *103*, 119–139. doi:10.1037/a0021284

Jiménez Catalán, R. M., & Terrazas Gallego, M. (2005). The Receptive Vocabulary Of English Foreign Language Young Learners. *Journal of English Studies*, *5-6*, 173–191.

Keuleers, E., Brysbaert, M., & New, B. (2010). SUBTLEX-NL: A new measure for Dutch word frequency based on film subtitles. *Behavior Research Methods*, *42*(3), 643–650. doi:10.3758/BRM.42.3.643

Kieffer, M. J., & Lesaux, N. K. (2008). The role of derivational morphology in the reading comprehension of Spanish-speaking English language learners. *Reading and Writing*, *21*, 783–804. doi:10.1007/s11145-007-9092-8

Kondrak, G., Sherif, T. (2006). Evaluation of several phonetic similarity algorithms on the task of cognate identification. Proceedings of the Workshop on Linguistic Distances. Sydney: Association for Computational Linguistics. 43–50. Retrieved from: http://dl.acm.org/citation.cfm?id = 1641983

Krashen, S. (1982). *Principles and practice in second language acquisition*. Oxford: Pergamon.

Lin, A. (2013). Classroom code-switching: Three decades of research. *Applied Linguistics Review, 4*, 195–218. doi:10.1515/applirev-2013-0009

Lin, A. (2015). Conceptualising the potential role of l1 in CLIL. *Language, Culture and Curriculum, 28*(1), 74-89.

Lindgren, E., & Muñoz, C. (2013). The influence of exposure, parents, and linguistic distance 30 on young European learners' foreign language comprehension. *International Journal of Multilingualism, 10*(1), 105–129. doi:10.1080/14790718.2012.679275

Lobo, V. R. (2013). *Teaching L2 English at a very early age : a study of Dutch schools*. Radboud University. Retrieved from http://www.lotschool.nl

Longman. (2012). *Longman Dictionary of Contemporary English for Advanced Learners* (6th Edition). Pearson Education.

Meara, P. (1992). *EFL vocabulary test*. Swansea, UK: Centre for Applied Language Studies.

Meara, P., & Buxton, B. (1987). An alternative to multiple choice vocabulary tests. *Language Testing, 4*: 142-154.

Mehisto, P. (2008). CLIL counterweights: Recognising and decreasing disjuncture in CLIL. *International CLIL Research Journal, 1*, pp. 93–119.

Merisuo-Storm, T. (2007). Pupils' attitudes towards foreign-language learning and the development of literacy skills in bilingual education. *Teaching and Teacher Education, 23*(2), 226-235. doi:10.1016/j.tate.2006.04.024

Moreno, S., Bialystok, E., Wodniecka, Z., & Alain, C. (2010). Conflict resolution in sentence processing by bilinguals. *Journal of Neurolinguistics, 23*, pp. 564-79.

Nuffic. (2015). Education system in The Netherlands. Retrieved from: https://www.nuffic.nl/en/publications/find-a-publication/education-system-the-netherlands.pdf

Nuffic. (2016). Bilingual education in Dutch schools: a success story. Retrieved from: https://www.nuffic.nl/en/publications/find-a-publication/bilingual-education-in-dutch-schools-a-success-story.pdf

Olson, R. K., Forsberg, H., Wise, B., & Rack, J. P. (1994). Measurement of word recognition, orthographic, and phonological skills. In G. R. Lyon (Ed.), *Frames of reference for the assessment of learning disabilities* (pp. 243–277). Baltimore: Brookes Publishing Co.

Potapova, I. (2016). Cognate identification methods: Impacts on the cognate advantage in adult and child spanish-English bilinguals. *International Journal of Bilingualism, 20*(6), 714-731. doi:10.1177/1367006915586586

Schepens J, Dijkstra T, Grootjen F, & van Heuven WJ. (2013). Cross-language distributions of high frequency and phonetically similar cognates. *PloS One*, *8*(5), E63006. doi:10.1371/journal.pone.0063006

Schijf, G. M. (2006). Leenwoordenlijst Engels [Loanwords English]. Retrieved from www.muiswerk.nl/downdoc/leenwoorden-2006.doc.

Siegel, L. S., Share, D. L., & Geva, E. (1995). Evidence for superior orthographic skills in dyslexics. *Psychological Science, 6*, 250–254.

Siegel, L. S., Share, D. L., & Geva, E. (1995). Evidence for superior orthographic skills in dyslexics. *Psychological Science, 6*, 250–254.

Simpson Baird, A., Palacios, N., & Kibler, A. (2016). The cognate and false cognate knowledge of young emergent bilinguals. *Language Learning, 66*(2), 448-470. doi:10.1111/lang.12160

Sorace, A. (2005). Selective optionality in language development. In L. Cornips & K. P. Corrigan (Eds.), *Syntax and variation. Reconciling the Biological and the Social* (pp. 55–80). Amsterdam: John Benjamins.

Toyoda, E. (2017). Intercultural knowledge, awareness and skills observed in a foreign language classroom. *Intercultural Education, 27*(6), 505-516. doi:10.1080/14675986.2016.1256600

Van der Leij, A., & Morfidi, E. (2006). Core deficits and variable differences in Dutch poor readers learning English. *Journal of learning disabilities, 39*, 74–90.

Van der Leij, A., Bekebrede, J., & Kotterink, M. (2010). Acquiring reading and vocabulary in Dutch and English: The effect of concurrent instruction. *Reading and Writing, 23*(3-4), 415-434.

Van der Slik, F.W.P. (2010). Acquisition of Dutch as a second language. The explanative power of cognate and genetic linguistic distance measures for 11 west European first languages. *Studies in Second Language Acquisition, 32*, 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. *The Quarterly Journal of Experimental Psychology*, *67*(6), 1176–1190. doi:10.1080/17470218.2013.850521

Van Rossum, M., & Vismans, R. (2006). Trading culture, teaching culture: the role of the Dutch language tutor in the acquisition of intercultural skills at beginners level. *Dutch Crossing, 30*(1), 143-174. doi: 10.1080/03096564.2006.11730876

Verhoeven, L. (1993b). *Lezen met Begrip 1–Schaal Betekenis Relaties [Comprehension—scale relations between meanings]*. Arnhem, The Netherlands: CITO.

Verhoeven, L. (1995). *Drie Minuten Toets Handleiding [Three Minutes Test Manual]*. Arnhem, The Netherlands: CITO.

Verspoor, M., De Bot, K., & Van Rein, E. (2011). English as a foreign language: The role of out-of-school language input. In A. De Houwer & A. Wilton (Eds.), *English in Europe Today: Sociocultural and educational perspectives* (pp. 147–166). Amsterdam & Philadelphia: John Benjamins.

Webb, S. (2008). Receptive and productive vocabulary sizes of L2 learners. *Studies in Second Language Acquisition, 30*, 79–95. doi:10.1017/S0272263108080042

Wells, J. C. (2008). *Longman Pronunciation Dictionary* (Third edition). Pearson Education.

Wolf, D. (2002). "On the importance of CLIL in the Context of the Debate on Plurilingual Education in the European Union." In D. Marsh (ed.). *CLIL/EMILE – The European Dimension: Actions, Trends and Foresight Potential* (Public Services Contract 2001-3406/001-001). University of Jyväskylä, Jyväskylä, pp. 44-8.

Yarkoni, T., Balota, D., & Yap, M. (2008). Moving beyond Coltheart's N: A new measure of orthographic similarity. *Psychonomic Bulletin & Review*, *15*(5), pp. 971-979. doi:10.3758/PBR.15.5.971

Zink, I., & Lejaegere M. (2002). *Lijsten voor Communicatieve Ontwikkeling*. In Aanpassing en hernormering van de MacArthur CDIs van Fenson et al. Acco: Leuven.

II. Appendix

ITEM	ENGLISH	DUTCH	OVERALL P-VALUE
1	Ball	Bal	1
2	Dog	Hond	1
3	Spoon	Lepel	0,96
4	Foot	Voet	1
5	Duck	Eend	0,98
6	Banana	Banaan	1
7	Shoe	Schoen	1
8	Cup	Kopje	0,99
9	Eat	Eet	1
10	Bus	Bus	1
11	Flower	Bloem	1
12	Mouth	Mond	0,99
13	Pencil	Potlood	0,99
14	Cookie	Koekje	1
15	Drum	Trommel	1
16	Turtle	Schildpad	0,98
17	Red	Rood	1
18	Jump	Spring	0,98
19	Carrot	Wortel	0,96
20	Read	Lees	0,99
21	Toe	Teen	0,94
22	Belt	Riem	0,89
23	Fly	Vlieg	0,99
24	Paint	Schilder	1
25	Dance	Dans	1
26	Whistle	Fluitje	0,91
27	Kick	Schop	0,98
28	Lamp	Lamp	1
29	Square	Vierkant	0,97
30	Fence	Hek	0,77
31	Empty	Leeg	0,97
32	Happy	Blij	0,99
33	Fire	Vuur	1
34	Castle	Kasteel	1
35	Squirrel	Eekhoorn	0,85
36	Throw	Gooi	0,99
37	Farm	Boerderij	0,99
38	Penguin	Pinguin	0,99
39	Gift	Cadeau	1
40	Feather	Veer	0,95
41	Cobweb	Spinnenweb	0,85
42	Elbow	Elleboog	0,99
43	Juggle	Jongleer	0,82
44	Fountain	Fontein	0,97
45	Net	Net	0,99
46	Shoulder	Schouder	1
47	Dress	Aankleed	0,95
48	Roof	Dak	0,96
49	Peek	Gluur	0,65
50	Ruler	Liniaal	0,94
51	Tunnel	Tunnel	0,99
52	Branch	Tak	0,62
53	Envelope	Envelop	0,99
54	Diamond	Diamant	0,69
55	Calendar	Kalender	0,98
56	Buckle	Gesp	0,26
57	Saw	Zaag	0,88
58	Panda	Panda	0,99
59	Vest	Gilet	0,3
60	Arrow	Pijl	0,89
61	Pick	Pluk	0,98
62	Target	Doelwit	0,9
63	Drip	Druppel	0,97
64	Knight	Ridder	0,95
65	Deliver	Bezorg	0,95
66	Cactus	Cactus	1
67	Dentist	Tandarts	0,97
68	Float	Drijf	0,87
69	Claw	Klauw	0,97
70	Uniform	Uniform	0,97
71	Gigantic	Gigantisch	0,9
72	Furry	Harig	0,53
73	Violin	Viool	0,94
74	Group	Groep	0,98
75	Globe	Aardbol	0,91
76	Vehicle	Voertuig	0,68
77	Chef	Kok	1
78	Squash	Pompoen	0,59
79	Ax	Bijl	0,72
80	Flamingo	Flamingo	1
81	Chimney	Schoorsteen	0,68
82	Sort	Sorteer	1
83	Waist	Taille	0,77
84	Vegetable	Groente	0,88
85	Hyena	Hyena	1

86	Plumber	Loodgieter	0,48
87	River	Rivier	0,97
88	Timer	Timer	0,89
89	Catch	Vang	0,99
90	Trunk	Stam	0,82
91	Vase	Vaas	0,96
92	Harp	Harp	0,97
93	Bloom	Bloei	0,81
94	Horrified	Geschokt	0,78
95	Swamp	Moeras	0,89
96	Heart	Hart	0,97
97	Pigeon	Duif	0,84
98	Ankle	Enkel	0,99
99	Flame	Vlam	0,76
100	Wrench	Moersleutel	0,42
101	Aquarium	Aquarium	1
102	Refuel	Bijtank	0,83
103	Safe	Kluis	0,97
104	Boulder	Rotsblok	0,49
105	Reptile	Reptiel	0,75
106	Canoe	Kano	0,97
107	Athlete	Atleet	0,88
108	Tow	Wegsleep	0,66
109	Luggage	Bagage	0,65
110	Direct	Leid	0,98
111	Vine	Wijnstok	0,69
112	Digital	Digitaal	0,92
113	Dissect	Ontleed	0,81
114	Predatory	Roofzuchtig	0,78
115	Hydrant	Brandkraan	0,2
116	Surprised	Verrast	0,95
117	Palm	Palmboom	0,98
118	Clarinet	Klarinet	0,79
119	Valley	Vallei	0,85
120	Kiwi	Kiwi	1
121	Interview	Interview	0,92
122	Pastry	Gebak	0,73
123	Assist	Help	0,91
124	Fragile	Fragiel	0,41
125	Solo	Solo	0,93
126	Snarl	Grom	0,67
127	Puzzled	Verbaasd	0,87
128	Beverage	Dranken	0,23
129	Inflated	Opgepompt	0,41

130	Tusk	Slagtand	0,12
131	Trumpet	Trompet	0,74
132	Rodent	Knaagdier	0,14
133	Inhale	Inhaleren	0,74
134	Links	Schakels	0,62
135	Pollute	Vervuil	0,48
136	Archeologist	Archeoloog	0,87
137	Coast	Kust	0,81
138	Inject	Injecteer	0,9
139	Fern	Varen	0,79
140	Mammal	Zoogdier	0,62
141	Demolish	Sloop	0,72
142	Isolation	Afzondering	0,48
143	Clamp	Klem	0,81
144	Dilapidated	Bouwvallig	0,79
145	Pedestrian	Voetganger	0,49
146	Interior	Interieur	0,71
147	Garment	Kledingstuk	0,31
148	Depart	Vertrek	0,67
149	Feline	Katachtig	0,31
150	Hedge	Heg	0,54
151	Citrus	Citrusvrucht	0,88
152	Florist	Bloemist	0,87
153	Hover	Zweef	0,52
154	Aquatic	Aquatisch	0,5
155	Reprimand	Berisp	0,55
156	Carpenter	Timmerman	0,45
157	Primate	Primaat	0,26
158	Glider	Zweefvliegtuig	0,53
159	Weary	Vermoeid	0,65
160	Hatchet	Handbijl	0,28
161	Transparent	Doorzichtig	0,72
162	Sedan	Sedan	0,32
163	Constrained	Belemmerd	0,55
164	Valve	Afsluiter	0,33
165	Parallelogram	Parallellogram	0,51
166	Pillar	Pilaar	0,83
167	Consume	Consumeer	0,54
168	Currency	Valuta	0,43
169	Hazardous	Gevaarlijk	0,17
170	Pentagon	Pentagon	0,34
171	Appliance	Apparaat	0,27
172	Poultry	Gevogelte	0,07
173	Cornea	Hoornvlies	0,26

174	Peninsula	Schiereiland	0,33
175	Porcelain	Porselein	0,73
176	Detonation	Ontploffing	0,47
177	Cerebral	Cerebraal	0,24
178	Perpendicular	Loodrecht	0,39
179	Submerge	Onderdompel	0,07
180	Syringe	Injectiespuit	0,35
181	Lever	Hefboom	0,32
182	Apparel	Kleding	0,09
183	Talon	Klauw	0,04
184	Cultivate	Cultiveer	0,29
185	Wedge	Wig	0,29
186	Ascend	Omhooggaan	0,28
187	Depleted	Leeggehaald	0,3
188	Sternum	Borstbeen	0,11
189	Maritime	Maritiem	0,34
190	Incarcerate	Opsluit	0,12
191	Dejected	Teneergeslagen	0,26
192	Quintet	Kwintet	0,38
193	Incandescent	Gloeiend	0,03
194	Confide	Toevertrouw	0,11
195	Mercantile	Commercie	0,1
196	Upholstery	Stoffering	0,02
197	Filtration	Filtratie	0,21
198	Replenish	Aanvul	0,07
199	Trajectory	Traject	0,1
200	Perus	Doorlees	0,08
201	Barb	Weerhaak	0,08

202	Converge	Convergentie	0,16
203	Hone	Slijp	0,16
204	Angler	Hengelaar	0,08
205	Wildebeest	Gnoe	0,02
206	Coniferous	Conifeer	0,05
207	Timpani	Pauk	0,02
208	Pilfer	Snaai	0,02
209	Pestle	Vijzel	0,05
210	Repose	Uitrust	0,02
211	Cupola	Koepel	0,04
212	Derrick	Boortoren	0,02
213	Convex	Convex	0,03
214	Embossed	Relief	0
215	Torrent	Stortvloed	0,02
216	Dromedary	Dromedaris	0,05
217	Legume	Peulvrucht	0,01
218	Cairn	No translation	0,01
219	Arable	Bebouwbaar	0,03
220	Supine	Achteroverliggend	0
221	Vitreous	Glasachtig	0,01
222	Lugubrious	Naargeestig	0,01
223	Caster	Wieltje	0
224	Terpsichorean	Terpischore	0
225	Cenotaph	Gedenkteken	0
226	Calyx	Bloemkelk	0,02
227	Osculate	Verwant	0,01
228	Tonsorial	Haarsnijders	0