The role of crosslinguistic influence in Greek child and adult learners' acquisition of English as a Foreign Language

Faidra Faitaki, DPhil Candidate
Department of Education and St Edmund Hall
University of Oxford

Supervisors: Professor Victoria Murphy and Professor Kate Nation

Thesis submitted in part-fulfilment of the requirements for the degree of Doctor of Philosophy in Education

Στον παππού μου

Acknowledgements

Enrolling on the DPhil programme in September 2017, I did not know what to expect. In August 2021, with the programme coming to an end, I can say with certainty that the past four years have been an academic and emotional rollercoaster. On an academic level, most things did not go as planned: the topic of the DPhil had to be modified, its sample had to be diversified, its methods had to be adapted and, as a result, the end-product looks nothing like I thought it would at the beginning. On a personal level, the past four years were turbulent – not least because of the end of a long relationship, the death of two grandparents and three waves of a devastating pandemic. Yet, despite the academic and emotional bumps (or indeed, because of them), completing this degree has been the biggest achievement of my life.

Achievements this size take a village to accomplish, and I am grateful to each and every one of the people who contributed to this project. The list is long but, first and foremost, I would like to thank my supervisors: Professor Victoria Murphy and Professor Kate Nation. From the application to the submission stage, Vicki was my constant source of support: she talked me down before I made a flawed decision and talked me up whenever I lost faith in myself. She is living proof of the fact that an outstanding academic can also be a caring advisor, and I aspire to follow in her footsteps throughout my professional life. Being an expert in a different field, Kate always considered my work from a perspective that I had not (and perhaps would have never) considered. Her suggestions were critical as well as inspirational, and this DPhil thesis would not be the same without her input.

In addition to my supervisors, I would like to thank Professor Dorothy Bishop, Professor Ianthi Tsimpli, Dr Elizabeth Wonnacott and Dr Robert Woore, who examined the thesis at its various milestones and, through their comments, helped me develop my thinking, refine my writing, and improve the overall quality of my research. I would like to extend my thanks to the members of the R.E.A.L. group and the Applied Linguistics group in the

Department of Education for the fruitful discussions that allowed me to flesh out rough ideas and build on more concrete ones. Most of all, I would like to thank (Dr!) Henriette Arndt, (Dr!) Annina Hessel, (Dr!) Kari Sahan, and (soon-to-be Dr!) Kedi Simpson, who offered great advice on pressing academic problems, and much-needed (caffeine-based) distraction during some of the most challenging times I experienced over the years.

Yet, this DPhil project would still (and perhaps forever) be in progress without the assistance of the administrators and teachers of: I. M. Panagiotopoulos School, St Catherine's British School, Merimna Paidiou Nursery and Preschool, Bradmore Road Nursery and Thomas' Clapham. From making me toasties and driving me home, to sharing vital information about the students and letting me take over the staffrooms with my testing equipment, the school staff made my research visits both easier and more pleasant. In particular, I want to express my gratitude to Ms Yvonne Mengos, Ms Eva Kaliskami and Ms Julie Tavoulari, who went above and beyond to help me collect the data I needed, even amidst a lockdown. Moreover, and although the chances of them ever reading this thesis are slim, I thank all the children who took part in the project. It was because of them –their exceptional linguistic abilities, their spontaneous hugs, and their unintentional jokes– that data collection was hands-down the best part of the entire DPhil process.

Moving on, I would like to mention all the people who played a role in the project behind the scenes. Living in Oxford would not have been nearly as nice had Priyanka Bawa, Tegan Ekanayake, Claudia Hartman and Carey Pike not helped me fill the city with vibrant colours and beautiful memories. I am also grateful to all of my beloved school friends who have been standing by me despite the geographical distance; though I will not list all of them (if I do, these acknowledgements will go on forever!), I need to mention Konstantinos Chatzivaritis, for being an incomparable source of support and laughter (wine and sour cherry ice-cream!) and my best friend for over 15 years. Moreover, I would be amiss not to thank my

brilliant TheatreLab famiglia for the enticing explorations and magical moments we experienced together – both on and off the stage. A special shout-out goes to Irini Fasia and Christiana Mitrollari for being the most supportive of bubbles at the most frustrating of times. Finally, I would like to thank Alex Mahal –the latest but most valuable addition to the support team– for enduring my tantrums, indulging my whims, and encouraging me to keep going – making the last six months before submission as amazing as they could be!

Last, but not least, I would like to acknowledge my family's immense contribution to the thesis – its start and its completion. It was their unshaking belief in the value of education that convinced me to pursue a DPhil at a time when I was not sure that academia was right for me. Throughout the course of the programme, my (grand)parents were present from afar: they listened to my problems and suggested the best solutions; they understood my worries and found the most suitable remedies; they appreciated my emotions and gave me space to express them. I thank them from the bottom of my heart for equipping me with the tools I needed to complete this journey, and for helping me repair the tools whenever they broke. This thesis is (also) dedicated to them, with all my love.

Table of Contents

Table of Contents

Thesis Abstract

Crosslinguistic influence (CLI) is the influence that a bilingual's first language (L1) exerts on their second language (L2) and vice versa. Most research on the causes and effects of the phenomenon focuses on simultaneous bilingual children and syntax-pragmatics interface linguistic structures that are deemed susceptible to CLI. As such, it does not determine whether the phenomenon affects the development of L2 learners and/or the use of linguistic structures that are not susceptible to CLI. To this end, this doctoral project aimed to uncover the role of CLI in Greek children and adults' use of CLI-prone and not CLI-prone structures in English.

In Experiment 1, 140 Greek-English sequential bilingual preschool children (35 attending a total immersion programme programme and 35 attending a partial immersion programme, as well as 34 Greek monolingual and 36 English monolingual controls) were tested on their use of sentential subjects (a syntax-discourse interface structure that is susceptible to CLI), once in Greek and once in English. The bilingual children were found to omit sentential subjects in English more often than their English monolingual peers. This finding could be attributed to CLI from the children's dominant L1 to their non-dominant L2.

Experiment 2 was conducted parallel to Experiment 1, and aimed to ascertain if the subject omissions observed in Experiment 1 were indeed due to CLI. The same children were tested on their use of the subordinate clause marker 'to' (a syntactic structure that is not susceptible to CLI and, thus, could serve as a control) in the two languages. The analyses revealed that bilingual children omitted more clause markers in English than their English monolingual peers. This finding hints that bilingual children are attentive to the (cross-linguistic) distributional properties of the structures they are exposed to.

Hence, Experiment 3 was designed to investigate whether the subject omissions observed in Experiment 1 also appeared at a more advanced state of L2 acquisition. To this end, a group of 31 Greek adult L2 learners were tested on their use of referential expressions,

and their performance was compared to that of 30 English L1 speakers. Like their child counterparts, the adult L2 learners were also found to omit more sentential subjects from their utterances than the L1 controls – an effect attributed to CLI from the learners' L1, Greek, to their L2, English.

Considered together, the results of these three experiments highlight that sequential bilingual children pattern with adult L2 learners in terms of their production (and CLI patterns therein). Moreover, they do not support traditional accounts of CLI which predict the phenomenon to arise in the case of structural overlap and/or just for interface linguistic structures. To provide a uniform account of omissions in L2 acquisition, the effects observed across the three experiments could be explained under a usage-based account that considers the structures' distribution in the L1 (and L2) input as well as the attentional mechanisms that sequential bilingual and L2 learners have developed through their exposure to the L1.

Word Count: 491

COVID-19 Statement

According to the initial plan, this doctoral project would have involved two studies, focusing on Greek-English sequential bilingual children. The first study would have consisted of two experiments (one testing the use of sentential subjects and another testing the use of subordinate clause markers) and involved Greek children who lived in England and learned English as an Additional Language (EAL). The second study would have constituted a partial replication of the first, such that the same two experiments would have been conducted with Greek children who lived in Greece and learned English as a Foreign Language (EFL) as a result of attending total or partial immersion preschool programmes. A possible third study, expanding on the findings of the first two, was envisioned (albeit not designed in detail). This third study would have involved exploring the usage-based framework that is sketched out at the end of the thesis.

Yet, the breakout of COVID-19 prevented aspects of this plan from being implemented. At the time of the first lockdown, half of the EAL children that formed part of the first study's sample had only completed the experiments in one of their languages. As testing was taking place in the children's homes, it was impossible to visit them in order to complete the experiments in their other language; this rendered the data collected in the first round of testing unusable. Likewise, the EFL children that formed part of the second study's sample had also only completed the experiments once. However, this time, it was possible to complete the second round of testing online; this allowed salvaging the data that had been collected in the first round of testing, and made this study the focus of the thesis. Furthermore, a third experiment (investigating the role of CLI in Greek adult learners' acquisition of EFL) was designed. Although this experiment was different from the third study that was originally envisioned, it is nonetheless a valuable addition to the project as it extends the results and conclusions that the child EFL learner experiments gave rise to.

Word Count: 340

List of Abbreviations

ANOVA:	Analysis of Variance
AoO:	Age of Onset
BPVS:	British Picture Vocabulary Scales
CHILDES:	Child Language Data Exchange System
CLI:	Crosslinguistic Influence
CUREC:	Central University Research Ethics Committee
DVIQ:	Diagnostic Verbal Intelligence Quotient
EFL:	English as a Foreign Language
L1:	First Language
L2:	Second Language
MLU:	Mean Length of Utterance
PPVT:	Peabody Picture Vocabulary Test
SES:	Socioeconomic Status
TROG:	Test for the Reception of Grammar

List of Figures

List of Tables

Chapter 1: Introduction

Bilingualism research took off in the 1980s as researchers sought to provide an answer to the (withstanding) question of whether bilingual children's two languages develop as a unified linguistic system or as two separate linguistic systems (Werker & Byers-Heinlein, 2008). The first studies on dual language acquisition suggested that, at the start of their lives, bilingual children learn their two languages without discriminating between them (Redlinger & Park, 1980; Vihman, 1985; Volterra & Taeschner, 1978). According to Volterra and Taeschner (1978), the process of dual language acquisition involves three stages: in the first stage, children possess a single system for both languages; in the second stage, they begin to differentiate between the languages on a lexical but not a grammatical level; in the third stage children also start to differentiate between the two languages' grammars, as well as to dissociate each language from the parent who speaks it.

This account, which is now referred to as the Unitary System Hypothesis, derived its support from studies that investigated bilingual children's code-mixing (i.e., the use of different languages in the same utterance). Sceptical about the use of code-mixing as evidence in favour of the Unitary System Hypothesis, Genesee (1989) examined the previous studies' design and findings in detail. He argued that the code-mixing patterns reported in the previous studies were characterised by linguistic sophistication and contextual sensitivity that presupposed the existence of two separate linguistic systems. Thus, he put forth the Dual System Hypothesis, which suggests that bilingual children separate the mental representations of their two languages from the start of their lives.

More recent work, using methodologies that tap into online processing, has elucidated the process of language differentiation further, and lent additional support to the Dual System Hypothesis (Byers-Heinlein, 2014). It is now known that bilingual infants can discriminate between the intonation patterns (Byers-Heinlein, Burns & Werker, 2010), phonological

information (Sebastián-Gallés & Bosch, 2002), and visual cues (Sebastián-Gallés, Albareda-Castellot, Weikum & Werker, 2012) of two different languages from the first months of life. Being able to discriminate between two languages' elements does not imply that infants place these elements into distinct mental categories, but it does suggest that they are sensitive to observable differences between languages from the start of development (Byers-Heinlein, 2014). This observation is corroborated by studies on older bilinguals which show that they can selectively produce the sounds (Poulin-Dubois & Goodz, 2001), words (Genesee & Nicoladis, 2007), and grammatical structures (Meisel, 2001) of both their languages as appropriate in the discourse context. These results hint that bilingual children represent phonological, lexical and grammatical information of their language in separate systems.

As language differentiation, and thus the Dual System Hypothesis, became established in the field, researchers turned their attention to the degree of contact that exists between bilingual children's two mental linguistic systems (Serratrice, 2013). This contact is referred to as crosslinguistic influence (CLI) and can be defined as the influence that the first language (L1) exerts on the second language (L2) and vice versa. Like code-switching, CLI might result in mixing the elements of two languages but, unlike code-switching, CLI is a subconscious process that does not involve speaking in a different language. As the theoretical backdrop of CLI is the idea that bilinguals' linguistic representations are separate but systematically interacting, CLI should also be systematic – otherwise, the separateness of linguistic representations that the Dual System Hypothesis posits will be called into question.

The quest for the systematic conditions that give rise to the interaction between two languages has led to various conceptual and empirical investigations. Yet, to this day, there is little consensus regarding the criteria that underpin the phenomenon of CLI. Moreover, most of the investigations face theoretical and methodological issues, including a focus on specific linguistic structures and learner populations. The aim of this doctoral project is to engage with

these issues, and to explore the role of CLI in the acquisition of a (growing) population that merits more attention: Greek children who learn English at preschool. These children can be classified as sequential bilingual learners, as their L2 acquisition starts sometime after their L1 acquisition. Despite the fact that the population of sequential bilingual children is growing across the globe due to educational reforms, research on their linguistic development (and CLI patterns therein) is not as extensive as the research on simultaneous bilingual children, who learn their two languages at the same time, or older L2 learners.

In this chapter, I will provide a thorough introduction to the project. First, I will operationalise CLI, discuss two (sets of) accounts that have been proposed to explain the presence and direction of the phenomenon, and highlight the limitations of existing studies that have explored the role of CLI in child and adult language development. The literature review will target simultaneous bilingual children L2 learning adults; though these groups differ from the sequential bilingual learners that are the focus of thesis in various respects (e.g., the starting age of L2 acquisition and the amount of L2 exposure), the comparison between them and sequential bilingual learners is believed to offer a starting point that both situates and facilitates the research conducted as part of the thesis. Following the literature review, I will describe the (changing) Greek educational landscape, in which the DPhil's experiments are embedded. Hence, I will showcase the linguistic structures that I intend to investigate in the thesis, and compare their use in English and Greek. The final section of the chapter contains an overview of the three experiments that the thesis consists of as well as an overall outline.

1.1. Theoretical Background

After three decades of research, there is widespread agreement that bilingual children's languages are distinct but interacting (however, for a counterargument, see: García & Wei, 2014). Nevertheless, there is still some disagreement regarding the precise conditions under which CLI appears. To understand these conditions, it is essential to establish what CLI is and

to determine what form(s) it can take. Paradis and Genesee (1996) were the first to provide an operationalisation of CLI. In their seminal paper, which aimed to determine if the development of bilingual children's two languages is autonomous or interdependent, the authors defined CLI as "the systemic influence of the grammar of one language on the grammar of the other language during acquisition" (Paradis & Genesee, 1996: 3). They explained that interdependence can manifest itself as transfer or acceleration. Transfer occurs when the two languages are structurally different; then, bilingual children transfer the grammatical properties of one language to the other language, thus producing illicit sentences in the latter. Acceleration occurs when the two languages are structurally similar; then, bilingual children's knowledge of a grammatical structure in one language facilitates the acquisition of the same structure in the other language. Paradis and Genesee (1996) argued that interdependence might also delay bilingual children's linguistic development. Yet, the overall delay is not related to the relative structure of the languages that bilingual children are learning but, rather, represents an effect of bilingualism as a whole.

To explore the possible effects of interdependence, Paradis and Genesee (1996) focused on the linguistic development of three English-French simultaneous bilingual children in Canada. They targeted three morphosyntactic structures which differ in English and French: finiteness, negation and the distribution of pronominal subjects. The three bilingual children were filmed interacting with their English-speaking father, French-speaking mother and both parents together at the ages of 2;0, 2;6 and 3;0. The authors reported that children acquired the three structures on par with monolinguals, and argued that children's knowledge of grammatical properties in French does not lead to transfer or accelerate their acquisition of the same grammatical properties in English. In other words, their results did not support the idea that the simultaneous bilingual children's development is interdependent.

1.1.1. Explaining Crosslinguistic Influence

While Paradis and Genesee's (1996) investigation did not provide direct evidence in favour of linguistic interdependence, CLI effects have been widely reported in the literature since. Most of the research on the topic suggests that the presence of CLI is determined by structural linguistic factors, that is, the relative structure of the two languages that bilinguals are learning (e.g., Hulk & Müller, 2000; Sorace & Filiaci, 2006). An alternative proposal is that the appearance of CLI is related to the quantity of input that bilinguals receive in either language (e.g., Argyri & Sorace, 2007; Yip & Matthews, 2000). Meanwhile, there is a growing belief that qualitative characteristics of the input contribute to the presence and magnitude of CLI effects (e.g., Daskalaki, Blom, Chondrogianni & Paradis, 2020; see also: Paradis & Navarro, 2003). The latter explanation has not received as much attention in the literature to date and will therefore be excluded from the discussion. Nevertheless, investigating the effects of input quality seems to be a promising endeavour that can provide a fresh perspective into the conditions that give rise to CLI[1].

1.1.1.1. Linguistic Structure

Hulk and Müller (2000) were the first to support the idea that bilinguals' two languages can both develop autonomously and influence each other. The authors argued that CLI arises "if (i) one of the languages opens the possibility for two grammatical analyses of a particular construction, and (ii) the other language seems (to the child) to support one of these two possible analyses" (Hulk & Müller, 2000: 228). So, if bilingual children's Language A allows

[1] It should be noted that, sometimes, the term 'CLI' is replaced with the term 'transfer'. I consider the two terms to differ in terms of their operationalisation, following Paradis and Genesee (1996) whose account (discussed in the previous section) argued transfer to be one of the possible manifestations of CLI. In addition, the two terms differ in terms of their meaning. Indeed, the connotations that 'transfer' carries are negative: the term implies that the interaction between two languages is one-directional, from the L1 to the L2, and rigid. By contrast, 'CLI' implies that the interaction of the two languages is two-directional (from the L1 to the L2 or from the L2 to the L1) and fluid. In light of these observations, I will use the term CLI throughout the thesis in order to reflect that the uncertainty around the target population's CLI patterns (given the fact that sequential bilingual children are acquiring their L2 while their L1 knowledge is still developing).

two manifestations of the same structure (and thus forms a 'superset'), and Language B allows only one of them (which forms the 'subset'), children are expected to use the overlapping structure across the board, even where Language A would require the non-overlapping structure. The authors highlighted that, for CLI to appear, another condition should be met: the target subset-superset structure should be regulated in the C-domain, which corresponds to the mental 'interface' between the domain of syntax and the domain of pragmatics.

To test their account, Hulk and Müller (2000) recorded two simultaneous bilingual children, one French-Dutch and one Italian-German, in interactions with their parents. Their aim was to investigate two linguistic structures that are constrained by the syntax-pragmatics interface: object-drop and root infinitives. Germanic languages (like German and Dutch) allow object-drop in clause-initial position, but Romance languages (like Italian and French) seem, at least to children, to allow it in varied contexts. By contrast, both Germanic and Romance languages allow root infinitives in the same contexts. Thus, the C-domain condition is met by both linguistic structures, while the subset-superset condition is only met by object-drop. For this reason, Hulk and Müller (2000) predicted CLI from the Germanic to the Romance language only in the case of object-drop. Their findings confirmed their predictions.

Formalising and extending Hulk and Müller's (2000) observation regarding the importance of the C-domain, Sorace and collaborators (Sorace, 2006; Sorace & Filiaci, 2006; Tsimpli & Sorace 2006) proposed the Interface Hypothesis. The main tenet of the Interface Hypothesis was that "language structures involving an interface between syntax and other cognitive domains are less likely to be acquired completely than structures that do not involve this interface" (Sorace, 2011: 1). Tsimpli and Sorace (2006) further divided interfaces into grammar-internal and grammar-external (also see Sorace & Serratrice, 2009); while the former involve linguistic phenomena that fall within the realm of semantics, the latter regulate phenomena that pertain to the domain of pragmatics. Hence, they argued that external

interfaces are more difficult to process as they require learners to integrate syntactic and pragmatic information (Sorace & Serratrice, 2009).

Putting their idea to the test, Tsimpli and Sorace (2006) studied the speech of 27 adult L1 Russian L2 Greek learners, focusing on the use of focus and the realisation of pronominal subjects. The rules that underpin the use of these two phenomena differ in Russian and Greek. Moreover, while focus is a syntax-semantics interface phenomenon, subject pronouns are regulated in the syntax-pragmatics interface. For this reason, the authors predicted that the learners would have more trouble with subject pronouns than with focussed constituents. Indeed, the L2 Greek learners behaved more L1 Greek speakers with regards to focus than with regards to pronouns, confirming the authors' predictions. Furthermore, Tsimpli and Sorace (2006) explained that learners' performance appeared to be independent of their length of residence in Greece (and, thus, their length of exposure to Greek), suggesting that interface structures remain problematic even for learners with high proficiency. Yet, the authors' claims about the additional difficulties associated with the processing of external interfaces cannot be evaluated based on this study alone, since its offline methods cannot tap into the learners' real-time processing (Sorace, 2011).

This methodological limitation is not exclusive to Tsimpli and Sorace's (2006) study; in fact, most of the work on the IH involves offline methods and concentrates on the bilinguals' knowledge representations of interface structures, rather than on the processing of these structures. The focus on knowledge representations has given rise to 'representational accounts' of the Interface Hypothesis (for the term, see: Sorace, 2011). Representational accounts are founded on the belief that monolingual and bilingual speakers possess different representations of interface structures due to CLI. Then, the language that possesses the more economical representations is argued to influence the language that possesses the less economical representations.

This idea was investigated by Serratrice, Sorace, Filiaci and Baldo (2009). The authors assessed English-Italian (N=59) and Spanish-Italian (N=31) bilingual children's knowledge of the rules that underpin the use of generic expressions (e.g., 'roses are red') as opposed to specific expressions (e.g., 'the rose is red'), using an acceptability judgement task. Following Chierchia's Nominal Mapping Parameter (1998), they explained that English and Italian differ in terms of the features that nominal arguments contain (see Jackendoff, 2002). In English, nouns are marked with the features [+arg +pred] and can denote either predicates or arguments; in Italian, nouns are [-arg +pred] and, thus, require a determiner to function as arguments. Projecting a determiner is not an economical process from a cognitive perspective; so, English-Italian bilingual children are expected to transfer the properties of the more economical language (in this case, English) to the less economical language (Italian). Serratrice et al (2009) reported that their English-Italian bilingual participants were less accurate than the English monolingual and the Spanish–Italian bilingual participants in rejecting bare nouns in generic contexts in Italian – an effect they attributed to the presence of CLI from English to Italian.

Yet, representational accounts are limited in that their explanatory efficacy only extends to situations when the bilingual is learning two languages that differ in terms of their feature specification (Sorace, 2011). This observation led to the proposal of 'processing resources accounts'. These accounts explore differences in monolingual and bilinguals' online processing strategies during the course of interpreting and/or producing interface structures. They maintain that bilingual speakers' variable usage patterns result from their inefficient processing of (syntax-discourse) interface structures. Their limited processing efficiency might stem from less detailed knowledge representations of the structures and/or a less automated access to these knowledge representations (Sorace, 2006).

These explanations were tested in Hopp's (2009) investigation of scrambling. In German, the order in which words appear in a sentence can be altered, such that subjects might

appear after the verb if they are stressed. In scrambled sentences, the role of each constituent (that is, whether it is the subject or the object of the sentence) is signalled through case-marking. Thus, scrambling represents an interface phenomenon, as it involves various linguistic domains. To assess the knowledge of the morphosyntactic and semantic constraints of scrambling, Hopp (2009) tested a total of 29 L1 Russian, 28 L1 English and 34 L1 Dutch L2 German learners in two experiments: one involved an offline acceptability judgement task and the other involved an online reading task. The results suggested that some of the learners had not mastered the phenomenon, and that proficiency played a role in shaping the learners' performance in the two experiments, such that the 'near-native' learners outperformed the 'advanced' ones. Hopp (2009) also noted a divergence between the participants' results in the offline and the online task. To account for this divergence, he argued that the learners possess the knowledge representations required to interpret scrambled constituents, but are not as efficient in integrating different sources of information in real-time.

The generalised processing approach adopted by Hopp (2009) and other proponents of processing resources accounts allows for a uniform explanation of the discrepancies that exist between bilingual and monolingual learners' performance, irrespective of the structure of their language(s). However, even processing resources accounts presuppose the existence of 'interfaces' between linguistic and extra-linguistic domains, and consider the interfaces to be problematic for (even 'near-native') language learners.

The idea of interfaces is linked to Generative Grammar (Chomsky, 1959; 1981; 1995). According to (some of the earlier frameworks of) Generative Grammar, humans' mental linguistic faculty consists of four levels of representation: the Lexicon; the D(eep)-Structure, which combines the lexical items; the S(urface)-Structure, where the D-Structure is mapped; and two interface systems, the Phonological Form and the Logical Form, that allow the end-product of the S-Structure to be interpreted (Chomsky, 1981). Generative Grammar views

pragmatics, the domain which is responsible for the interpretation of utterances in the discourse context, as external to this mental linguistic faculty. The separateness of the domain explains why linguistic processes that are regulated in the syntax-pragmatics interface are considered to be more difficult for learners.

While a lot of the research on CLI has adopted a Generative Grammar approach, few studies have been conducted under its 'opponent': the usage-based approach. The latter constitutes a move away from generative and closer to connectionist approaches (Shirai, 2018), as it maintains that there is a profound association between linguistic structure and use, such that linguistic structure is acquired as a result attaining to various aspects of the input (Bates & MacWhinney, 1989; Bybee, 1985; Diessel, 2019; Goldberg, 2006; Langacker, 1987; Tomasello, 2003). For instance, Bybee (1985) highlighted that, when acquiring their L1, children 'calculate' the frequency with which certain lexical and grammatical items occur. In turn, frequency impacts their acquisition, as frequent items acquired faster and greater ease than infrequent ones (Ambridge, Kidd, Rowland, & Theakston, 2014; Ambridge & Lieven, 2011; Lieven, 2010).

The usage-based approach encompasses multiple frameworks, many of which have been employed to explain the process of L2 acquisition. For instance, associative learning theories (which fall under the usage-based umbrella) have been proposed to account for persisting mistakes in L2 learners' language use (see: Ellis, 2002; 2006; 2019; Ellis, Hafeez, Martin, Chen, Boland & Sagarra, 2014; Ellis & Sagarra, 2010). These theories suggest that L2 learners attain to particular aspects of their L2 input based on their L1 experience. Thus, their acquisition and use of linguistic structures in the L2 will depend on whether these structures are present in the L1: if the structures are present in the L1, the learners are likely to look for and produce them in the L2 as appropriate; if not, the learners are not (as) likely to look for or produce them as their L1 attentional mechanisms will 'block' them. These mechanisms

develop as a result of the reduced plasticity that characterises L2 learning (for a review of simulation studies conducted to explore the idea of prior neural commitment in L2 acquisition under a connectionist framework, see: Shirai, 2018). Despite the clear synergy between the predictions of associative learning theories and the appearance of transfer effects in L2 learners' speech, usage-based frameworks have not been explicitly linked with CLI to date. Nevertheless, these frameworks will be mentioned in various chapters of the thesis, and further discussed in Chapter 5.

1.1.1.2. Linguistic 'Dominance'

Although linguistic structure has been widely reported as the main source of CLI, it is not the sole cause: numerous studies have highlighted the importance of linguistic dominance in determining the presence, magnitude and/or direction of CLI. These studies suggest that CLI appears due to language-internal reasons, but individual differences are bound to impact the acquisition process.

One such difference is the children's linguistic dominance. The important role of linguistic dominance or input quantity in shaping CLI effects has been noted since the beginning of bilingualism research. For example, Yip and Matthews (2000) investigated the speech of their Cantonese-English bilingual son between the ages of 1;5 and 3;6 and reported unidirectional transfer from Cantonese to English in various phenomena, such as null objects, in-situ wh-questions and prenominal relative clauses. The authors used Mean Length of Utterance (MLU; see Brown, 1973) as a proxy for dominance and reported that, given his higher MLU in Cantonese, their son was Cantonese-dominant; thus, language dominance was considered responsible for his observed errors.

However, not all researchers adopt the same definition (and measurement) of dominance. For instance, Genesee, Nicoladis and Paradis (1995) defined dominance as 'language preference' and measured it as the total amount of time for which children used the

language. However, language use is not believed to be a reliable estimate of dominance, as some children might refrain from using a language but still be dominant in it (De Cat, 2020). Defining dominance as morphological complexity and measuring it through MLU scores (e.g., Yip & Matthews, 2000) is problematic since the presence and use of morphemes is subject to cross-linguistic variation (Marinis, 2003). Likewise, defining dominance as 'proficiency' and measuring it through children's performance in vocabulary tests (e.g., Nicoladis, 2012) might under- or over-represent their knowledge and lead to discrepancies across the two languages, since even translations of the same tests do not always measure the same constructs (Antonijevic, Durham & Ní Chonghaile, 2017). A definition of dominance as 'input quantity' (e.g., Argyri & Sorace, 2007; Daskalaki, Chondrogianni, Blom, Argyri & Paradis, 2019; Hervé, Serratrice & Corley, 2016) might be more reliable, as using children's relative exposure to the two languages to calculate dominance can limit the effect of extraneous factors and production demands that previous definitions of the term involve (Unsworth, 2013a; 2015). For this reason, the (more descriptive) term 'input quantity' will be used throughout the thesis.

Since Yip and Matthews' (2000) work, linguistic dominance, or input quantity, has been widely reported as a determinant of CLI effects (e.g., Argyri & Sorace, 2007; Daskalaki, et al., 2019; Hervé, et al., 2016; Kupisch, 2012; Nicoladis, 2012). In general, input quantity accounts of CLI purport that the language in which bilingual children are dominant exerts influence on their other language, provided that the two languages use a linguistic structure in a different fashion. In other words, while the appearance of CLI is not limited to a set of specific (ambiguous and/or interface) linguistic structures, a certain degree of cross-linguistic asymmetry is still required for CLI to arise. However, it should be highlighted that the discrepancies in the operationalisation of dominance across different studies question the generalisability of their findings across different contexts (Kupisch, 2012).

1.1.2. Investigating Crosslinguistic Influence

Linguistic structure and input quantity are the most prominent explanations of CLI that have been proposed in the literature. In addition to the issues associated with each of the accounts that were discussed earlier, there are some overarching problems with the investigation of CLI. In particular, studies on the phenomenon tend to focus on specific linguistic structures and learner populations that are known to be impacted by CLI.

1.1.2.1. Lessons from Interface Structures

The focus on specific linguistic structures, such as those regulated in the interface between syntax and pragmatics, has provided valuable information thus far. A number of studies have looked into the acquisition of syntax-discourse interface structures over the years. Most of these appear to focus on the realisation of sentential subjects in various linguistic combinations (e.g., for Greek: Argyri & Sorace, 2007; for Hebrew: Hacohen & Schaeffer, 2007; for Italian: Serratrice, Sorace & Paoli, 2004; for German: Schmitz & Müller, 2008; for Spanish: Paradis & Navarro, 2003). These studies use different methods, but give rise to similar findings.

Longitudinal designs were especially popular at the earlier stages of CLI research (Serratrice, et al., 2004; Pinto, 2007; Paradis & Navarro, 2003; Schmitz & Müller, 2008; Schmitz, Patuto & Müller, 2012). For instance, Serratrice et al (2004) studied the speech of one English-Italian bilingual child, Carlo, between the ages of 1;1 and 4;8. Their aim was to track Carlo's development of sentential subjects in the two languages. To this end, the authors compared Carlo's speech to that of six Italian and five English monolingual children, found in various corpora within the CHILDES database (MacWhinney, 2000). Their results suggested that Carlo produced more overt subjects in Italian than the Italian monolingual children.

To account for these results (which are in line with Hulk and Müller's (2000) Subset-Superset Condition), the authors explained that the child notices that, in Italian, overt subjects are used for specific discourse purposes (e.g., the shift to a new topic), while in English they

have a discourse purpose and a syntactic purpose (and are, thus, 'underspecified' for syntax-discourse interface functions). The authors argue that, due to his contact with English, Carlo experienced the under-specification of Italian overt subjects, which led him to use overt subjects in contexts where Italian would use a null subject. The same pattern of results has been obtained in all the longitudinal and/or corpus-based studies of bilinguals who acquire a language that allows null subjects (like Italian, Spanish and Greek) together with a language that does not (like English).

This finding has been corroborated using experimental designs that tap into children's comprehension of sentential subjects (e.g., Serratrice, 2007; Serratrice, Sorace, Filiaci & Baldo, 2012; Sorace, Serratrice, Filiaci & Baldo, 2009). Sorace et al (2009) asked the same participants who took part in Serratrice et al,'s (2009) experiment discussed above to complete an acceptability judgement task that explored their sensitivity to incorrect and/or inappropriate uses of sentential subjects in Italian and English. Their results suggested that the English-Italian bilinguals were more willing to tolerate the use of overt subjects in non-topic shift contexts (where null subjects would be preferred) than the Italian monolinguals.

Fewer studies appear to experimentally assess children's production of sentential subjects. For example, Argyri and Sorace (2007) tested 32 Greek-English bilingual children, half of whom were Greek-dominant and half of whom were English-dominant, in an oral elicitation task. The authors investigated if children's use of null and overt subjects in preverbal and postverbal position and wh-embedded interrogatives in Greek was influenced by their English. The authors reported that there was some influence of English on Greek in both structures, and that it was English-dominant bilinguals who transferred the properties of English on Greek, while transfer was not apparent in the Greek-dominant group. Thus, the authors' findings highlighted the important role of input quantity (Sorace & Serratrice, 2009).

Research on the acquisition of sentential subjects has also been conducted with adult learners. In one of the first explorations on the topic, White (1985) focused on L1 Spanish and L1 Italian L2 learners of English, divided in three proficiency levels: beginner, intermediate and advanced. She tested all learners using an acceptability judgement task, and reported that the beginner learners accepted 64.7% of the structures that involved illicit null subjects in English, while the intermediate and advanced group performed better but still not close to the English monolingual control group.

More recently, Prenza's (2014) study of 90 beginner, intermediate and advanced Greek learners of L2 English gave rise to similar results. The author reported that, in general, the L2 learners accepted ungrammatical null subjects more often relative to an English monolingual control group. In particular, it was the beginner and intermediate learners who accepted and produced null subjects at chance level. The advanced learners did significantly better on both the productive measure (a cloze test) and the receptive measure (a paced grammaticality judgement task). Yet, they were still outperformed by the English monolinguals. Similar findings were reported by Prenza and Tsimpli's (2013) study, which built on Tsimpli and Roussou's (1991) earlier work.

In sum, the results of the aforementioned studies reveal children who learn a null-subject language together with a non-null-subject language appear to accept and produce more overt subjects in the former than pragmatically appropriate. By contrast, adult speakers of a null-subject L1 who learn a non-null-subject L2 accept and produce more null subjects in the latter than grammatically appropriate. Based on their findings, all of these studies provide evidence in favour of structural linguistic and input quantity accounts of CLI. In particular, they suggest that structural constraints play a pivotal role in shaping the direction of the phenomenon. At the same time, they highlight the importance of input quantity in determining the magnitude of CLI effects.

These studies are insightful with regards to CLI insofar as they test the same syntax-discourse interface structures. These structures are known loci of CLI; thus, their use to 'prove' the presence of CLI could be argued to resemble a self-fulling prophecy[2]. To truly determine the effect(s) of CLI in bilingual development, it would be imperative to (also) test non-interface structures that are not deemed to be vulnerable to CLI. These structures would provide a 'control' condition from which to compare the bilinguals' use of non-interface structures with their use of interface structure. In turn, this process could elucidate the role of CLI in the bilinguals' development, and demonstrate whether interfaces are indeed more vulnerable than other linguistic domains.

1.1.2.2. New Lessons from Sequential Bilinguals

Another issue with the current state of CLI literature is that it tends to focus on specific populations, like simultaneous bilingual children (i.e., children who learn their two languages from birth) and/or L2 learning adults. However, simultaneous bilingual children represent a small portion of bilingual learners in total. Indeed, numerous children learn their L2 after their L1 has started developing – sometime between the ages of 3;0 and 6;0 (Schwartz, 2004; Unsworth, 2005). These children are referred to as sequential (or successive) bilinguals. Researchers agree that sequential bilingualism is different from simultaneous bilingualism, but the (age) boundaries that separate simultaneous and sequential bilinguals as well as sequential bilinguals and child L2 learners are unclear and might differ across linguistic domains and processes (Meisel, 2008; Unsworth, 2013b).

[2] It should be noted that most of the studies mentioned in the previous paragraphs, including Tsimpli and Roussou's (1991) and Prenza and Tsimpli's (2013) work, were not conducted to examine the phenomenon of CLI itself but, rather, the distribution of null and overt subjects from a syntactic perspective. As such, they do not speak for or against the Interface Hypothesis or the Subset-Superset Account, which highlight the special status of interface structures in adult L2 and child simultaneous acquisition respectively. Nevertheless, as their findings point towards the presence of (negative) transfer effects from the L1 to the L2, they can also be considered to provide evidence in support of the (persisting) vulnerability of syntax-discourse interface structures.

Studies that investigate CLI in the context of sequential bilingualism give rise to interesting results; these shed light into how the causes of CLI that have been proposed in the literature might apply to this population. For instance, Zdorenko and Paradis (2011) examined the production of definite and indefinite articles (i.e., 'the' and 'a') by 40 sequential bilingual children between the ages of 5;0 and 6;11. The learners were divided in four groups based on their L1 background: 10 were L1 Chinese speakers, 10 were L1 Hindi/Urdu/Punjabi speakers, 10 were L1 Spanish speakers and the final 10 had Arabic as their L1. The authors used a narration task to elicit the target structures. They reported that CLI appeared in the form of article omissions, and was evident in the speech of children whose L1s do not use articles (i.e., Chinese and Hindi/Urdu/Punjabi). In contrast, children whose L1s possess an article system did not appear to transfer the properties of their L1 to the L2. Zdorenko and Paradis (2011) also examined the effect that the children's length of exposure to English had on their use of articles. They reported that the learners' use of the definite article improved overtime. They concluded that, even when present, the effects of CLI are ephemeral.

These results do not demonstrate clear evidence for traditional accounts of CLI. Linguistic structure only appeared to matter in the extreme case when the children had to acquire a new linguistic category (and its features) from scratch. Likewise, the amount of input children received in the L1 and/or the L2 mattered with respect to the acquisition of definite articles, while indefinite articles remained problematic. These observations prompted the authors to suggest that the sequential bilingual children's acquisition of definite and indefinite articles appears to progress according to L1 developmental norms.

In a more controlled experiment, Hervé et al (2016) tested 78 English-French bilingual children in the UK and France with the aim of measuring their use of left dislocations. These are syntax-pragmatics interface structures, whose use differs between English and French; while both languages allow left dislocations, their use in French is extensive and unrestrained.

The researchers primed children with dislocated or non-dislocated sentences (e.g., 'the astronaut, he is in space' versus 'the astronaut is in space'), while showing them pictures, and then asked them to describe similar pictures in their own words. They reported that children produced more left-dislocated utterances after the relevant prime, but children with higher exposure to French produced more left dislocations in this language regardless of the prime. By contrast, children with higher exposure to English produced few left dislocations in general.

Hervé et al's (2016) findings support input quantity accounts of CLI, as bilingual children were found to pattern with monolingual speakers of their dominant language. Yet, the study offers little and unclear evidence in favour of the idea that bilingual children's production of left dislocations is affected by the different grammatical constraints of French and English. To test the effect of linguistic structure, it would be useful to compare children's production of interface structures (which induce CLI) and narrow syntax structures (which do not).

This task was undertaken in Daskalaki et al's (2019) replication of Argyri and Sorace's (2007) study. Daskalaki et al (2019) tested 62 second- and third-generation Greek heritage learners in Canada and the US, focusing on subject realisation in non-topic-shift contexts (where Greek requires null subjects), and subject placement in wide focus contexts and in wh-interrogatives (where Greek requires post-verbal subjects). They expected CLI in the case of subject realisation, which is an ambiguous interface phenomenon, but not in the case of subject placement, which is an unambiguous syntactic phenomenon. Daskalaki et al (2019) reported that bilingual children correctly produced null subjects in non-topic shift contexts a lot of the time and post-verbal subjects in wh-interrogative clauses some of the time. Yet they also used a high number of illicit pre-verbal subjects in wide focus contexts. The authors also noted that increased exposure and language use improved children's accuracy.

These findings support the role of input quantity, as children with higher exposure and increased language use performed more target-like in all three structures. Yet, they challenge

structural linguistic accounts, since subject realisation does not seem to be as susceptible to CLI as much as subject placement. However, it should be noted that the linguistic structures that the authors chose to investigate are not 'equal': the subject realisation condition involves the production of a single noun phrase, while the subject placement condition involves the production of an entire sentence with the correct word-order, which might be a more difficult task for learners. To test the effect of structural factors without misrepresenting learners' performance, the target structures should ideally be equally complex.

As the three examples discussed in this section illustrate, studies conducted with sequential bilingual children suggest that the transfer of grammatical properties occurs from the language in which the learners receive more input to the language in which they receive less input. This effect is in line with input quantity accounts of CLI. By contrast, the relative structure of the bilinguals' languages does not seem to drive the transfer of grammatical properties as much. In fact, even in cases when L1 structure seems to affect children's linguistic behaviour, its effects are unclear and/or fleeting. Therefore, it appears that the explanatory power of structural linguistic accounts of CLI is more limited in this learner population. This observation highlights the need for additional research on sequential bilingual children.

1.2. Educational Background

In addition to the theoretical reasons in favour of focusing on sequential bilingual children's acquisition, investigating the linguistic development of sequential bilingual children is of practical importance. Across the globe, the starting age of L2 acquisition is decreasing (OECD, 2014). This trend is widespread in Europe, where children often start learning an L2 at preschool (Eurydice, 2017). Fuelled by the belief that 'the younger the language learner, the better the language learning', the European Council set its member countries the objective to teach children at least two foreign languages from an early age during its 2002 meeting in Barcelona, (Eurydice, 2017). This belief stems from the idea that language acquisition is

difficult or impossible after a certain age which tends to coincide with the start of puberty (see Lenneberg, 1967). Although this idea has been dispelled on theoretical and empirical fronts (for a review, see: Muñoz & Singleton, 2011), it remains crucial in shaping language learning policies (Enever, 2011; Genesee, 2016; Rixon, 2013).

Indeed, driven by the belief that 'the younger, the better', parents are dissatisfied when FL learning is not part of their country's national preschool provision, and seek alternative means of introducing their children to other languages (Murphy & Evangelou, 2016; Rixon, 2013). This phenomenon occurs in a manic level in Greece, where the national provision did not include L2 tuition (Dendrinos, Zouganeli & Karavas, 2013). Officially, Greek schools are expected to introduce children to their first foreign language, English, at second grade (from 8;0) and to their second foreign language (which can be French, German or Spanish), at fifth grade (from 11;0) (Eurydice, 2017). Nevertheless, it is known that many Greek children are exposed to English before 8;0, as a result of attending private preschools (Dendrinos et al., 2013). These preschools are allowed to circumvent the national guidelines and can include L2 teaching in their curricula (Law Nº 1566/1985).

Thus, in Greece there exist monolingual immersion preschools as well as bilingual immersion preschools. The 'monolingual' preschools offer a total immersion programme, where children are 'submersed' in English: L1 English teachers deliver the British National Curriculum in English, and without using the children's L1 to scaffold language learning. The 'bilingual' preschools offer a partial immersion programme, such that L1 English and/or L1 Greek teachers deliver a custom curriculum that combines the two languages, such that some classes are in Greek and some are in English. In both kinds of preschools, there is no explicit instruction of the L2. Rather, children are expected to 'pick up' the language through mere exposure to input – much like their English monolingual peers.

Although at present both immersion programmes are exclusively operated in private institutions, the Greek educational landscape is changing. In September 2020, the Minister of Education launched a pilot action that involved a select group of state preschools around Greece introducing English language lessons as part of their provision (Law Nº 4692/2020). The English language lessons lasted two hours per week, involved a series of creative and interactive activities, and were delivered with the help of a qualified English language teacher (who was an L1 Greek speaker). In June 2021, the Minister decided to extend the pilot action across Greek state preschools, such that English language lessons are to be incorporated in the curriculum of Greek preschools from the academic year 2021-2022 onwards.

Yet, a lot about the incorporation of English tuition in the curriculum of Greek preschools remains unspecified: for instance, the details of the training that the preschool and the visiting English language teachers will receive in order to become equipped for teaching in a different language and/or to an unfamiliar age group are not presented in the bill; the information about the resources allocated to the preschools and their staff in order to facilitate the provision of English language lessons is limited; in addition, the expected linguistic and academic outcomes of the pilot action have not been specified.

These are important omissions, since the teachers' training, the allocation of resources and the creation of appropriate curricula contribute to the success of a dual language programme (Murphy & Evangelou, 2016). If done well (that is, delivered by well-trained teachers, with sufficient resources and well-devised curricula that allow bilingual children's two languages to develop in parallel), language learning at preschool can have multiple benefits, such as better oral language and literacy skills in the L2, greater levels of confidence, and higher aspirations to continue school (Faitaki, Hessel & Murphy, 2020; Genesee, 2009; Murphy, 2014). Yet, dual language programmes are often designed and delivered in a rushed and unstructured fashion, in an effort to deal with parental pressure (Rixon, 2013); then,

children's lexical and grammatical knowledge in one or both of their developing languages might remain unsupported and not reach its full potential (Murphy & Evangelou, 2016).

As dual language programmes are a new (but seemingly spreading) trend in Greece, the linguistic progress of the pupils that attend them has not been subject to systematic investigation. Current available evidence (or lack thereof) indicates that this thesis constitutes the first attempt to explore and compare different types of English immersion programmes in Greece. Finding out more about the pupils' linguistic development as part of these programmes could help design suitable curricula and resources, which can then be used to support sequential bilingual children's linguistic skills in Greece's existing or prospective language learning programmes (Strand & Hessel, 2018).

1.3. Linguistic Background

The previous sections highlighted the theoretical and practical motivations to investigate the development of Greek-English sequential bilingual children. At the same time, Greek and English constitute a useful language combination for studying CLI, as their similarities and differences both give rise to the cross-linguistic transfer of grammatical properties and allow for its investigation.

1.3.1. Referential Expressions

A linguistic dimension where Greek and English differ is the use of referential expressions, that is, the terms used to refer to entities and objects (e.g., personal pronouns and proper nouns). On the one hand, and as discussed earlier on in the chapter, Greek (like Spanish and Italian) is a null-subject language – in other words, it allows referential expressions to remain phonologically null when in subject position; on the other hand, English is a non-null-subject language – as such, it requires the subject referents to be overt in almost all sentential contexts (Alexiadou & Anagnostopoulou, 2000). The difference between the use of sentential subjects in English and Greek is illustrated by (1) and its Greek translation in (2).

1. *(He) is eating the watermelon.
2. (Aftos) troi ena karpouzi.
   He eats a.NEUT watermelon
   'He's eating a watermelon'

In addition to the distribution of null and overt subjects, the two languages are somewhat different with respect to the referential expressions that various sentential contexts require. Example (3) illustrates the referential expressions used for topic introduction, topic maintenance across sentences and topic shift: the subject is first introduced using an overt Determiner Phrase (DP); topic maintenance across sentences is achieved using an overt pronoun; and the switch to a new subject requires the use of an overt DP (i.e., 'the apples').

3. Alex is in the garden. He is picking apples. The apples are ripe.

The optimal referential expressions for each of these four discourse contexts in Greek can be observed in the translated version of the English example above, which is presented in (4). Note that, once again, the subject is introduced using an overt DP. However, on this occasion, topic maintenance across the first two sentences is achieved using a null argument. Rather, the switch to a new subject is achieved using an overt DP as is the case in English.

4. O Alex ine ston kipo. [...]
   The.MASC.SG Alex is in-the.NEUT.SG garden
   [...]Mazevi ta mila. [...]
   Picks the.NEUT.PL apples
   [...] Ta mila ine orima.
   The.NEUT.PL apples are ripe

In sum, as Table 1 highlights, English and Greek use similar referents to signal topic introduction and topic shift, as well as topic maintenance within the same clause. A relevant difference between the two languages concerns the choice of referents used for topic

maintenance across clauses[3]: this discourse context requires an overt argument in English (e.g., a pronoun), but a null argument in Greek.

Table 1. Referential expressions used in different discourse contexts in English and Greek.

	Introduction	**Shift**	**Maintenance (Across Clauses)**
English	Overt (DP)	Overt (DP or Pronoun)	Overt (Pronoun)
Greek	Overt (DP)	Overt (DP or Pronoun)	Null

The acquisition of referential expressions has received substantial attention in the literature. It is known from naturalistic studies that both L1 English and L1 Greek children use sentential subjects (in their various language-specific versions) with sufficient accuracy from around 2;6 (e.g., for English: Bloom 1990; for Greek, Tsimpli, 2005). Children appear to be familiar with the discourse-appropriate forms of referential expressions from the age of 2;0 onwards (for reviews, see: Allen, Skarabela & Hughes, 2008; Graf & Davies, 2014).

However, it is known that L1 English children have a tendency to omit subject pronouns from their utterances before 2;6 (Bloom, 1990; Valian, 1991). These omissions dissipate over time, as children become more familiar with the rules that determine the use of subjects in their language. In addition, there is experimental evidence that L1 Greek children differ from adults in terms of their preferred interpretation of null and overt subject arguments as well as in terms of the pronoun resolution strategies they use during sentence comprehension.

These issues were evidenced in an experimental investigation carried out by Papadopoulou, Peristeri, Plemenou, Marinis and Tsimpli (2014). The authors tested 75 L1 Greek children (20 in Grade 1, 22 in Grade 2 and 31 in Grade 3) alongside 36 L1 Greek adults

[3] This is obviously not the sole point where the two languages diverge: as the use of null subjects is a central, macro-parametric difference between the two languages, it impacts many other properties of the language; these include verbal morphology (as evidenced in examples 2 and 4), subject placement and word order.

on two experiments: one focusing on null and one on overt subject resolution. The experiments used a computer-mediated sentence-picture matching task that involved participants listening to a sentence and selecting whether the sentence captured the meaning of the picture that appeared on the screen or not. The results of the two experiments suggested that all children, like their adult counterparts, could differentiate between the two structures and were aware of the grammatical constraints that underpin the distribution of null and overt pronouns. However, at least younger children, were not able to integrate the pragmatic constraints associated with null/overt pronoun use in real time in order to derive the appropriate meaning (for a similar observation, see: Clahsen & Felser, 2006).

The acquisition and use of referential expressions has also been a prominent topic in bilingualism research, as focusing in null- and non-null-subject languages allows for the investigation of CLI (Serratrice, 2013). As the various studies reported earlier in this chapter suggest, bilingual children tend to follow the same developmental trajectory as their monolingual peers with respect to subject realisation, but their acquisition is often impacted by CLI (see Tsimpli, 2014). They also appear to use the discourse-appropriate forms of referential expressions like their monolingual peers (Paradis & Navarro, 2003; Serratrice et al., 2004).

1.3.2. Subordinate Clause Markers

A linguistic dimension where Greek and English are similar is the use of subordinate clause markers, that is, the markers used to introduce a subordinate clause. The two languages make use of various types of subordinate clauses (e.g., adjectival, adverbial). In both Greek and English, subordinate clauses are introduced using specific markers, which also help identify the type of the subordinate clause (e.g., adjectival clauses start with relative pronouns and adverbial clauses start with subordinating conjunctions).

One such marker used in English is the morpheme ‘to’. ‘To’ heads infinitival clauses (thus called due to the fact that they contain non-finite verbs). The linguistic status of ‘to’ has

perplexed researchers: some of the first investigations into the nature of the morpheme suggested that it is a complementiser (Postal & Pullum, 1978), while subsequent work has considered it to be a tensed particle (Chomsky, 1981) or even an auxiliary (Pullum, 1982). As these linguistic debates are beyond the scope of the thesis, I will be referring to 'to' as a subordinate clause marker, ignoring the additional properties that the morpheme might have.

In Greek, there is an equivalent of 'to': 'na'. This marker can have a pragmatic function (e.g., deixis) as well as a grammatical function, since it heads subordinate clauses. Joseph (1981) offered the first detailed account of the structure 'na'. He observed that the subjunctive 'na' (a) functions as a clitic, as it is unstressed and connected to the verb it introduces (b) precedes the verb or a clitic pronoun governed by the verb and (c) appears in the beginning of the sentence only when it introduces a verb. Thus, he contrasted the structure with the deictic 'na' which (a) is stressed and does not have to be connected to a verb, (b) precedes a DP, and (c) appears in the beginning of the sentence. Philippaki-Warburton and Veloudis (1985) explained that the subjunctive 'na' appears with verbs that communicate 'potential worlds'. Indeed, 'na' often accompanies the verbs 'want', 'wish' or 'desire' – all of which connote an unfulfilled action. Following on from this suggestion, Christidis (1985) argued that the subjunctive 'na' assumes a deictic function (pointing to the verbs' potential worlds), thus resembling its deictic counterpart.

However, the opinions regarding the linguistic status of 'na' are (also) divided: Philippaki-Warburton and Veloudis (1985) analysed the structure as a marker of subjunctive clauses, Tsimpli (1990) argued it to be a modality marker that agrees with (un)tensed phrases, Agouraki (1991) analysed the structure as a complementiser (see Fiotaki, 2014), while Christidis (2002) stated that 'na' should be referred to a particle as this term allows the various meanings and functions of the structure to co-exist. Once again, I will steer clear from the linguistic debates and simply refer to 'na' as a subordinate clause marker in the thesis.

In sum, both ‘to’ and ‘na’ are subordinate clause markers; however, ‘to’-clauses contain non-finite verbs, while ‘na’-clauses can also contain tensed verbs (Tsimpli, 1990). Despite this difference, the markers are required to head subordinate clauses; without their presence, the clauses that follow them are not legitimised. This is illustrated in the English sentence in (5) and its Greek equivalent in (6):

1. He wants *(to) see a film
2. Theli *(na) dhi mia tenia
   Wants to sees.SUBJ a.FEM film

Given the different opinions regarding the status of the markers in both English and Greek, it will not come as a surprise that their acquisition and use by monolingual and/or bilingual children has not been investigated extensively. Research on L1 English children suggests that they acquire ‘to’ from 3;0, and use it consistently from 3;6 (Bloom, Tackeff & Lahey, 1984; Diessel, 2004; Kirjavainen & Theakston, 2011; Kirjavainen, Lieven & Theakston, 2017; Kirjavainen, Lieven, Theakston & Tomasello, 2009). However, there is a period during which L1 English children omit ‘to’ from their non-finite subordinate clauses.

Research on when and how ‘na’ is acquired by L1 Greek children is limited, but it is known from observational studies that subordinate clauses are present in children’s utterances from 3;0 (Tsimpli, 2005). An experimental study by Mastropavlou and Tsimpli (2010) suggests that ‘na’ is also omitted at a high rate by children under 6;0. The authors used speech samples collected over three 45-minute sessions to assess the acquisition of subordinate clauses in 18 L1 Greek children: eight had Specific Language Impairment (SLI) and served as the experimental group, while the remaining 16 served as controls: half of these were matched with the SLI group in terms of age (mean: 5;1) and the other half in terms of linguistic development (mean age: 3;2). The results indicated that ‘na + verb complement’ clauses were the most frequent subordinate clauses in all of the children’s samples. Nevertheless, there was a high

rate of omissions of 'na' in the SLI and the language-matched groups' data, suggesting that a stage of 'na' omission is normal at the earliest stages of acquisition (as is the case for L1 English children), and that the stage is extended in the case of SLI. Thus, the present thesis constitutes one of the first experimental investigation of the structure in L1 Greek children – not least under the lens of bilingualism.

1.4. The Thesis at a Glance

Taking the theoretical, educational and linguistic considerations discussed throughout this chapter into account, the main goal of the thesis was to explore whether CLI affects the linguistic development of Greek-English sequential bilingual children. A secondary goal was to investigate the relative strength of linguistic structure and input quantity in accounting for the CLI patterns observed in this group. As the Greek-English sequential bilingual children were pooled from a total immersion preschool and a partial immersion preschool, an additional aim of the thesis became to explore the extent to which the children who attend the different preschools differ in terms of lexical and grammatical proficiency.

To meet its aims, the project comprised three experiments. Experiment 1 investigated Greek-English sequential bilingual children's use of sentential subjects. Due to their status as 'syntax-discourse interface structures', and the fact that their manifestation differs in English and Greek, sentential subjects are considered prone to CLI (Sorace, 2011). Experiment 2 explored the same children's use of subordinate clause markers. As these markers are only regulated in the domain of syntax and appear in a similar form in English and Greek, they are not considered prone to CLI. To explore whether the children's (mis)use of the structures persists at a more advanced stage of L2 acquisition, a third experiment was conducted to investigate the use of sentential subjects (the more 'problematic' of the two structures under investigation) in adult Greek learners of EFL.

Capturing the design of the doctoral project, the thesis is structured as follows: this chapter explained the theoretical, empirical, and pedagogical motivations behind the project; Chapter 2 will proceed to outline the background, methods, and results of Experiment 1; Likewise, Chapter 3 will present the details of Experiment 2, while Chapter 4 will focus on Experiment 3; finally, Chapter 5 will conclude the thesis with an overview of the findings and a discussion of what they mean in relation to the project's overarching aim.

Chapter 2: Bilingual Children's Use of Sentential Subjects

Experiment 1 was conducted to uncover the factors that determine the appearance of crosslinguistic influence (CLI) in sequential bilingualism, focusing on the use of sentential subjects (and, in particular, the third person singular subject pronouns 'he' and 'she'). This structure was chosen since it differs between Greek and English. Greek is a null-subject language, so it allows sentential subjects to be voiced (overt) or unvoiced (null) – with the former being used in special contexts, like the introduction of a new topic. In contrast, English is a non-null-subject language and requires subjects to be overt in all contexts, bar the continued reference to a topic within the same clause.

Investigating the structure is also worthwhile from a different perspective: the realisation of subject arguments is a well-known locus of CLI effects and, thus, a well-investigated topic in the CLI literature (Serratrice, 2013). Most studies conducted thus far target either simultaneous bilingual or second language (L2) learners. These have revealed that the relative structure of the two languages that bilingual children are learning can shape the direction of CLI. In other words, the phenomenon is reported to appear in situations when the learners acquire a language that allows multiple manifestations of sentential subjects (i.e., a null-subject language) together with a language that allows one of these manifestations (i.e., a non-null-subject language). In these situations, simultaneous bilingual children appear to transfer the grammatical properties of the unambiguous language to the ambiguous language, thus overusing the overlapping manifestation: overt subjects (Argyri & Sorace, 2007; Hacohen & Schaeffer, 2007; Schmitz et al., 2012; Serratrice et al., 2004; Paradis & Navarro, 2003). By contrast, L2 learners of an unambiguous non-null-subject language appear to transfer the grammatical properties of their ambiguous null-subject first language (L1), thus overusing its 'default' manifestation: null subjects (Prenza, 2014; Prenza & Tsimpli, 2014; Tsimpli & Roussou, 1991; White, 1985; 1986).

In addition to linguistic structure, the appearance of CLI effects might depend on the learners' linguistic dominance, that is, the quantity of input that they receive in either language. Input quantity accounts of CLI suggest that the phenomenon appears as the transfer of grammatical properties from the language in which the learners receive more input (and are more dominant) to the language in which they receive less input (and are less dominant). Linguistic dominance has been found to decrease the magnitude of CLI effects in simultaneous children (e.g., Argyri & Sorace, 2007). Recent evidence (e.g., Daskalaki et al., 2019) suggests that input quantity accounts of CLI might be better at accounting for the CLI patterns observed in sequential bilingual children's speech than structural linguistic accounts, as these children receive more input in their L1 and have limited proficiency in their L2. Indeed, studies on adult L2 learners suggest that increasing proficiency in the L2 can make CLI effects dissipate (e.g., Prenza & Tsimpli, 2014).

This information is important in light of the fact that the experiment targeted two groups of Greek-English sequential bilingual children. One group attended a monolingual (or total) immersion preschool programme, where the curriculum is delivered in English (the children's L2). The other group attended a bilingual (or partial) immersion preschool programme, where half of the lessons take place in English and the other half take place in Greek. As the preschools are based in Greece, and the pupils' L1 tends to be Greek, the exposure to English also tends to be limited to the school context. Thus, both groups of Greek-English sequential bilingual children who took part in the experiment were dominant in Greek, but children in total immersion received more input in English than children in partial immersion.

With these issues in mind, Experiment 1 was created to answer the following questions:

1) How does the bilingual children's use of subject pronouns in their two languages differ:
    a) from that of their monolingual peers?

b) depending on whether they attend the total immersion or the partial immersion preschool programme?

2) How does the bilingual children's linguistic proficiency in Greek and English affect their use of the target structure in the two languages?

With regards to (1) structural linguistic accounts of CLI would predict transfer from English to Greek, leading to the overproduction of overt subjects in Greek, while input quantity accounts of CLI would predict transfer from Greek to English, resulting in the underproduction of overt subjects in English. As the present study targets sequential bilingual children, input quantity accounts should provide a better explanation: as such, I expect a main effect of group in the English assessment, such that: (a) bilingual children produce fewer overt subjects than English monolinguals, and (b) children in total immersion perform more like the English monolinguals than children in partial immersion. I do not expect a main effect of group in the Greek assessment: rather, I predict that the bilingual children will produce a comparable number of overt (and null) subjects relative to Greek monolinguals, and that there will be no differences between children in total and partial immersion. As far as (2) is concerned, I expect linguistic proficiency (operationalised as children's performance in a vocabulary and grammar test) to play a role in children's performance in English (their L2), such that children with higher lexical and grammatical knowledge in English are predicted to perform better and more like their English monolingual peers.

These predictions are associated with the experimental condition of the experiment which, as will be discussed in the following section of the chapter, tested the manifestation (or lack thereof) of the subject argument. In addition, the experiment made use of a comparison condition that assessed the disambiguation of the subject referent through the use of a Determiner Phrase (DP). As the condition necessitated the manifestation of the subject argument, the predictions for the comparison condition are similar to the predictions for the

experimental condition. In particular, for question (1), I expect a main effect of group in the English but not in the Greek assessment while, for question (2), I expect children's English proficiency to affect their control item performance in the English assessment (where group differences are predicted to arise). Moreover, I expect a main effect of condition: as the comparison condition involves additional cognitive constraints than the experimental condition, children should produce fewer overt subjects in the former than in the latter regardless of their linguistic status. Finally, given the theoretical background of the project, I do not expect to attest an interaction between group and condition.

2.1. Methods

In this section I will present the methods used to address the aforementioned questions. I will describe the characteristics of the children who took part in the experiments in greater detail, outline the design and procedures involved in the experiment, and explain how the collected data was prepared for analysis.

2.1.1. Participants

The participants were recruited through their schools, using stratified random sampling (Rose, McKinley & Briggs Baffoe-Djan, 2020). This sampling process involves the researcher(s) dividing the population into groups based on their characteristics, and then selecting participants for each group at random. Following this procedure, an extensive online search was conducted to populate the groups of the study. The search allowed identifying numerous schools whose pupils met the criteria for inclusion in England (N=13) and Greece (N=12). All suitable schools were emailed an invitation to take part in the study (Appendix A). Three schools in England and four schools in Greece accepted the invitation. An in-person or online meeting was arranged with the heads of these schools to discuss the details of their potential involvement in the project further.

Through the meetings it became clear that one of the three English schools was not a good fit, as most of its pupils were under the age of 4;0. This school was excluded from the sample which ended up comprising two private schools, one in Oxford and one in London. Likewise, one of the Greece-based schools was also deemed unsuitable: though it operated an English language programme where pupils received five hours of English instruction per week (which would be interesting to investigate further), it accepted a small number of pupils per academic year and did not demonstrate a strong interest in being involved. Thus, the three Greece-based schools that formed part of the sample were: a semi-private Greek school in Thessaloniki, as well as a private total immersion school and a private partial immersion school located in Athens.

The child sample initially consisted of 158 children between the ages of 4;0 and 5;11. Of these, 83 were Greek-English bilingual, 37 were English monolingual and 38 were Greek monolingual. The bilingual group was further divided into the total immersion group, consisting of 41 children between 4;10 and 5;11, and the partial immersion group, which included 42 children between 5;0 and 5;11. According to the results of a power analysis, conducted based on the results of a pilot study (that will be discussed later on in this chapter) using the G*Power 3 programme (Faul, Erdfelder, Lang & Buchner, 2007), 33 participants per group would be required in order to achieve 95% power at the α=.05 level [(1-β)=.95] for main effects. Thus, the number of recruits per group allowed for the experiment to be sufficiently powered, at least as far as group comparisons are concerned[4].

To be included in the sample, monolingual children had to be between the ages of 4;0 and 6;0, and to fit in one of the monolingual language groups (English or Greek). Bilingual children also had to be between 4;0 and 6;0, but they also had to have Greek as their only L1,

[4] The power analysis was conducted as part of a preregistration of an earlier version of the experiment. The preregistration was uploaded at the Open Science Framework (OSF), but had to be withdrawn as aspects of the design it included could not be implemented due to the COVID-19 lockdowns.

speak Greek at home, and learn English through school. Since bilingual children's exposure to English was limited to the school context, the age at which they joined the preschool (i.e., from 3;0) was also their starting age of L2 acquisition.

Yet, some of the children who formed part of the initial sample did not meet the inclusion criteria and/or did not manage to participate in all of the tasks. Three children (one from the total immersion group, one from the partial immersion group and one from the Greek group) were unable to take part, because one of their parents spoke a language other than Greek at home; 11 bilingual children (four from the partial immersion group and five from the total immersion group) did not participate in the English treatment due to timing constraints; two children from the partial immersion group were eliminated because they were incapable of completing the experimental tasks in English; and five children (two from the English group and three from the Greek group) were eliminated because they were uncooperative and/or unable to pass the training items in (some of) the tasks. Thus, the final sample consisted of 140 children: 70 Greek-English bilingual (35 in the total immersion group and 35 in the partial immersion group), as well as 36 English monolingual and 34 Greek monolingual children. Despite the attrition, the number of participants per group was still sufficient to achieve the required statistical power. The characteristics of the participants are presented in Table 2.

Table 2. Characteristics of the Child Participants.

	Male	**Female**	**Mean Age (Months)**	**SD Age (Months)**	**Total**
Monolingual English	22	14	61	6.80	36
Monolingual Greek	17	17	60	5.63	34
Bilingual (Total Immersion)	16	19	64	4.35	35
Bilingual (Partial Immersion)	16	19	66	3.50	35
Total	71	69	62.75	5.07	140

2.1.2. Design

The dependent variable of this experiment was the number of (overt and null) subjects that children used. The within-subjects independent variables were the linguistic treatment (English vs Greek) and the condition (experimental vs comparison), while the between-subjects independent variable was language group (bilingual Greek-English in total immersion vs bilingual Greek-English in partial immersion and/or vs monolingual English and/or vs monolingual Greek)[5]. Children's lexical and grammatical proficiency functioned as both a dependent variable (facilitating between-group comparisons) and as an independent variable (predicting children's performance in the experiment).

The number of sentential subjects resulted from counting the uses of the structures in a Question-and-Answer task, which required participants to answer a question prompt that was framed so as to enforce the use of a target structure. In general, these tasks allow the researchers substantial control over the participants' production, provide a semi-naturalistic discourse context, and are suitable for testing children who would perhaps struggle with processing a lot of (linguistic and contextual) information at the same time (Ambridge & Rowland, 2013). An additional benefit is that they can be used together with structural priming. Structural priming refers to the tacit tendency of two interlocutors to use similar structures during a discourse interaction (Branigan & Pickering, 2017). Applied to the task at hand, structural priming was achieved by using a structure in the question prompt in order to influence children to use the same structure in their response.

[5] Dividing the bilinguals into two groups, total immersion and partial immersion, and comparing them to English and Greek monolinguals allowed for an investigation of 'input quantity'. Children in total immersion receive more input in the L2 (English) and less input in the L1 (Greek) than children in partial immersion. Therefore, comparing their results to English and Greek monolingual children's in the English and Greek assessment respectively could shed light on the role of input quantity, as well as on the (expected) relationship between input quantity and proficiency.

To get an estimate of children's linguistic proficiency, tests of receptive vocabulary and grammar were used. The motivation behind measuring children's understanding of words was the fact that receptive vocabulary knowledge is known to be a good indicator of proficiency (Miralpeix & Muñoz, 2018) as well as a reliable predictor of academic achievement (Bialystok, Luk, Peets & Yang, 2009). The decision to also measure children's receptive grammar knowledge was driven by an interest in obtaining a holistic estimate of children's proficiency, as well as by the observation that the two linguistic structures under investigation (i.e., sentential subjects and subordinate clause markers) pertain to the realm of syntax and/or its intersection with pragmatics. All the tests that were chosen in the context of the study are normed and reliable instruments that can be used with children of the target age-group. Yet, their norming process involved monolingual children and, as such, the use of the tests with bilingual children is questionable (Faitaki & Murphy, 2020). To mitigate against this potential validity issue, the tests' raw scores were used in lieu of the standardised scores.

2.1.2.1. Vocabulary Tests

The third edition of the British Picture Vocabulary Scales (BPVS3; Dunn, Dunn & Styles, 2009) was used to measure children's receptive vocabulary in English. The BPVS3 uses a multiple-choice format to assess children's knowledge: the assessor says a word, and the participant is asked to select one of four pictures that illustrates the meaning of the word. The test consists of 168 words, organised in 14 sets of increasing difficulty. Children start from a specific set depending on their age, and proceed to complete as many sets as possible until the procedure is terminated or the test comes to an end. For the former to happen, the participants should make eight (or more) mistakes in a set.

Children's receptive vocabulary in Greek was measured using the Greek edition of the Peabody Picture Vocabulary Test - Revised (PPVT-R Greek; Simos, Sideridis, Protopapas & Mouzaki, 2011). The PPVT-R has the same multiple-choice format as the BPVS3. The original

edition of the PPVT-R consists of 175 items (five are used to train the participants), but the Greek edition consists of 173 items as two words (and their corresponding image plates) were eliminated by the test creators. Taking language-specific considerations into account, the creators of PPVT-R Greek also changed the target stimuli of 44 plates and modified the order in which the items are presented. Unlike the BPVS3, the PPVT-R (and its Greek counterpart) is not divided into sets: the test commences from a 'starting' item based on children's age and proceeds until the child makes six mistakes in a sequence of eight consecutive items.

2.1.2.2. Grammar Tests

The second edition of the Test for the Reception of Grammar (TROG2; Bishop, 2003) was used to measure children's receptive grammar in English. The TROG2 consists of 80 test items, arranged in 20 blocks of four, so that each block assesses children's understanding of a specific grammatical construction four times. Like the vocabulary tests described above, the TROG2 also uses a multiple-choice format. Regardless of their age, all participants start from the test's first item (after succeeding in the test's two practice items), and proceed to complete as many blocks as possible until the test comes to an end – unless the participants are unable to correctly identify all the items in four consecutive blocks, which terminates the procedure.

Children's receptive grammar in Greek was measured using the 'Comprehension of Morphosyntax' subtest of the Diagnostic Verbal Intelligence Quotient (DVIQ; Stavrakaki & Tsimpli, 2000). As a whole, the DVIQ assesses children's productive and receptive skills, as well as their ability to recall linguistic forms. The test contains subtests on children's vocabulary, morphosyntax, and metalinguistic awareness. Like all of the aforementioned instruments, the 'Comprehension of Morphosyntax' subtest makes use of a multiple-choice format. The subtest consists of a list with 31 sentences, testing various constructions, presented in a random order within the list. The administration of the subtest commences from the test's first item and proceeds until the last item, regardless of the children's age or performance.

2.1.2.3. Experiment Materials

The target structure of Experiment 1 was third person singular subject pronouns 'he' and 'she'. The stimuli consisted of 23 items: 10 experimental (five for 'he' and five for 'she'), 10 comparison, and three practice items (of which two had the same format as the experimental items and one had the same format as the comparison items) (Appendix B). The comparison items were used to identify whether children's production/omission of sentential subjects was specific to the experiment's target structure(s) or extended to all nominal arguments. The stimuli were presented alongside a matching picture; indeed, 23 pictures were hand-drawn and coloured (Appendix C).

In both English and Greek, all stimuli were between 10 and 11 words long and all target utterances (i.e., the utterances that children were hoped to produce) were expected to be between four and six words long – although children could use more or fewer words in their utterance. For English, the verbs in all of the stimuli were among the 50 most frequent in children's input according to the Manchester Corpus (Theakston, Lieven, Pine & Rowland, 2001) found in CHILDES database. Moreover, all the other words were within children's age norms (Kuperman, Stadthagen-Gonzalez & Brysbaert, 2012). For Greek, creating items that were at the right level was more difficult because the Greek corpora available on CHILDES are limited, and there are no available Greek age-of-acquisition ratings. Therefore, the items created for the English treatment were translated into Greek. To ensure that the translations contained words that were frequent and familiar to children, two preschool teachers (from the Greek monolingual school) were asked to review the items and words therein. The teachers confirmed the items to be appropriate for children's age and linguistic development.

All items were about the same two characters, Nick and Anne, as well as their friends and relatives, who were introduced before the experiment started. The items were composed of two parts: a prime and a question. The prime stated the actions that the characters performed

on a regular basis. It also served as an aid for the children, as it used the verb that they were expected to produce in their target utterance. The question ("what's happening today?") was aimed to elicit children's responses. To answer the question, children were expected to produce (a version of) the target utterance, describing the actions that the characters were performing in the picture. The format of the prime sentence was different for the experimental and the comparison items: the former contained the character's name, thus warranting the use of a subject pronoun in the target utterance, while the latter contained either a conjoined DP (e.g. "Nick and his teacher") or a possessive DP ("Nick's parents"), so as to force children to use a disambiguating DP in their target utterance. As such, the experimental condition tapped into a grammatical issue (testing the manifestation of the subject argument) while the comparison condition tapped into a pragmatic issue (testing the disambiguation of the subject argument's referent). Two example items (and pictures) can be found in Table 3 and Figure 1.

The stimuli were arranged into two different semi-random lists, while the pictures that corresponded to the stimuli were duplicated and placed in two picture booklets. The order of the pictures differed in each booklet, as it corresponded to one of the semi-random item lists. Children saw one of the booklets in their testing session and were equally divided between the two lists, both as a whole and per group. The bilingual children who saw List 1 in the English treatment saw List 2 in the Greek treatment and vice versa.

Table 3. Experiment 1 Example Stimuli

Item	Prime	Question	Target
Experimental	Anne eats one of her favourite fruits in the evenings.	What's happening today?	She's eating an apple.
Comparison	Anne and her friend climb something in the forest every weekend.	What's happening today?	Anne's friend is climbing a tree.

Figure 1. Example pictures for Experiment 1: Experimental item (left) & comparison item (right)

2.1.3. Procedures

Having created or chosen the experimental materials and additional instruments, ethical clearance for the DPhil as a whole was obtained from the Central University Research Ethics Committee of the University of Oxford (Appendix D). Then, the battery was piloted with and administered to the target populations.

2.1.3.1 Obtaining Consent

Since Experiment 1 focused on children, their parents were the ones who had to provide consent. Thus, an Information Sheet and a Consent Form (Appendices E and F) were sent to the parents through the school. The consent forms adopted an opt-in protocol – in other words, parents had to sign and return the consent form to the school in order for children to take part. Having obtained consent from the parents, children were also asked whether they assented to take part prior to the administration of the tests and tasks.

2.1.3.2. Piloting

Having received ethical clearance for the DPhil, it was possible to pilot the experiments. Piloting the child battery was essential since the existing research with the experiment's target

population (Greek-English sequential bilingual children) is limited. A substantive pilot was conducted in 2018 in order to inform the design of the child study. The pilot focused on the same population (i.e., bilingual children in total and partial immersion programmes programmes) and linguistic structures as those that were tested in Experiments 1 and 2 (i.e., subject pronouns and subordinate markers). Its aim was to determine the best method for assessing children's linguistic production (contrasting a repetition task with an elicitation task), and to explore whether children's performance was affected by their linguistic knowledge.

The results of the pilot (which are presented in Appendix G) suggested that elicitation was the most reliable and valid method in the particular context and highlighted the role of both lexical and grammatical proficiency in shaping children's performance. However, children's overall performance on the elicitation task was low. Taking this finding into consideration, it was decided that the task should be made easier and more engaging so as to allow all participants, but especially those with lower exposure to English (i.e., children in partial immersion), to show their full potential. Moreover, it was decided to include comparison items so as to further explore children's production and/or measure their baseline performance.

These decisions led to changes in the materials for Experiment 1 (and 2). Yet, the new materials needed to be tried out before being put into use. Thus, a second (much smaller) round of piloting was conducted with 10 Greek preschool children with English as an Additional Language in the UK. These children were chosen because they resembled the target population in terms of age and L1, but not in terms of educational context; thus, testing these children allowed for getting a sense about the target population's performance in the various measures of the battery, without 'sacrificing' potential participants. Five of the children were tested in the English treatment and five were tested in the Greek treatment. The piloting indicated no additional problems with the linguistic proficiency tests or the language production tasks.

2.1.3.3. Administration

Testing started after the completion of the piloting. As the previous sections make clear, the child protocol was designed for face-to-face administration, and was administered twice for bilingual children, once in Greek and once in English, while Greek and English monolingual children completed the tasks and tests in their respective L1s.

The instructions were given in the children's L1 – that is, English for the English monolinguals, and Greek for the Greek monolinguals and the bilinguals. The measures were completed in the following order: the BPVS3/PPVT-R Greek was administered first, and was followed by Experiment 1 and TROG2/DVIQ. This order was chosen in order to keep children engaged. Since bilingual children had to complete the battery twice, it was decided that the Greek assessment would be administered first and the English assessment would follow (at least) two months later, so as to allow bilingual children to accumulate more exposure in English. To prevent familiarisation effects in this assessment, the order of the measures (i.e., the standardised tests and experimental tasks) was reversed.

The standardised tests were administered as explained in the tests' manuals, while the experimental tasks were introduced as games. For Experiment 1, children were told that the researcher will explain what the characters like to do 'every day', while they had to look at the pictures and tell the researcher what the characters were doing 'today'.

In the English treatment, administration of the BPVS3 and the TROG2 lasted 10 to 15 minutes depending on children's performance, and Experiment 1 five to ten minutes. In the Greek treatment, administration of the PPVT-R Greek lasted a maximum of 10 minutes, while the 'Comprehension of Morphosyntax' subtest of the DVIQ took a mere five minutes to complete. Of course, the experiment took the same time as in the English treatment.

The monolingual English and Greek children were tested in an unused classroom in their schools. The bilingual children were tested in a spare office and/or a storage room in their

schools for the Greek treatment (which was administered first), while the English treatment took place online, using the schools' virtual learning platforms due to COVID-19 restrictions that made visiting the schools in person impossible. During face-to-face administration, the experimenter was alone with the child – except in the partial immersion school, where a teacher was present in the room throughout the procedure, without intervening. During the online administration of the English assessment in the bilingual groups, a teacher had to be with the children throughout the procedure for children's safeguarding but also to help with technical issues (e.g., setting up the platform). The teachers did not interfere with the procedure.

Moreover, a certificate of completion was designed (Appendix H): it depicted the main characters of the two experiments holding balloons; after completing each of the activities, the children could put a stamp inside one of the balloons. Children wrote their name on the certificate, and received it as a token of appreciation at the end of the procedure. Stamping a certificate was not possible in the virtual leg of testing, so arrangements to give children a sticker were made with the teacher who facilitated the procedure in the two bilingual schools.

2.1.4. Preparations for Data Analysis

Preparing the data for data analysis involved two procedures: transcribing the audio recordings and calculating the participants' scores of standardised tests and questionnaires. The transcriptions of the audio-recordings focused on the children's target utterances and marked unintelligible segments, pauses and code-switching. As the data was collected in schools (or during a video call) the sound quality was rather poor; thus, 10% of the audio-recordings (5% in English and 5% in Greek) were blind-transcribed to ensure the accuracy of the transcriptions. There was 91.5% agreement between the transcribers (87.3% for the English transcripts and 95.7% for the Greek transcripts). The few items that caused disagreement were discussed with the independent transcriber, and a few small changes were made as a result.

Hence, the transcripts were coded. The coding scheme differed for the two linguistic treatments. In the English assessment, one point was awarded for the provision of an overt subject in the experimental and comparison conditions, while zero points were awarded for the omission of an overt subject. This coding scheme was lenient regarding the production of subject pronouns with incorrect morphological features. In particular, if children produced a subject pronoun with the wrong gender marking (that is, 'he' instead of 'she' and vice versa) or the wrong case marking ('her' instead of 'she'), they would still receive a point. This decision was taken since the focus of this research is to investigate subject realisation and not gender marking. Furthermore, children seem to be oblivious to their use of incongruent gender marking at the start of language acquisition (Eilers, Tiffin-Richards & Schroeder, 2018), which suggests that the acquisition of morphological gender might be a slow and difficult process that is not directly related to the acquisition of pronouns. In the Greek assessment, one point was awarded for the omission of an overt subject in the experimental condition (provided that the verb was marked for person and number to reflect the features of the omitted subject) and for the provision of an overt subject in the comparison condition, and zero points awarded otherwise. In both assessments and conditions therein, one-word utterances were excluded from the analyses. To ensure the reliability of the coding, 5% of the transcriptions were coded by two independent raters (an L1 English speaker for the English assessment and an L1 Greek speaker for the Greek assessment), using the above coding schemes. Interrater agreement was high for both sessions: 97% in the English assessment, and 92% in the Greek assessment.

To complete the preparations, the Greek and English receptive vocabulary and grammar tests had to be scored. The formula for calculating raw scores (as per the tests' manuals) was: the total number of items the child completed minus the total number of mistakes he/she made.

2.2. Results

In this section, I will present the results of the experiment[6]. I will first discuss the findings that emerged from the Greek and then the English assessment to address the first question. Hence, I will present the data collected as part of the vocabulary and grammar tests) that were used in both assessments in an effort to address the second question of the experiment.

2.2.1. Greek Assessment

In the Greek treatment, the bilingual children's performance was compared to that of the Greek monolingual controls. As shown in Figure 2, the performance of all groups in the experimental item condition (where the target structure was a null subject) was high: the Greek children scored 8.06 (SD=2.39), while the bilingual children scored 8.43 (SD=2.29). Turning to the comparison condition (where the target structure was an overt subject DP), both of the groups' scores were rather low. In particular, the Greek children scored 2.26 (SD=2.73), while the bilingual children scored 3.19 (SD=3.19). Figure 3 depicts the scores of the bilingual groups themselves. The total immersion group scored 8.23 (SD=2.71) and the partial immersion group scored 8.63 (SD=1.78) in the experimental condition; in the comparison condition, the former achieved a mean score of 3.63 (SD=3.23) and the latter achieved a score of 2.74 (SD=3.13).

After ensuring that the distribution of each group's scores was normal and establishing whether the variances of the groups were homogenous (see Appendix I for information on all the variables associated with the Greek assessment), it was possible to conduct two 2x2 ANOVAs, which allow examining within-participant variables (i.e., condition) alongside between-participant variables (i.e., group).

The first 2x2 ANOVA compared the monolingual Greek and the bilingual Greek-English children's performance. As predicted, the ANOVA indicated that the interaction

[6] The data can be accessed at: https://osf.io/r4zt7/.

between group and condition was not significant and represented a small effect [F(1,102)=.30, p=.59, η^2_p=.003], suggesting that the two groups' performance was not different (from a statistical perspective) in the two conditions. Also as predicted, there was a main and large effect of condition [F(1,102)=120.26, p<.001, η^2_p=.54], which indicates that there were differences between the two conditions, such that all children performed better in the experimental than in the comparison condition. In addition, contrary to predictions there was as a main and small effect of group: F(1,102)=6.15, p=.02, η^2_p=.06. However, an independent samples t-test that was conducted to explore this effect further suggested that there were no statistical differences between the Greek monolingual and the Greek-English bilingual children [for the experimental condition: t(102)=-.76, p=.45; for the control condition: t(102)=-1.44, p=.15]. Thus, the discovered main effect is likely an artefact of the analyses.

The second 2x2 ANOVA explored the performance of the bilingual groups in the Greek assessment. Once again as predicted, there was no group*condition interaction: F(1,68)=1.26, p=.27, η^2_p=.02. Likewise, there was a main and large effect of condition: F(1,68)=83.86, p<.001, η^2_p=.55. Finally, according to predictions, no main effect of group was detected in this set of analyses: F(1,68)=.53, p=.47, η^2_p=.01.

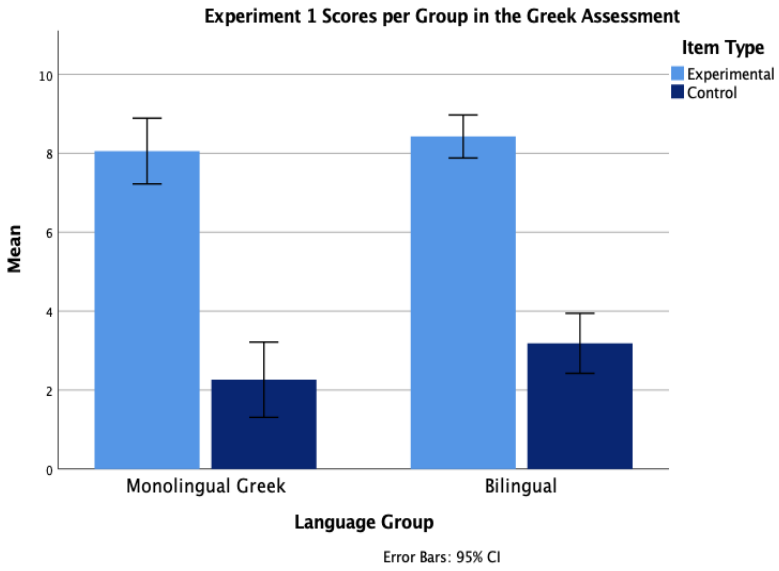


Figure 2. Monolingual and bilingual children's experimental and comparison item scores in the Greek assessment of Experiment 1.

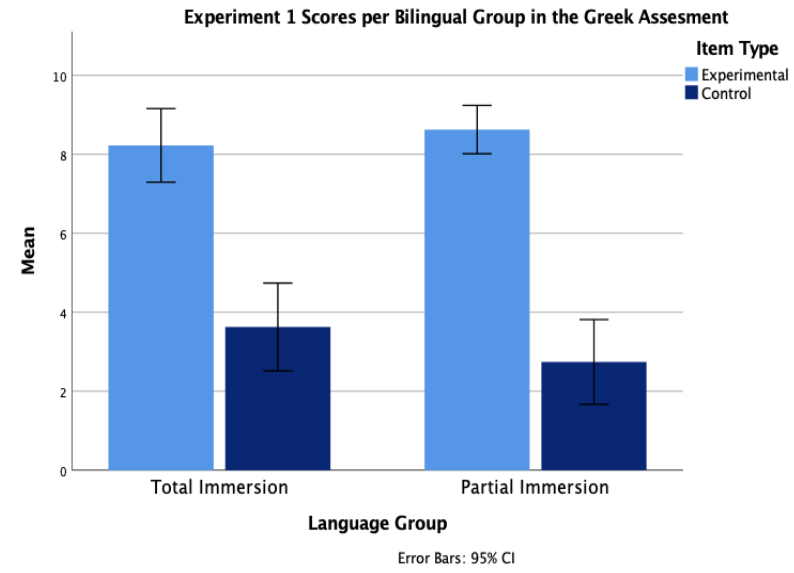


Figure 3. Total and partial immersion children's experimental and comparison item scores in the Greek assessment of Experiment 1.

Observing the discrepancies between all of the groups' high experimental item and low comparison item scores, a bivariate correlation was conducted to identify the relationship between children's performance in the two conditions. Figure 4 illustrates the results of the correlation, which revealed a strong, negative and significant relationship between children's performance in experimental and comparison items, $r=-.58$, $p<.001$. In other words, children who were using more null subjects in the experimental condition were also producing fewer overt subjects in the comparison condition. The correlation was two-tailed, suggesting that children who were using fewer null subjects in the experimental condition were also producing more overt subjects in the comparison condition but, also, that children who were using more overt subjects in the comparison condition were also producing fewer null subjects in the experimental condition. This finding indicates that children had a preference for a specific manifestation of the subject argument (whether null or overt), which they used without considering the rules that underpin the distribution of null and overt subjects in Greek (for similar results regarding the processing preferences of age-matched children with regards to null and overt subjects relative to adults, see Papadopoulou et al., 2014).

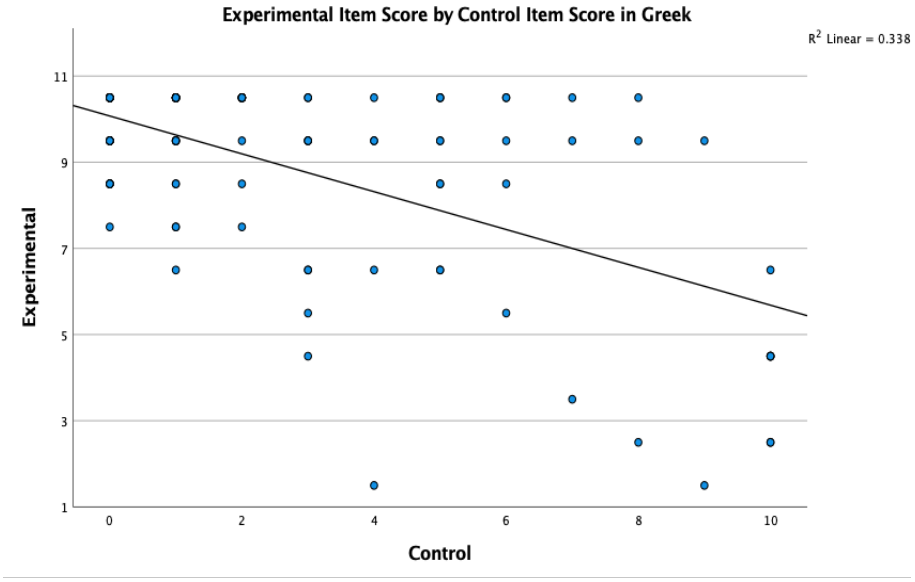


Figure 4. Correlation between the children's experimental and comparison item scores for the Greek assessment of Experiment 1.

2.2.2. English Assessment

In the English assessment, the bilingual children were compared to age-matched English monolingual controls. As shown in Figure 5, the English children achieved a mean score of 8.25 (SD=1.86) in the experimental item condition (where the target structure was an overt third person singular subject pronoun), while the bilingual children scored 4.20 (SD=2.94). Similar patterns were observed in the groups' comparison item performance (where the target was an overt subject DP): the mean score of the English group was high, 8.14 (SD=2.79), while the bilingual group was much lower: 3.90 (SD=3.01).

The groups' scores were found to be normally distributed (see Appendix J for the normality and homogeneity of variances for all the variables associated with the English assessment). Thus, it was possible to conduct a 2x2 ANOVA in order to compare the English monolingual and the Greek-English bilingual children in the two conditions. As expected, the ANOVA did not reveal an interaction between language group and condition [$F(1,104)=.24$, $p=.62$, $\eta^2_p=.002$], suggesting that the two groups performed alike in the experimental and the comparison condition of the experiment. Contrary to expectations, the analysis did not indicate a main effect of condition [$F(1,104)=1.15$, $p=.29$, $\eta^2_p=.01$], highlighting that the experimental

and comparison condition scores were comparable. Nevertheless, it did indicate a main and large effect of group [$F(1,104)=58.04$, $p<.001$, $\eta^2_p=.38$], according to which the two groups' scores differed significantly. A independent samples t-test that was conducted to explore this effect further highlighted that the English monolingual and the Greek-English bilingual children differed in both the experimental condition [$t(104)=7.52$, $p<.001$] and in the comparison condition [$t(104)=6.92$, $p<.001$].

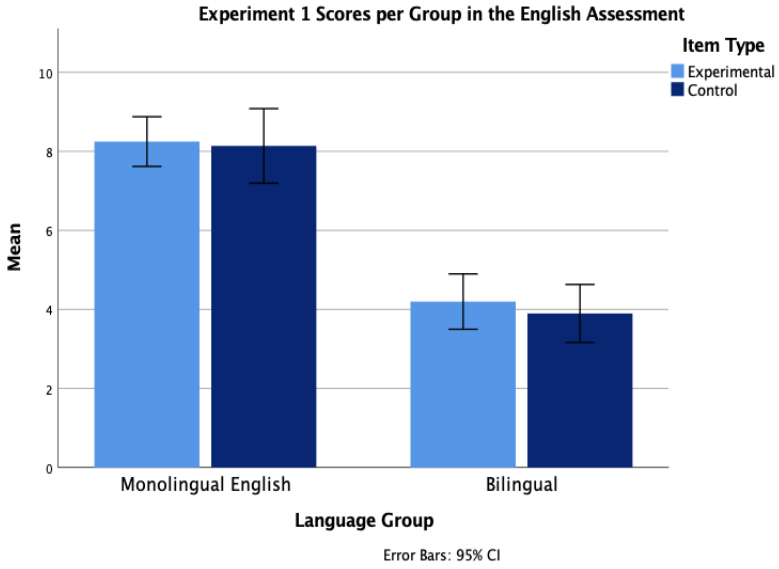


Figure 5. Monolingual and bilingual children's experimental and comparison item scores in the English assessment of Experiment 1.

Zooming in on the bilingual groups themselves, the total immersion group scored 4.91 (SD=2.58) in the experimental condition and 4.60 (SD=2.81) in the comparison condition, while the partial immersion group scored a bit lower across the board: 3.49 (SD=3.14) in the experimental condition and 3.20 (SD=3.12) in the comparison condition. Following a 2x2 ANOVA, no group*condition interaction was found: $F(1,68)=.004$, $p=.95$, $\eta^2_p=.000$. Moreover, the analysis did not indicate a main effect of condition: $F(1,68)=1.84$, $p=.18$, $\eta^2_p=.03$. However, it did reveal a main effect of group: $F(1,68)=4.47$, $p=.04$, $\eta^2_p=.55$. Two independent samples t-test with language group as the independent variable and experimental or comparison item scores as the dependent variables suggested that the children in total

immersion differed from children in partial immersion in the experimental condition [t(68)=2.08, p=.04] but not in the comparison condition [t(68)=1.94, p=.06].

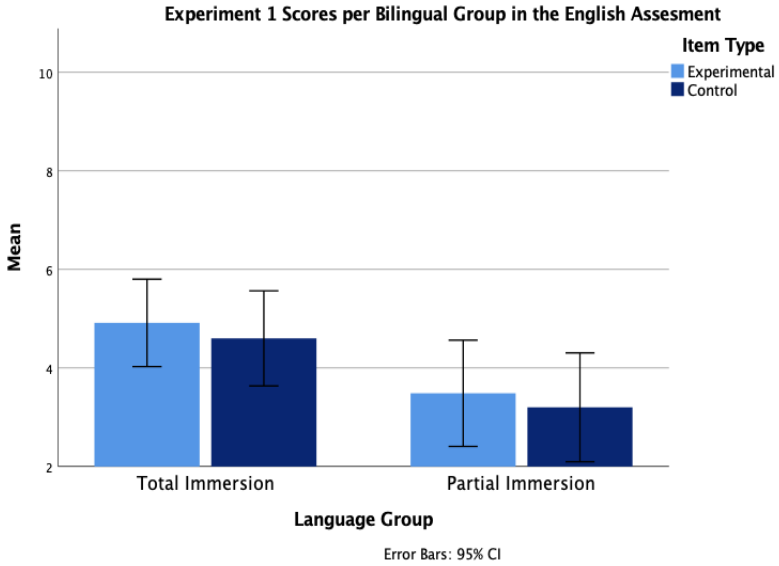


Figure 6. Total and partial immersion children's experimental and comparison item scores in the English assessment of Experiment 1.

Noting the similarities between the groups' experimental and comparison item scores, a two-tailed bivariate Pearson's correlation was run to establish the relationship between children's performance in the two conditions. As depicted in Figure 7, a very strong, positive and significant correlation was found between the two variables (r=.86, p<.001). This result suggests that children who achieved high scores in the experimental condition also achieved high scores in the comparison condition (and vice versa). Thus, contrary to the Greek assessment, it can be argued that children's performance in the English assessment was dependent on knowing the rules that determine the use of referential expressions in English: children who correctly used a subject pronoun in the experimental condition were more likely to correctly use an overt DP in the comparison condition.

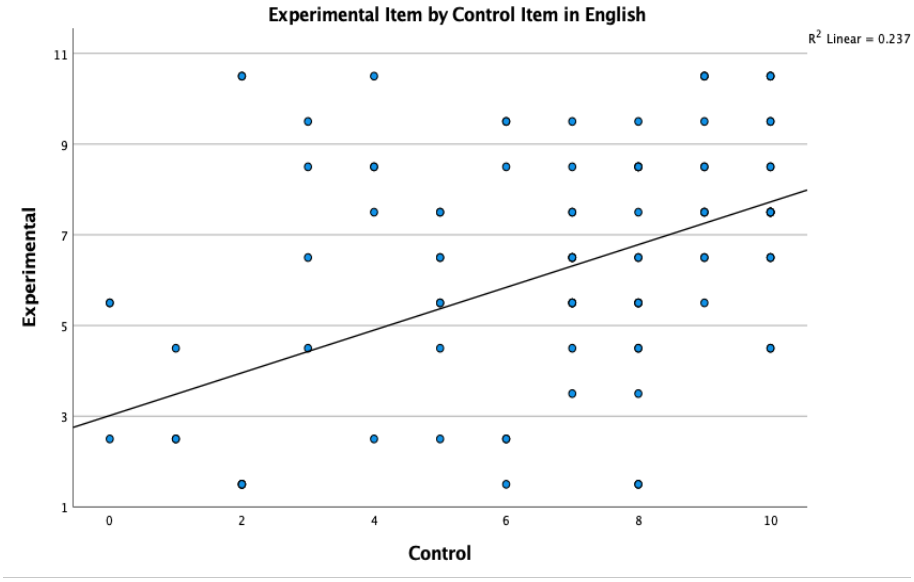


Figure 7. Correlation between the children's experimental and comparison item scores for the English assessment of Experiment 1.

2.2.3. Background Measures

Before discussing the effect(s) of proficiency on children's performance in the assessments, it is important to note the children's performance in the lexical and grammatical proficiency tests.

2.2.3.1. Greek Proficiency Scores

As shown in Figure 8, the Greek monolingual children scored 66.06 (SD=16.19) in the PPVT-R Greek, while the bilingual children in the total immersion group scored 46.66 (SD=14.53) and those in the partial immersion group scored 72.09 (SD=12.94). The three groups' performance on the test was compared using a One-Way ANOVA with the PPVT-R Greek raw scores as the dependent variable and the Language Group as the independent variable. The ANOVA was significant and represented a large effect: $F(2,101)=28.95$, $p<.001$, $\eta^2=.36$.

To identify the locus behind the observed effects, a series of Tukey HSD post-hoc tests were conducted. These revealed that the partial immersion group scored higher than the other two groups; however, the mean difference between partial immersion children and Greek monolingual children, 6.03, was not significant ($p=.21$), but the mean difference between

partial and total immersion children, 25.43, was: p<.001. Children in total immersion also differed from Greek monolinguals (mean difference=-19.40; p<.001).

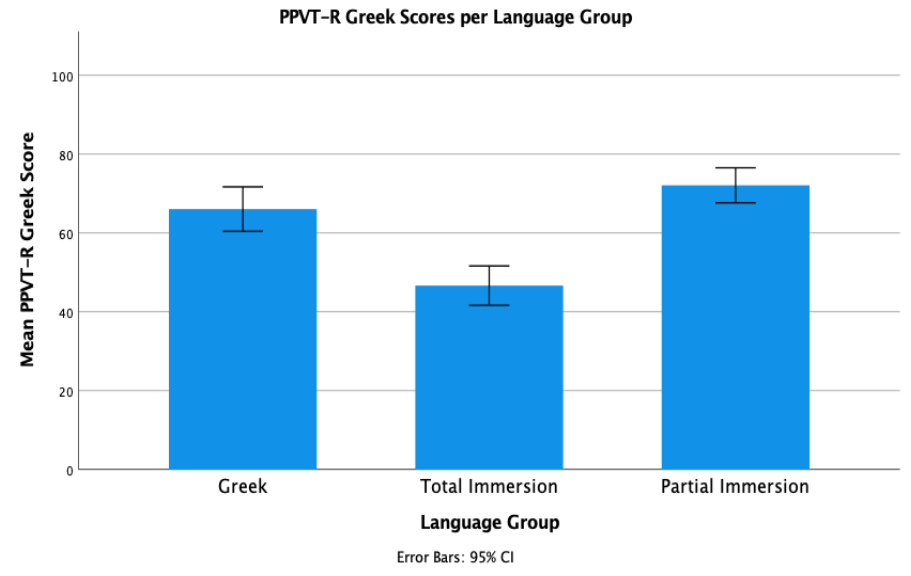


Figure 8. Children's Scores on the Greek vocabulary test.

As depicted in Figure 9, the picture is equally mixed with regards to the results of the DVIQ. The average raw score of the Greek group was 21.59 (SD=3.53), the average raw score of the total immersion group was 22.34 (SD=2.82) and that of the partial immersion group was 23.60 (SD=3.08). In this case, the One-Way ANOVA with DVIQ raw scores as the dependent variable and Language Group as the independent indicated a significant difference between the groups that corresponded to a medium effect: $F(2, 101)=3.59$, $p=.03$, $\eta^2=.06$.

The Tukey HSD post-hoc tests suggested that, once again, children in partial immersion performed better than their peers in the other two groups. This time, the mean difference between the immersion group and the Greek group, 2.01, was significant: p=.03. However, the mean difference between the immersion group and the total immersion group, 1.26, was not (p=.218). In addition, children in total immersion did not differ from Greek monolinguals (mean difference=.75; p=.58).

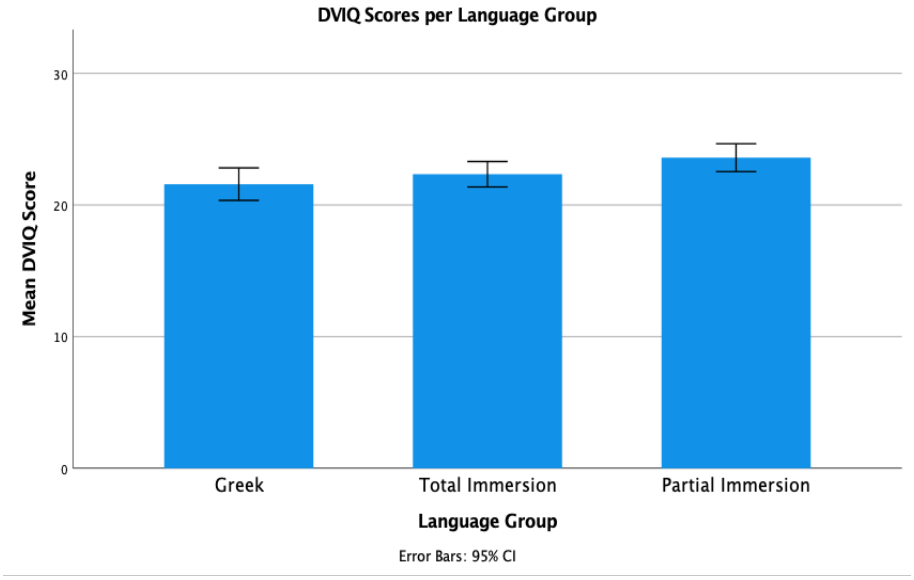


Figure 9. Children's Scores on the Greek grammar test.

2.2.3.2. Effects of Proficiency (& Age) in the Greek Assessment

How are children's vocabulary and grammar test scores related to their age, and their scores in the Greek assessment of Experiment 1? To answer this question, two bivariate Pearson's correlations were conducted: one for the Greek group (Table 4) and one for the bilingual group (Table 5). As expected, the children's performance on the Greek vocabulary and grammar tests was related: Figures 10 and 11 illustrate that there was a strong, positive and significant correlation between the two sets of scores for the Greek group, and a moderate but also positive and significant correlation for the bilingual groups.

Table 4. Correlations between age, proficiency and Experiment 1 scores (experimental and comparison) for the Greek children in the Greek assessment.

	Age	Vocabulary	Grammar	Experimental	Comparison
Age	-	.68***	.50***	.26	-.25
Vocabulary	-	-	.66***	.27	-.27
Grammar	-	-	-	.06	.05
Experimental	-	-	-	-	-.79***
Comparison	-	-	-	-	-

Table 5. Correlations between age, proficiency and Experiment 1 scores (experimental and comparison) for the bilingual children in the Greek assessment.

	Age	Vocabulary	Grammar	Experimental	Comparison
Age	-	.52***	.01	.02	-.25
Vocabulary	-	-	.37***	.08	-.11
Grammar	-	-	-	.10	.22
Experimental	-	-	-	-	-.52***
Comparison	-	-	-	-	-

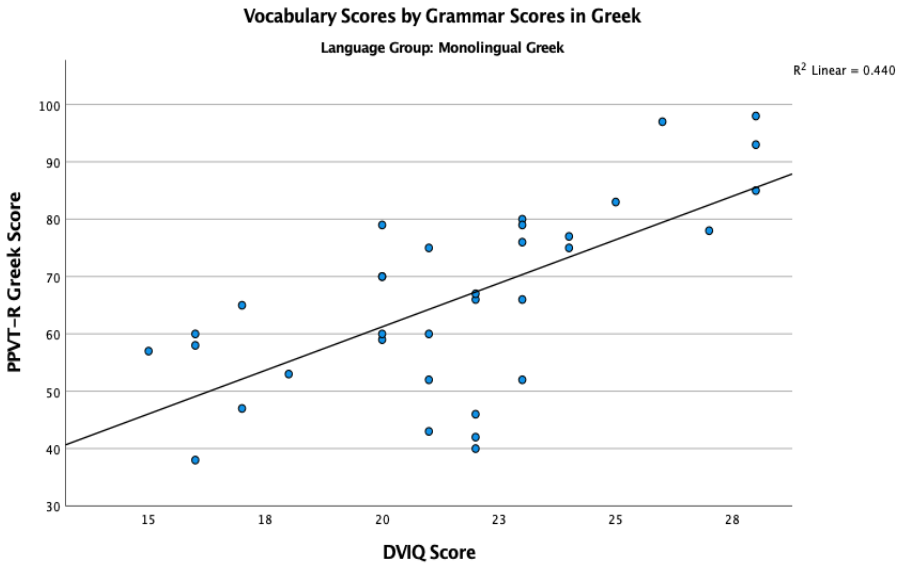


Figure 10. Correlation between the Greek children's vocabulary and grammar scores in Greek.

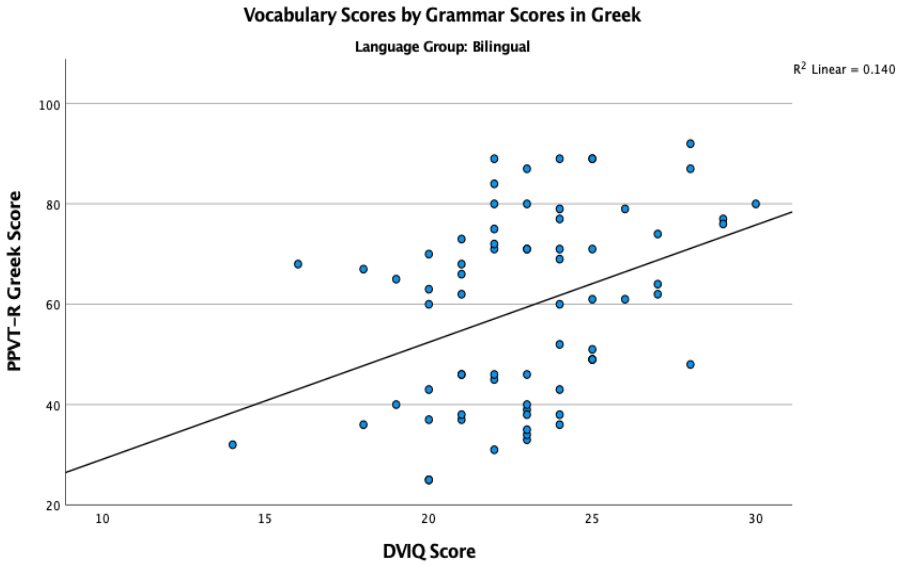


Figure 11. Correlation between the bilingual children's vocabulary and grammar scores in Greek.

Also as expected, age was found to be related to children's performance in the vocabulary and grammar tests: for the Greek group, there was a strong, positive and significant correlation between age and PPVT-R Greek scores, shown in Figure 12; there was also a strong, positive and significant correlation between age and DVIQ scores, depicted in Figure 13.

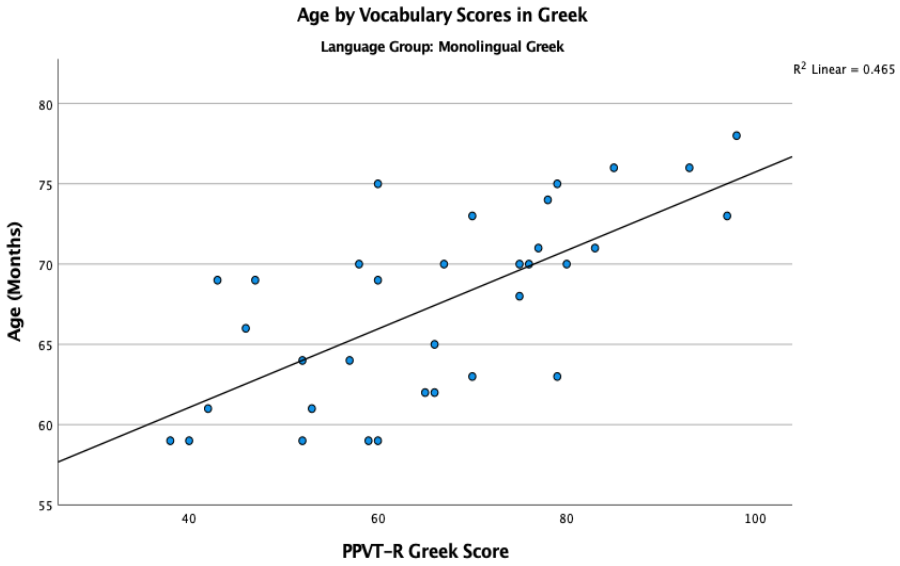


Figure 12. Correlation between the Greek children's age and vocabulary scores in Greek.

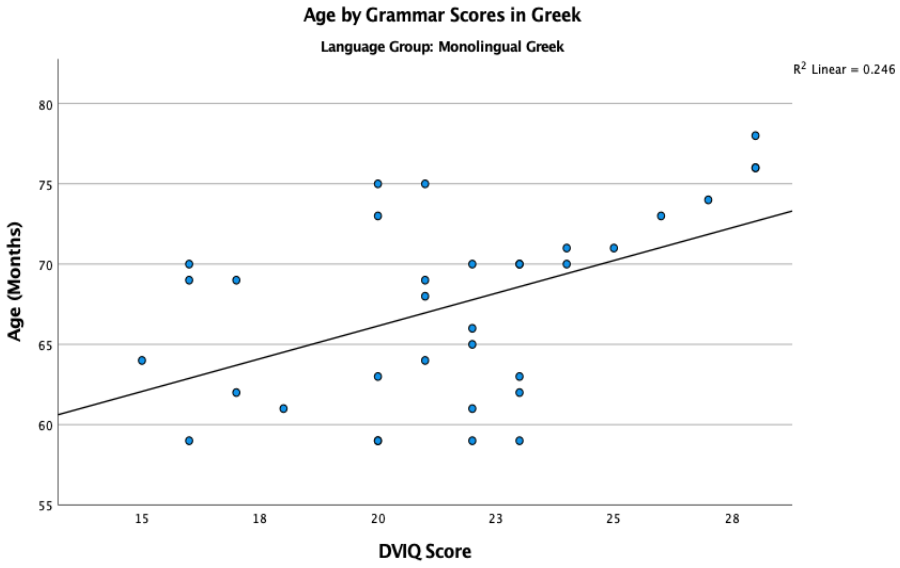


Figure 13. Correlation between the Greek children's age and grammar scores in Greek.

For the bilingual children there was only one strong, positive and significant correlation observed, that is, between age and PPVT-R Greek scores, visualised in Figure 14.

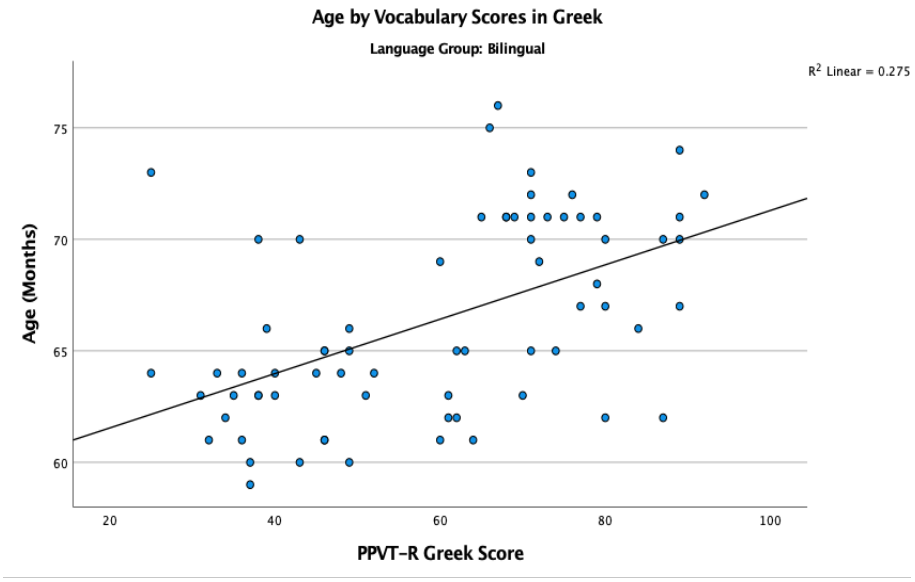


Figure 14. Correlation between the bilingual children's age and vocabulary scores in Greek.

2.2.3.3. English Proficiency Scores

In the English treatment, the English monolinguals outperformed the two groups of Greek-English bilinguals in both tests. With regards to BPVS3, and as illustrated in Figure 15, monolingual children scored 65.28 (SD=13.01), while bilingual children in total immersion scored 46.09 (SD=19.67) and bilingual children in partial immersion scored 27.86 (SD=12.06). A One-Way ANOVA with children's raw scores as the dependent variable and Language Group as the independent was significant, and corresponded to a very large effect: $F(2,103)=53.31$, $p<.001$, $\eta^2=.51$.

A series of Dunnett's T3 post-hoc tests revealed that these effects stemmed from differences between all three groups. In particular, English monolingual children performed significantly higher than children in total immersion (mean difference=19.19; $p<.001$) and better than children in partial immersion (mean difference=37.42; $p<.001$). Moreover, children in total immersion performed significantly better than children in partial immersion (mean difference=18.23; $p<.001$).

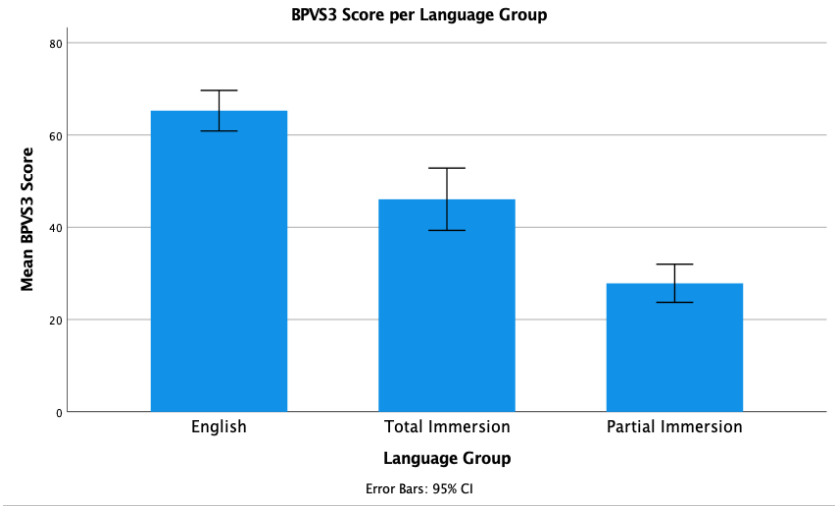


Figure 15. Children's scores on the English vocabulary test.

The same patterns can be observed in the TROG2 results. As depicted in Figure 16, the mean raw score of English monolingual children on this test was 50.19 (SD=12.57), while the mean raw score of total immersion children was 24.26 (SD=9.74) and that of partial immersion children was 14.57 (SD=5.38). A One-Way ANOVA revealed that the difference between the groups was significant and represented a very large effect: $F(2,103)=127.74$, $p<.001$. $\eta^2=.71$.

The Dunnett T3 post-hoc tests that were subsequently carried out revealed that the mean difference between the English group and the total immersion group (i.e., 25.94), between the English group and the partial immersion group (35.62), and between the total immersion group and the immersion group, 9.69, was also significant ($p<.001$).

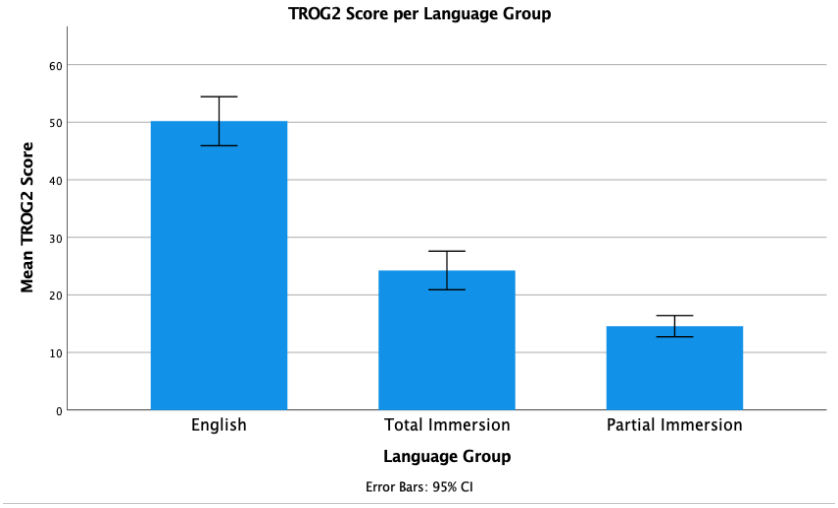


Figure 16. Children's scores on the English grammar test.

2.2.3.4. Effects of Proficiency (& Age) in the English Assessment

To explore the relationships between children's vocabulary and grammar, their age, and their scores in the English assessment of Experiment 1, two more bivariate Pearson's correlations were conducted: one for the English group (Table 6) and one for the bilingual group (Table 7). These revealed strong, positive and significant correlations between the vocabulary and grammar scores for both the English group and for the bilingual group, as can be seen in Figures 17 and 18 respectively.

Table 6. Correlations between age, proficiency and Experiment 1 scores (experimental and comparison) for the English children in the English assessment.

	Age	Vocabulary	Grammar	Experimental	Comparison
Age	-	.85***	.76***	.02	.28
Vocabulary	-	-	.68***	-.01	.18
Grammar	-	-	-	.37*	.48**
Experimental	-	-	-	-	.73***
Comparison	-	-	-	-	-

Table 7. Correlations between age, proficiency and Experiment 1 scores (experimental and comparison) for the bilingual children in the English assessment.

	Age	Vocabulary	Grammar	Experimental	Comparison
Age	-	-.14	-.09	-.15	-.12
Vocabulary	-	-	.80***	.42***	.38***
Grammar	-	-	-	.36***	.31**
Experimental	-	-	-	-	.85***
Comparison	-	-	-	-	-

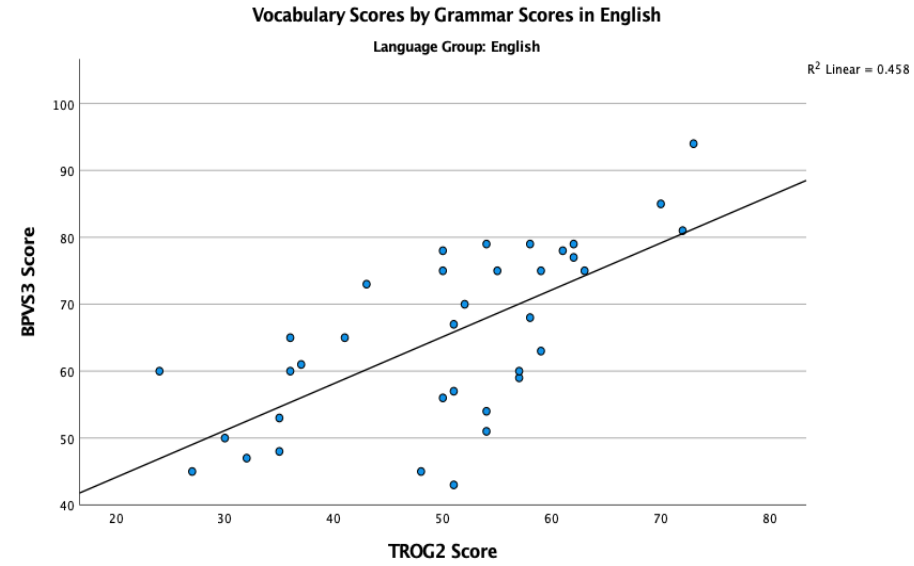


Figure 17. Correlation between the English children's vocabulary and grammar scores in English.

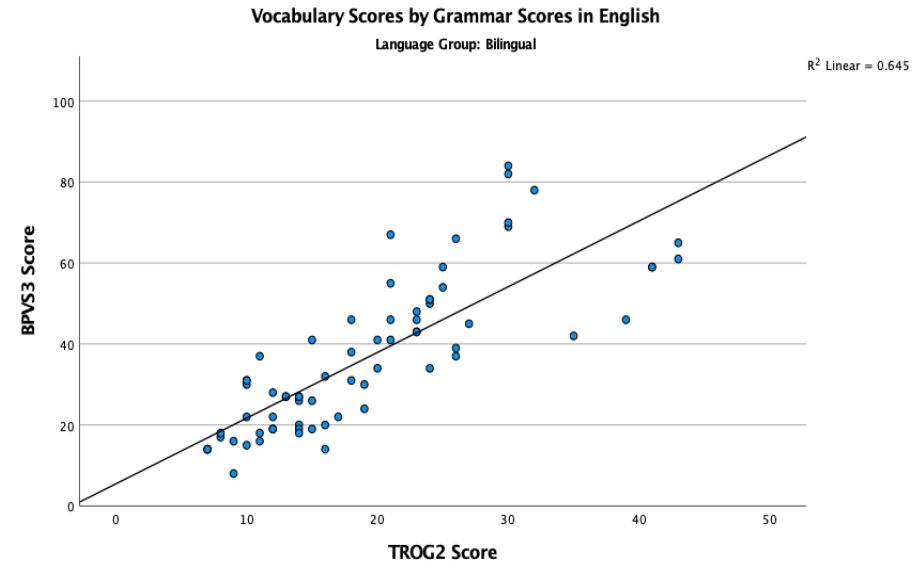


Figure 18. Correlation between the bilingual children's vocabulary and grammar scores in English.

The children's proficiency was found to be related to their performance in the experiment. For the English group, grammar scores had a moderate, positive and significant correlation with experimental and comparison item scores, as presented in Figure 19. For the bilingual group, there were moderate, positive and significant correlations between the children's vocabulary and the experimental and comparison item scores (see Figure 20), as well as between their grammar and their experimental and comparison item scores (see Figure 21).

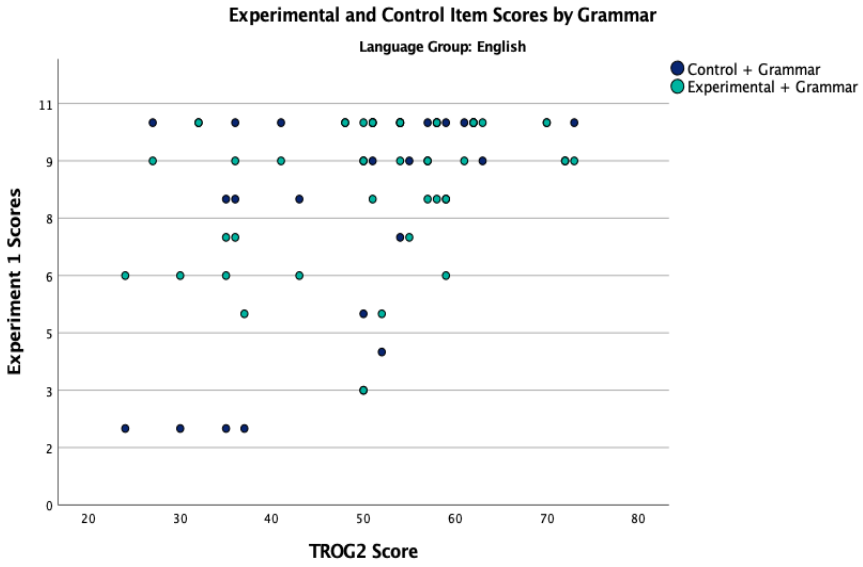


Figure 19. Correlation between the English children's experimental and comparison item scores and their grammar scores in English.

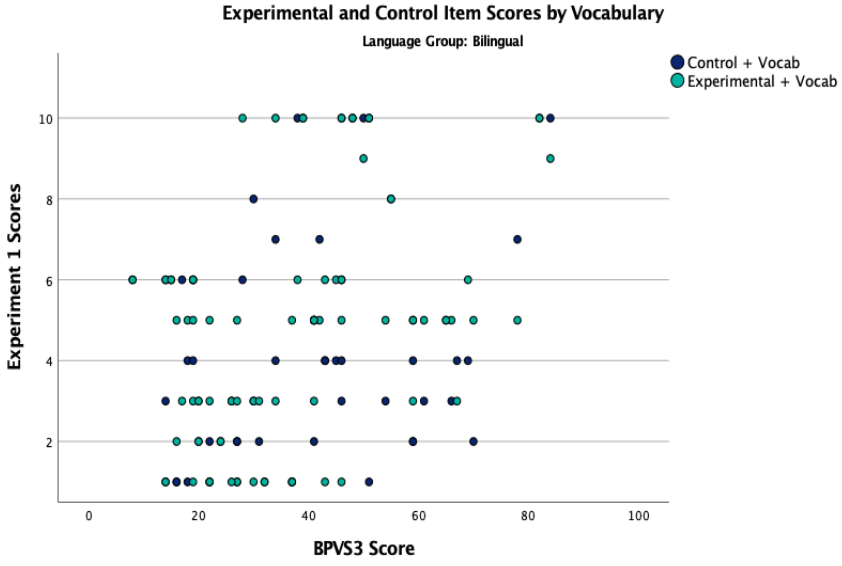


Figure 20. Correlation between the bilingual children's experimental and comparison item scores and their vocabulary scores in English.

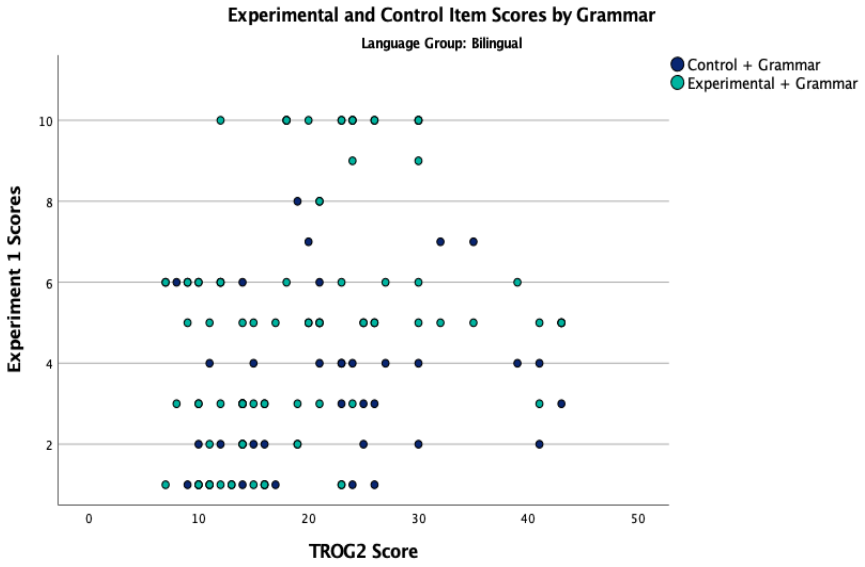


Figure 21. Correlation between the bilingual children's experimental and comparison item scores and their grammar scores in English.

The effect of age was limited to the English group, where the variable was found to have a strong, positive and significant correlations with vocabulary and grammar (these correlations are depicted in Figures 22 and 23).

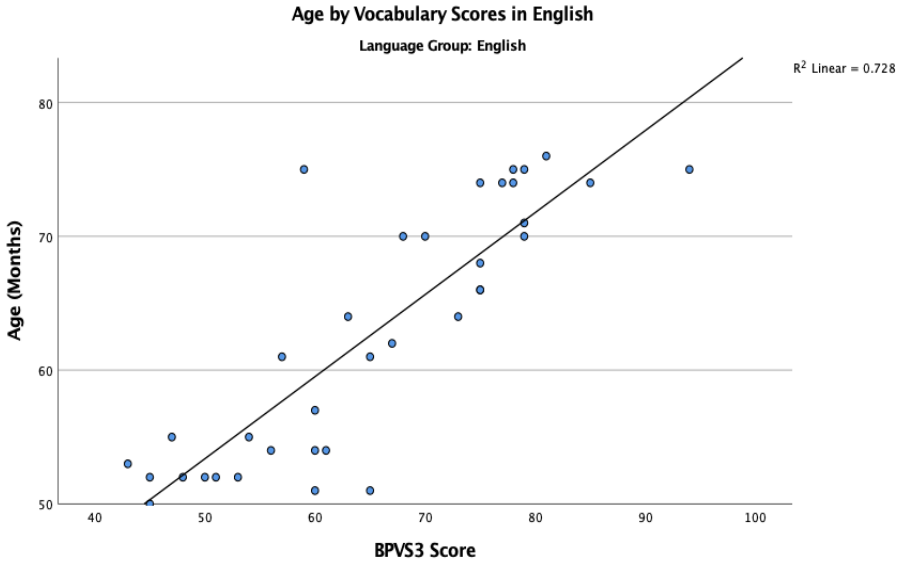


Figure 22. Correlation between the English children's age and vocabulary scores in English.

Figure 23. Correlation between the English children's age and grammar scores in English.

To explore the role that proficiency, as well as age, played in determining the bilingual children's experimental item scores in the English assessment further, a hierarchical linear regression analysis was performed. The same three models were tested: the predictor used in the first model was children's vocabulary, the predictors used in the second model were vocabulary and grammar, while the third model also took children's age into account. The results of the regression (presented in Table 8 below) suggested that all three models were a good fit and revealed vocabulary to be the sole significant predictor of children's performance.

In particular, Model 1 accounted for 17.5% of the variance [R^2=.18, F(1,68)=14.39, p<.001], with vocabulary having a significant effect on the bilingual children's experimental item scores. Model 2 accounted for an additional 0.2% of the variance [R^2=.177, F(1,67)=7.19, p=.001], but did not reveal vocabulary or grammar to be significant predictors of children's experimental item scores. Finally, Model 3 accounted for 18.6% of the variance [R^2=.19, F(1,66)=5.03, p=.003] but, once again, revealed none of the independent variables included in the model to be significant predictors[7].

[7] It should be noted that no collinearity was noted in either model. In particular, for Model 2, the VIF was 1.02; for Model 3, the VIF for age was 1.02, for vocabulary 2.85 and for grammar 2.82. The values are under the acceptable threshold (Menard, 2001).

Table 8. Predictors of Experiment 1 Scores in English

	Model 1			Model 2			Model 3		
Variable	B	*SE* B	β	B	*SE* B	β	B	*SE* B	β
Vocabulary	.07	.02	.42***	.06	.03	.36	.05	.03	.34
Grammar				.02	.06	.08	.03	.06	.08
Age							-.07	.08	-.10
R^2			.18			.18			.19
F Change in R^2			14.39***			.16			.75

2.3. Discussion

The analyses presented in the previous section, can provide answers to the experiment's questions. Question (1a) was: 'how does the bilingual children's use of subject pronouns in the two languages differ from that of their monolingual peers?'. Following the input quantity account of CLI that postulates the transfer of grammatical properties from the children's dominant language (in this case, Greek) to their non-dominant language (in this case, English), it was predicted that the bilingual children will use fewer overt subjects in English than the English monolingual children but a comparable number of subjects relative to the Greek monolingual children. This prediction was borne out in the findings: the bilingual children omitted the subject argument in both the experimental item condition (which required the use of an overt third person singular subject pronoun) and the comparison item condition (which required the use of an overt subject DP) more often than their English monolingual peers.

It should be noted that, despite the fact that overt subjects are required in English, English monolingual children also produced sentences with null subjects 14.4% of the time in the experimental condition. The percentage depended on specific participants, suggesting that some of the (younger) monolingual children who formed part of the sample had not yet acquired the subject realisation rules of their language. This is not surprising, given the fact

that monolingual children are known to omit the subject argument from their utterances at the start of language development, even if their L1 is a non-null-subject language, like English (e.g., Bloom, 1990). This phase often passes as children get older and acquire the rules that underpin the realisation and distribution of subject arguments in their language. However, the omissions of subjects by the English monolingual children could blur the present results – an issue that will be revisited and addressed in Chapter 4 of the thesis.

With regards to question (1b), 'how does the bilingual children's use of subject pronouns in the two languages differ depending on whether they attend the total immersion or the partial immersion preschool programme?', it was speculated that the bilingual children would perform in an input-based gradient in English, such that children in total immersion (who receive all of their school input in English) would manifest smaller differences from the English monolinguals than the children in partial immersion (who receive half of their school input in English). This prediction was borne out in the experimental condition of the English assessment (but not in the comparison condition).

This finding hints that the amount of input bilingual children hear in each of their languages has an effect on the direction of CLI, while other factors like proficiency and/or age are important in informing the magnitude of CLI effects (see De Cat, 2020). Nevertheless, it is important to highlight that the present experiment operationalised input quantity in a categorical fashion (capitalising on the natural groups that exist in the Greek educational landscape). A less crude measure of input quantity, such as a language exposure questionnaire (e.g., De Cat, 2020; Paradis, 2011; Unsworth, 2015), might allow nuanced effects on the individual level to be captured.

In relation to these points, the answer to question (2) of the experiment, 'how does the bilingual children's linguistic proficiency in Greek and English affect their use of the target structure in the two languages?' is pertinent. Given previous CLI research, it was expected that

children with higher linguistic proficiency would be less affected by transfer and, thus, omit sentential subjects at a lower rate in the English assessment. The correlation analyses conducted in response to this question corroborated this prediction: while proficiency did not appear to be related to (or predictive of) bilingual children's performance in the Greek assessment, it did play a role in their performance in the English assessment: both vocabulary and grammar were correlated with Experiment 1 scores, such that the better the children performed at the BPVS3 and the TROG2 the more accurate they were in providing the (appropriate) subject argument in the experimental condition. Thus, it can be argued that the quantity of input that children hear modulates the direction of CLI, but their proficiency affects the magnitude of the effect, such that more proficient children (in this case, especially in terms of vocabulary) are more resistant to L1 transfer.

Moreover, vocabulary was found to be a significant predictor of bilingual children's performance in English, accounting for 17.5% of the variance in their scores. The fact that vocabulary predicts the children's use of (null or overt) sentential subjects (and, in particular, the third person singular subject pronouns 'he' and 'she' that were tested in the experimental condition of Experiment 1) could suggest that children acquire subject pronouns in a lexical fashion. At face value, this speculative idea appears to echo usage-based frameworks of language acquisition (e.g., Tomasello, 2003), but requires further (well-powered) research in light of the fact that there were also high correlations between children's proficiency scores and their performance in the comparison condition which captured the children's referential choice, as opposed to their use of null and overt subjects).

In sum, the findings of Experiment 1 suggest that CLI in this group of Greek-English sequential bilingual children takes the form of transfer from the children's (dominant) L1, Greek, to their (non-dominant) L2, English. As such, it manifests itself as the erroneous overuse of null subjects in English. This finding is in line with the input quantity account of CLI.

However, the findings are not in line with (most) structural linguistic accounts of CLI. For instance, the Subset-Superset Condition (Hulk & Müller, 2000) would predict transfer from the 'unambiguous' language (English) to the 'ambiguous'[8] language (Greek); likewise, the Interface Hypothesis (and frameworks therein) would predict transfer from the most economical language, which is English, to the less economical language, Greek (Sorace, 2011). This pattern is not evident in the children's linguistic behaviour. Instead of stipulating that the Greek-English sequential bilingual children tested in this experiment are not abiding by (universal) economy rules, it seems more appropriate to suggest that sequential bilingual children might be different from simultaneous bilingual children and more similar to L2 learning adults. Indeed, the few studies on CLI conducted with sequential bilingual children highlight the importance of input in determining the nature of CLI effects (e.g., Daskalaki et al., 2019; Paradis & Blom, 2015; Zdorenko & Paradis, 2008) and emphasise the need for additional research within this ever-growing population.

[8] Of course, English can be ambiguous as well as far as pronominal reference is concerned: for example, consider the sentence: "Alex talked to Matt while he was doing the washing-up". The term 'unambiguous' used to describe English above (and the corresponding term 'ambiguous' to describe Greek) draw from Hulk and Müller's (2000) Subset-Superset Account, under which unambiguous languages provide one manifestation of a structure, and thus form a subset, while ambiguous languages provide more than one manifestation of the same structure, and thus form a superset.

Chapter 3: Bilingual Children's Use of Subordinate Clause Markers

Experiment 1 (discussed in the previous chapter) focused on the use of sentential subjects. It revealed that the Greek-English sequential bilingual children who participated in the task omitted more sentential subjects –whether overt third person singular subject pronouns or subject Determiner Phrases (DPs)– from their utterances relative to age-matched English monolingual children. However, they produced a similar number of sentential subjects as age-matched Greek monolingual children. These patterns were attributed to crosslinguistic influence (CLI) from the children's dominant first language (L1), Greek, to their less dominant second language (L2), English. CLI is known to appear as a function of linguistic structure, leading to 'structural linguistic accounts' of the phenomenon (e.g., Hulk & Müller, 2000; Sorace, 2006), and/or as a function of linguistic dominance, leading to 'input quantity accounts' of the phenomenon (e.g., Yip & Matthews, 2000). Although the two (sets of) accounts are not mutually exclusive (Argyri & Sorace, 2007; Sorace et al., 2009), they give rise to different predictions about the conditions that make CLI appear. Given the findings of Experiment 1, input quantity accounts are believed to provide a better explanation of the Greek-English sequential bilingual children's use of sentential subjects – in line with previous literature on sequential bilingual children's use of subjects (e.g., Daskalaki et al., 2019).

Sentential subjects are a known locus of CLI, as they are regulated in the mental interface between syntax and pragmatics (Serratrice, 2013). As such, they are often the focus of studies that investigate transfer among both sequential bilingual and L2 learners (Blom, Cornips & Schaeffer, 2017). In turn, these tend to reveal the presence of CLI effects. However, to ascertain the causes behind effects that are nominally attributed to transfer from the L1 or the L2, it is important to assess the same learners use of 'control' structures: these would be structures that are not prone to CLI and/or are not regulated in the syntax-pragmatics interface. A comparison between the learners' use of CLI-prone and non-CLI prone structures would

allow for (a) clarifying whether the observed effects are indeed the result of CLI and (b) confirming the stipulated vulnerability of interface domains.

This assessment is of particular importance in view of the findings of Experiment 1 since the bilingual children were found to omit (rather than produce) a structure. Omissions are frequent in the earliest stages of L1 acquisition (Bloom, 1990) but, also, in L2 acquisition (Slabakova, 2008). Eilers and Oller (1976) proposed that these omissions are due to children being caught in a processing bottleneck: forming a sentence at the start of L1 and L2 acquisition (in the L1 and the L2 respectively) requires a substantial amount of processing resources for remembering and combining words in strings; as a result, the learners' struggle and leave some of the words they intend to use out of their utterances. Omission errors of this kind tend to be developmental in nature, with L1 speakers and L2 learners alike becoming better at processing and producing (strings of) words as their age and/or linguistic proficiency increases.

Under this light, it is important to identify whether the omissions observed in Experiment 1 represent a CLI effect. To this end, Experiment 2 was designed to assess sequential bilingual children's use of a 'control' structure: subordinate clause markers and, in particular, the marker 'to'. This structure is not CLI-prone, as it only involves a syntactic component. Moreover, it appears in both English and Greek. Indeed, the two languages require specific markers to head subordinate clauses: the English 'to' heads non-finite clauses while the Greek equivalent, 'na', heads subjunctive clauses. 'Na' is a more versatile structure than 'to': it can head main interrogative and imperative clauses as illustrated in the Greek examples (1) and (2), while 'to' cannot. Moreover, as 'na' heads subjunctive clauses it can appear with inflected verbs (in the subjective mood) while 'to', which heads non-finite clauses can only appear with non-inflected verbs, as the English examples in (3) and the Greek example in (4) (first presented in Chapter 1, but repeated here for convenience) demonstrate. These examples

also highlight that the markers 'to' and 'na' must be present to license a subordinate clause and make the overall sentence grammatical.

1. Na pao?

   To go.1SG.INTER

   'Can I go?'

2. Na pas

   To go.2SG.IMP

   'You can go'

3. He wants *(to) see a film

4. Theli *(na) dhi mia tenia

   Wants to sees.SUBJ a.FEM film

L1 English children are known to use the marker 'to' consistently from 3;6 (Kirjavainen & Theakston, 2011; Kirjavainen, et al., 2017; Kirjavainen, et al., 2009). However, there is a period of time during which L1 English children omit 'to' from their non-finite subordinate clauses. Indeed, Kirjavainen et al. (2017) studied the use of infinitival 'to' by 75 L1 English children between 3;0 and 4;0. The researchers presented children with pictures, described using both target verb+to+verb and the alternative verb+DP strands (that is, 'I want to eat' vs 'I want food'). Children were asked to complete a sentence which ended on the intentional verb 'want', thus legitimising both strands. Younger children omitted 'to' more often in general, but were more likely to produce it following verb+to+verb primes – an effect that the researchers interpreted as an effect of structural priming that, in turn, results from children's tacit knowledge of structure. While research on L1 Greek children's acquisition of 'na' is limited, there is some evidence that they use the marker 'na' consistently from around 3;0, following a period of illicit 'na' omissions from complement clauses (Mastropavlou & Tsimpli, 2010).

Thus, L1 English and L1 Greek children master the structure (in its subordinate clause marker capacity) at more or less the same age.

Given its aims and the issues discussed above, Experiment 2 sought to answer the following questions (echoing, to some extent, Experiment 1):

1) Does the bilingual children's use of subordinate clause markers in English and Greek differ from that of their monolingual peers?
2) Do the children in total immersion perform differently from children in partial immersion?
3) Does the proficiency in English influence their performance and, if so, in what way?

With regards to question (1), no CLI-driven effects are expected to appear in the bilingual children's use of the subordinate clause marker 'to' in English or 'na' in Greek, since the distribution of the structure(s) is similar (at least) at the surface level of the two languages. Thus, the bilingual children are expected to produce (and/or omit) the structure at a similar rate to both their English and their Greek monolingual peers, meaning that there will be no main effect of group. Turning to question (2), which explores if children's omissions result from a processing limitation, a main effect of group is expected; in other words, there should be differences between children in total immersion and children in partial immersion in the English assessment, such that the former (who receive more input in English and, thus, are likely to be more proficient in the language) should omit the structure less often. Of course, no differences are expected in the Greek assessment, since both groups of bilingual children are dominant in Greek. Likewise, as far as (3) is concerned, increased lexical and grammatical knowledge is predicted to improve children's performance and reduce their omissions of the structure. However, it should be noted that proficiency is expected to play a role in the event that there are differences between the groups' performance.

The aforementioned predictions also hold in the control condition of Experiment 2 which, as will be discussed later on, tested children's production of verbs in finite constructions

(i.e., without the marker 'to'/'na'). With regards to question (1), I do not predict a main effect of group in the English or the Greek assessment. As far as (2) is concerned, I expect a main effect of group in the English assessment, such that children in total immersion are more accurate than children in partial immersion, but comparable in the Greek assessment (where no differences between the groups are predicted). With regards to (3), I expect proficiency to affect the children's performance in the English assessment (where the differences between the groups should be more pronounced). In addition to the group effects, I predict that there will be a main effect of condition, such that in all children (regardless of their linguistic status) are more accurate in the control condition in both languages, as the requirement for producing 'to'/'na' is expected to make the experimental condition more difficult. Finally, I do not expect an interaction between group and condition.

3.1. Methods

In this section I will present the methods of Experiment 2. As the participants were the same as those who completed Experiment 1, I will not present their characteristics (or the process through which they were recruited) in great detail. I will, however, outline the design, procedures and preparations for data analysis that Experiment 2 involved. It should be noted that Experiment 2 took place as part of the battery that included Experiment 1 and the standardised tests of vocabulary and grammar.

3.1.1. Participants

The sample consisted of the same 140 children that took part in Experiment 1. Of these, 70 formed part of the Greek-English sequential bilingual group (mean age: 5;4). In particular, 35 attended a total immersion programme (where all of the tuition takes place in English) and 35 attended a partial immersion programme (where the tuition takes place in both English and Greek). The bilingual children were compared to 36 English monolingual children (mean age: 5;1) and 34 Greek monolingual children (mean age: 5;0). All the children who took part in the

experiment were between 4;0 and 6;0. The bilingual children had Greek as their only L1, spoke Greek at home, and learned English through school from 3;0 onwards.

Prior to the administration of the experiment, a power analysis was conducted based on the results of the same pilot study that was mentioned in Chapter 2, and using the G*Power 3 programme (Faul et al., 2007). Taking stock of the large effects that resulted from the pilot participants' use of subordinate clause marker 'to', the power analysis indicated that a mere 13 participants per group would suffice in order to achieve 95% power at the α=.05 level [(1-β)=.96] for main effects. Therefore, the number of participants that ended up being included in each of the groups in Experiment 2 was more than sufficient for it to be well-powered for the group comparisons[9].

3.1.2. Design

The dependent variable under investigation was the number of (produced or omitted) subordinate clause markers in an experimental language production task (in particular, an explicit correction task) that required participants to 'fix' the description of a picture using a target structure. To ensure that children would make use of the target structure (instead of an easier want+noun combination, see: Kirjavainen et al., 2017), the task involved structural priming. As an experimental method, priming capitalises on the tacit human tendency to repeat the same structures as an interlocutor during the discourse and, therefore, is effective for 'inducing' the production of less frequent and/or more difficult linguistic structures (Faitaki & Murphy, 2020). The Experiment 2 participants were primed with the target structure 'want+to+verb', and were expected to (re)produce it in their utterances.

[9] The power analysis was required for a preregistration that involved an earlier version Experiment 2 (and was also based on findings from Experiment 1). As both experiments had to be adapted to COVID-19 times, the proposed design and analysis was not viable; thus, the preregistration was withdrawn from the Open Science Framework (OSF).

The within-subjects variables were: the linguistic treatment (i.e., English or Greek) and the condition (experimental vs. control). The between-subjects variables were: the language group (i.e., monolingual English, monolingual Greek, children in total immersion, children in partial immersion)[10] and the children's lexical and grammatical proficiency, measured through tests of receptive vocabulary and grammar in English and Greek (for information about the children's performance in these tests, please refer back to Chapter 2).

The stimuli for Experiment 2 consisted of 33 items: 10 experimental, 10 control, 10 filler and three practice items (of which one resembled the experimental items, one the control items and one the filler items) (Appendix K). While the experimental items required children to use 'to' followed by a non-finite verb, the control items involved the production of a finite verb alone. The filler items were similar to the experimental and control items in terms of structure: five had the same format as experimental items and the other five had the same format as control items. Furthermore, each item was presented alongside a matching picture (Appendix L).

In both English and Greek, the items were between 13 and 14 words long, and all the target utterances were expected to be between four and seven words long. The English items were created using the 50 most frequent in children's input according to Theakston et al's (2001) Manchester Corpus on the CHILDES database, while all other words in the items were checked against Kuperman et al's (2012) age-of-acquisition norms. The English items were translated into Greek, and their developmental and linguistic appropriateness was assessed by the same two teachers (one from the Greek monolingual preschool) that reviewed the items of Experiment 1.

[10] As discussed in Chapter 2, examining the performance of children in total immersion and partial immersion relative to that of the English and Greek monolingual controls in the two assessments was designed to allow for an investigation of input quantity. In other words, input quantity was treated as a categorical predictor with three levels in each treatment: monolingual, total immersion and partial immersion.

All Experiment 2 items concerned two characters, John and Mary, as well as their friends and relatives. The items were composed of a 'context' sentence, and a 'prime' sentence. The purpose of the former was to introduce the characters and the actions they performed on a regular basis, while the purpose of the latter was to describe the specific action that the characters wanted to perform on the particular day. The description in the prime sentence was either congruent or incongruent with the visual stimulus: if the description was incongruent, children were expected to produce a 'target' utterance, fixing the experimenter's erroneous utterance; if not, children did not have to produce an utterance. Given the nature of the task, the experimental and control items had incongruent primes, while the filler items corresponded to congruent primes that did not need correction. In addition to signalling whether a target utterance was required, 'prime' sentences served the function of priming children with the structure that they were expected to use in their target utterances.

As the examples in Table 9 (and matching Figures 24 and 25) illustrate, in the English treatment, the context sentence contained a verb in the simple present tense, while the prime contained a non-finite verb in the case of experimental items, and a finite verb in the present continuous tense in the case of control items. The verbs of target sentences were also expected to be marked like their experimental or control item counterparts. The tense markings were overall similar in Greek – a notable exception being the use of the present continuous tense in control items, since Greek lacks this tense and uses the simple present to mark actions happening at the moment of talking. Thus, the children who participated in the Greek treatment were expected to use simple present in the control condition.

Having developed the items, two semi-random lists were composed, and two matching picture booklets were compiled; the order of the pictures in the picture booklets corresponded to the order of the items on the list. All children saw one of the lists. An effort was made to alternate between lists during testing, so that the number of children who saw List 1 was the

same as the number of children who saw List 2. Bilingual children saw one of the lists in the English treatment and the other list in the Greek treatment.

Table 9. Experiment 2 Example Stimuli.

Item	Context	Prime	Target
Experimental	Every week, John gives his teacher something.	Today, he wants to get her balloons.	He wants to get her flowers
Control	Mary goes for a run on the weekends.	She runs with her cat.	She runs with her dog
Filler (Experimental)	Mary goes to the forest on Sunday.	She wants to climb a tree.	N/A
Filler (Control)	John tidies his toys every evening.	He puts them in a box.	N/A

Figure 24. Example pictures for Experiment 2: Experimental item (left), control Item (right).

Figure 25. Example pictures for Experiment 2: Filler items.

3.1.3. Procedures

As mentioned in the previous chapter, the DPhil project (including Experiment 2) had received clearance from the Central University Research Ethics Committee of the University of Oxford. For the children to participate in the experiment, their parents had to sign an opt-in consent form that made them eligible for the entire battery that also included Experiment 1 and the tests of receptive vocabulary and grammar. Moreover, prior to the administration of the battery, children were asked for their assent to participate.

The design of the experiment was tested and refined as part of a substantive pilot (whose aims are presented in Chapter 2, and whose results are presented in detail in Appendix G). Following a second round of piloting (aimed at checking whether the refined methods worked), the administration of the experiment could take place. Experiment 2, like the entire child battery, was administered twice for bilingual children, once in Greek and once in English, two months apart. The monolingual children completed it once, in their respective L1s.

As was the case with all the tests and tasks that formed part of the testing battery, the instructions for Experiment 2 were given in the children's L1 (that is, Greek for the bilingual

children and the Greek monolingual children, and English for the English monolingual children). In particular, children were told that the experimenter had made mistakes in the stories and were asked to look at the pictures and say what the 'real' stories were, while being audio-recorded. The experiment lasted between 10 to 15 minutes in both assessments. The order in which the measures of the battery were presented is illustrated in Figure 26.

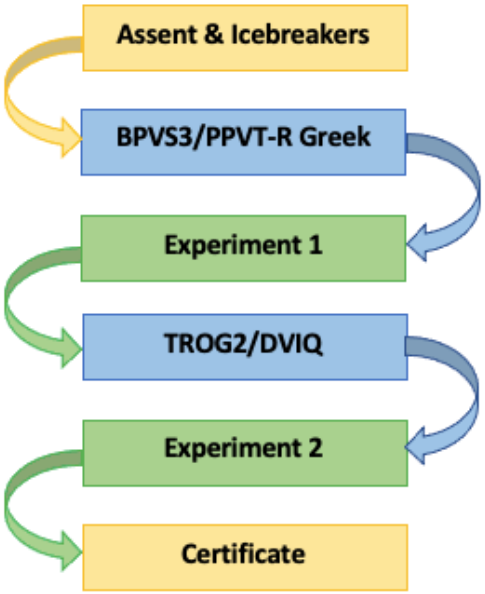


Figure 26. Order of Presentation of the Child Battery Measures.

All children were tested in an unused classroom, office or storage room in their schools. The experimenter was alone with the children in all schools except the partial immersion one, where a teacher was also present in the room during testing. For bilingual children, the administration of the English assessment had to take place online, due to COVID-19 restrictions. Then, a teacher had to be present throughout the procedure to ensure children's safeguarding and to provide technical assistance. The teachers did not interfere otherwise.

The same certificate used for the remainder of the battery was used to keep children motivated in completing Experiment 2. The children received the certificate at the end of their testing session, regardless of their performance in the tests and tasks. Since the administration of the English assessment for the bilingual group took place online, the children were simply given a sticker as a reward for their participation.

3.1.4. Preparations for Data Analysis

The audio-recordings obtained as part of Experiment 2 were transcribed. The transcriptions focused on the children's target utterances (ignoring other things that they might have said during the experiment's administration). Unintelligible segments, pauses and code-switching were marked, but this information was not used in the analyses.

The transcripts were then coded. A coding scheme was developed for both the Greek and the English assessment: in the experimental condition, one point was awarded for the production of the subordinate clause marker ('to' for English and 'na' for Greek), while zero points were awarded for the omission of the linguistic structure. In the control condition, one point was awarded for the production of a finite verb, while zero points were awarded if the verb was absent. Filler items were coded as 'correct' (and awarded one point) if children correctly identified that the picture they saw matched the description they heard, and as 'incorrect' (and awarded zero points) if children incorrectly noted that the picture they saw did not match the description they heard. Children's one-word utterances in the experimental and control conditions were excluded. Moreover, cases where the children misunderstood the experimental or control items as filler items (and did not provide a target utterance) were marked as 'N/A' and not included in subsequent analyses.

To ensure the reliability of the coding, 5% of the transcriptions were coded by two independent raters (an L1 English speaker for the English treatment and an L1 Greek speaker for the Greek treatment who also rated the coding of Experiment 1). Interrater agreement was 84% in the English assessment, and 86% in the Greek assessment. The items that caused disagreement were discussed in an individual meeting with each rater, and a consensus about their coding was reached. No changes in the coding were made as a result.

3.2. Results

I will first discuss the findings that emerged from the Greek and then the English assessment in order to address the first two questions question of the experiment[11]. Hence, I will draw on the relationship between the children's performance on the experiments on the background measures (i.e., the vocabulary and grammar tests) to address the experiment's third question.

3.2.1. Greek Assessment

In the Greek treatment, the bilingual children's performance was compared to that of the Greek monolingual controls. As shown in Figure 27, the performance of all groups in the experimental item condition (where the target structure was: na+verb$_{subjunctive}$) was high: the Greek children scored 7.32 (SD=1.77), while the bilingual children scored 6.99 (SD=1.91). Turning to the control condition (where the target structure was: verb$_{indicative}$), the two groups' scores were also quite high: the Greek children scored 7.35 (SD=1.67), while the bilingual children scored better than in the experimental condition: 8.17 (SD=1.93). As far as the performance of the two bilingual groups themselves is concerned, in the experimental condition the total immersion group scored 7.06 (SD=1.94) and the partial immersion group scored 6.91 (SD=1.90). Rather, in the control condition, the former achieved a mean score of 7.91 (SD=2.16) and the latter achieved a score of 8.63 (SD=1.67). These data are presented in Figure 28.

Looking at the two figures, it can be observed that all participants' performance in the filler items was (almost) at ceiling: the Greek group's mean score was 8.85 (SD=.96), while the total immersion group scored 9.11 (SD=.99) and the partial immersion group scored 9.06 (SD=1.08). The children's high performance on the filler items suggested that they were attentive during the presentation of the verbal and visual stimuli.

[11] The data can be accessed at: https://osf.io/r4zt7/.

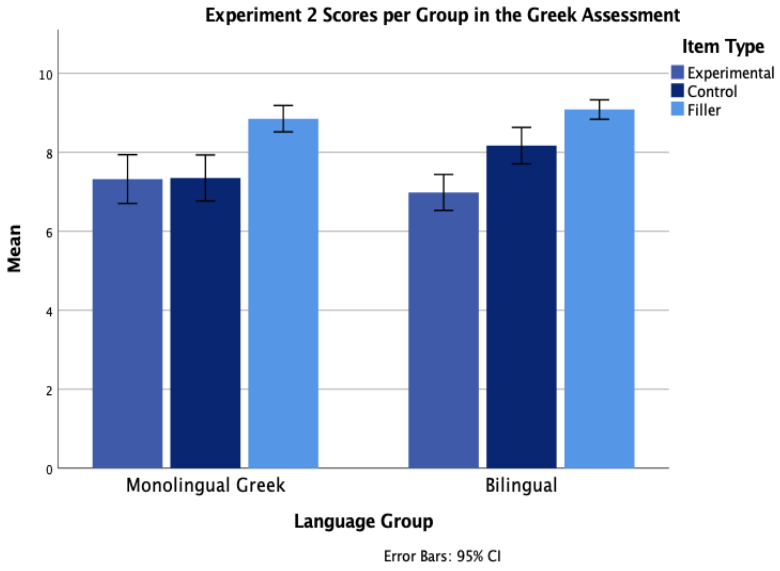


Figure 27. Monolingual and bilingual children's experimental, control and filler item scores in the Greek assessment of Experiment 2.

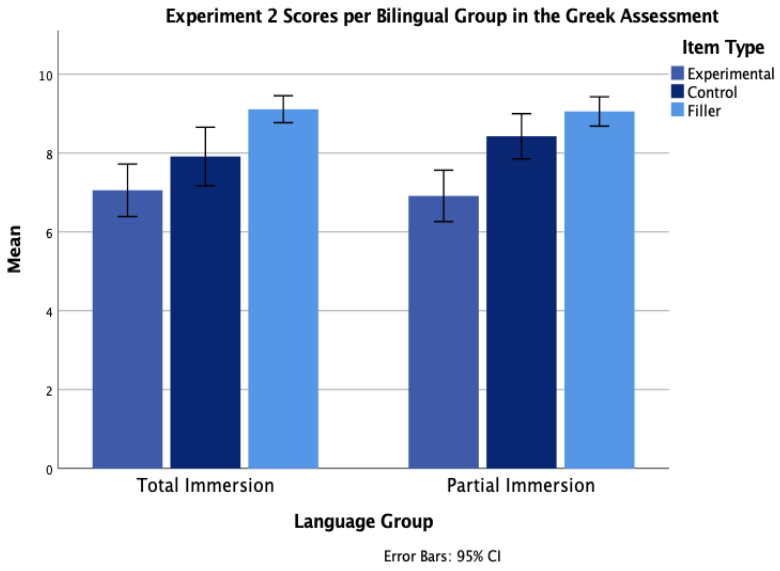


Figure 28. Total and partial immersion children's experimental, control and filler item scores in the Greek assessment of Experiment 2.

After ensuring the normality and homogeneity of variances of each group (see Appendix M), it was possible to compare their performance in the experimental and the control condition of the Greek assessment (note that the children's performance in the filler condition was not analysed statistically). Two 2x2 ANOVAs were conducted. The first compared the monolingual and the bilingual children; it indicated a significant interaction between language group and condition [$F(1,102)=4.60$, $p=.03$, $\eta^2_p=.06$], suggesting that the two groups performed

differently in the experimental and the control condition. In addition, there was a simple main effect of condition [F(1,102)=5.07, p=.03, η^2_p=0.5], suggesting that (as expected) the two conditions differed, such that the experimental condition resulted in lower scores than the control condition. Nevertheless (and also in line with expectations), there was no simple main effect of group [F(1,102)=.74, p=.40, η^2_p=.01] – according to which the two groups did not manifest statistically significant differences.

The second 2x2 ANOVA concerned the two bilingual groups. As expected, it did not reveal a significant group*condition interaction [F(1,68)=1.00, p=.32, η^2_p=.01] but it did indicate a main effect of condition: F(1,68)=13.00, p=.001, η^2_p=.16. Moreover, it did not reveal a main effect of group: F(1,68)=.33, p=.57, η^2_p=.01.

3.2.2. English Assessment

In the English assessment, the bilingual children were compared to age-matched English monolingual controls. As shown in Figure 29, the English children achieved a mean score of 6.69 (SD=2.53) in the experimental item condition (where the target structure was: to+verb$_{infinitive}$), while the bilingual children achieved a mean score of 4.93 (SD=3.04). Similar patterns were observed in the groups' control item performance (where the target was: verb$_{finite}$), although the scores were higher on the whole. The mean score of the English group was 8.64 (SD=1.22), while the bilingual group scored 5.24 (SD=3.12).

The results of the two bilingual groups can be found in Figure 30. Some differences can be noticed between the groups: in the experimental condition, the total immersion group scored 6.51 (SD=2.42), while the partial immersion group scored 3.34 (SD=2.78). In the control condition, the former group achieved a mean score of 5.77 (SD=2.73), while the latter group's mean score was 4.71 (SD=3.42).

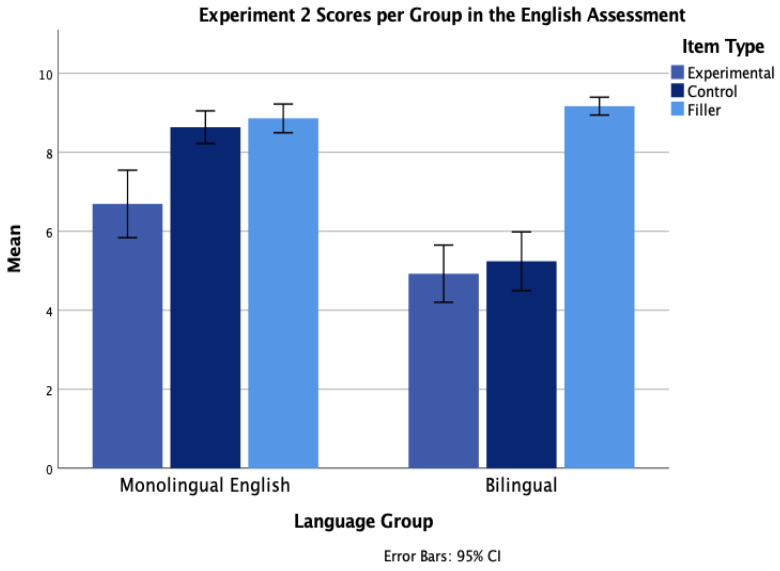


Figure 29. Monolingual and bilingual children's experimental, control and filler item scores in the English assessment of Experiment 2.

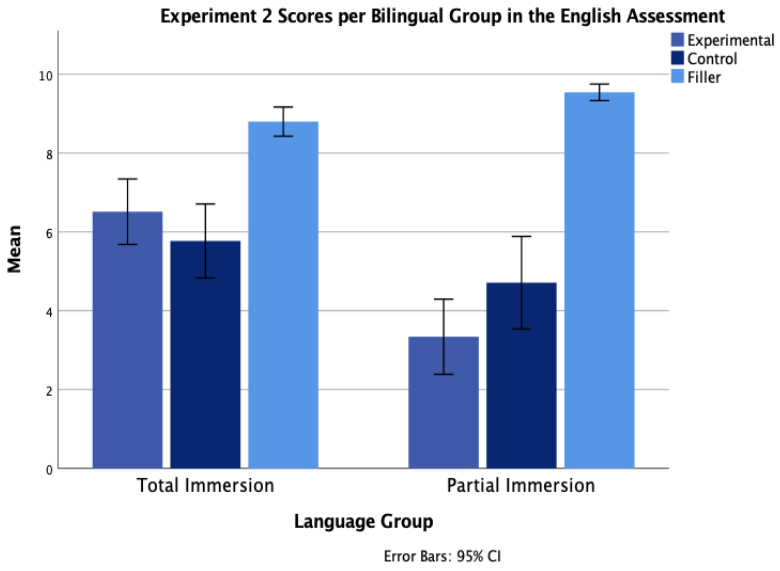


Figure 30. Total and partial immersion children's experimental, control and filler item scores in the English assessment of Experiment 2.

With regards to filler items, all groups achieved very high scores (as was the case in the Greek assessment). The English monolinguals scored 8.86 (SD=1.07), the children of the total immersion programme scored 8.80 (SD=1.08), while the children of the partial immersion programme scored 9.54 (SD=.61). The children's excellent performance in the filler condition can be taken as evidence of their attention throughout the experiment. Moreover, it suggests

that the differences in children's performance that were reported in the previous sections were not due to extra-linguistic factors.

Having checked the normality and homogeneity of variances of each group in this assessment (see Appendix N), two more 2x2 ANOVAs were conducted in order to examine the children's performance in the experimental and control conditions (the filler condition was excluded from the analysis). The first compared the monolingual and the bilingual children, and revealed an interaction between language group and condition [$F(1,104)=7.07$, $p=.01$, $\eta^2_p=.06$], such that the two groups performed differently in the experimental and the control condition. As expected, there was a simple main effect of condition [$F(1,104)=13.58$, $p<.001$, $\eta^2_p=.16$], suggesting that the two conditions' scores differed from one another (once again with children achieving higher scores in the control condition). Moreover, and as expected, there was a simple and large main effect of group: $F(1,104)=29.44$, $p<.001$, $\eta^2_p=.22$. To explore this effect further, two independent samples t-tests (with group as the independent variable and experimental or control item scores as the dependent variable) were conducted. These revealed that the English monolinguals differed from the Greek-English bilinguals in both the experimental and the control conditions: $t(104)=2.99$, $p=.003$ for the experimental condition; $t(104)=8.00$, $p<.001$ for the control condition.

The second 2x2 ANOVA focused on the two bilingual groups; its results suggested that there, once again, was an interaction between language group and condition: $F(1,68)=7.77$, $p=.01$, $\eta^2_p=.10$. Nevertheless, there was no simple main effect of condition: $F(1,68)=.69$, $p=.41$, $\eta^2_p=.01$. Finally, there was a simple main effect of group: $F(1,68)=13.83$, $p<.001$, $\eta^2_p=.17$. To unpack this effect, two additional independent samples t-tests were conducted. These revealed that the children in total immersion differed from the children in partial immersion in the experimental condition [$t(68)=5.10$, $p<.001$] but not in the control condition of Experiment 2 [$t(68)=1.43$, $p=.16$].

3.3. Background Measures

While the group comparisons presented above can answer the first and second question of the experiment, different tests are needed to answer the third question, focusing on the effect(s) of age and/or proficiency on children's performance in the experiment. Proficiency was measured through tests of receptive vocabulary and grammar. The (monolingual and bilingual) children's performance in each of these tests was presented in detail in Chapter 2. Thus, this chapter will focus on the analyses that explore the relationship between the children's proficiency and their use of subordinate clause markers.

3.2.3.1. Effects of Proficiency (& Age) in the Greek Assessment

Two bivariate Pearson's correlations were conducted to explore the children's performance in the Greek assessment of Experiment 2. Table 10 presents the analysis for the Greek group, while Table 11 presents the analysis for the bilingual group.

Table 10. Correlations between age, proficiency and Experiment 2 scores (experimental and control) for the Greek children in the Greek assessment.

	Age	Vocabulary	Grammar	Experimental	Control
Age	-	.68***	.50***	.20	.16
Vocabulary	-	-	.66***	.38*	.15
Grammar	-	-	-	.05	-.08
Experimental	-	-	-	-	.20
Control	-	-	-	-	-

Table 11. Correlations between age, proficiency and Experiment 2 scores (experimental and control) for the bilingual children in the Greek assessment.

	Age	Vocabulary	Grammar	Experimental	Control
Age	-	.52***	.01	-.18	.14
Vocabulary	-	-	.37***	-.10	.21
Grammar	-	-	-	.22	.01
Experimental	-	-	-	-	-.03
Control	-	-	-	-	-

Besides the aforementioned correlations between age, vocabulary and/or grammar, there was a moderate, positive and significant correlation between vocabulary and experimental item scores for the Greek group (Figure 31). No significant correlations with experimental and/or control items were found for the bilingual group.

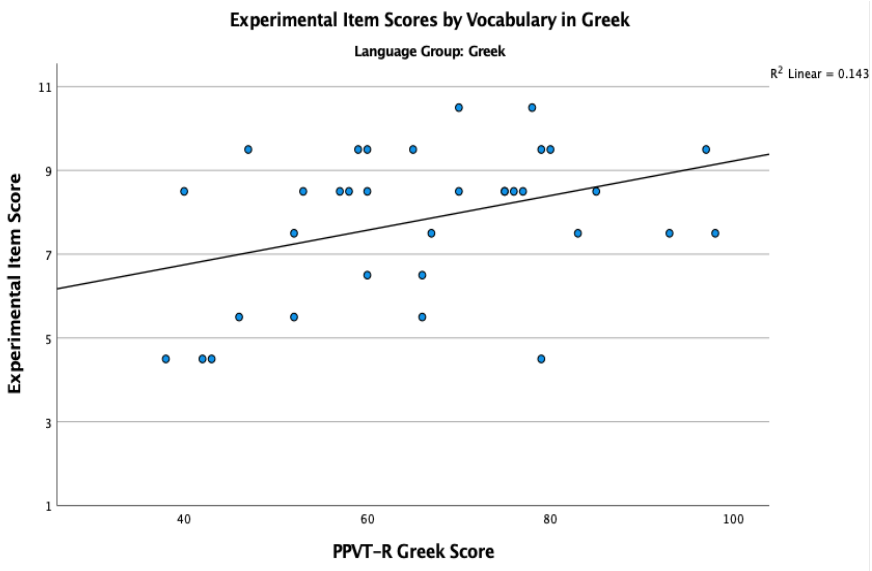


Figure 31. Correlation between the Greek children's experimental item scores and their vocabulary scores in Greek.

3.2.3.2. Effects of Proficiency (& Age) in the English Assessment

To explore the relationships between children's vocabulary and grammar, their age, and their scores in the English assessment of Experiment 1, two more bivariate Pearson's correlations were conducted: the first is for the English group and is presented in Table 12, while the second is for the bilingual group and is presented in Table 13.

Table 12. Correlations between age, proficiency and Experiment 2 scores (experimental and control) for the English children in the English assessment.

	Age	Vocabulary	Grammar	Experimental	Control
Age	-	.85***	.76***	.35*	.19
Vocabulary	-	-	.68***	.22	.20
Grammar	-	-	-	.39**	.26
Experimental	-	-	-	-	.51***
Control	-	-	-	-	-

Table 13. Correlations between age, proficiency and Experiment 2 scores (experimental and control) for the bilingual children in the English assessment.

	Age	Vocabulary	Grammar	Experimental	Control
Age	-	-.14	-.09	-.08	.04
Vocabulary	-	-	.80***	.25*	.21
Grammar	-	-	-	.30**	.31**
Experimental	-	-	-	-	.42***
Control	-	-	-	-	-

The tables highlight the correlations between age and/or the bilingual and monolingual children's lexical and grammatical proficiency in English that were first presented in Chapter 2. In addition, they suggest that vocabulary was only related to the bilingual group's performance, such that, as shown in Figure 32, the higher the children scored in the vocabulary test, the higher they (also) scored in the experimental item condition.

Grammar scores were related to both the monolingual and the bilingual children's performance: for the English group, there was a moderate, positive and significant correlation between grammar and experimental item scores (see Figure 33); for the bilingual group, there were moderate, positive and significant correlations between grammar and experimental item scores as well as with control item scores, illustrated together in Figure 34. These correlations suggest that monolingual and bilingual children with higher scores on the grammar test also achieved higher scores on (aspects of) the experimental task.

The effect of age was limited to the English group, where age correlated positively with experimental item scores (see Figure 35), such that older children were found to produce the marker 'to' with greater accuracy. No such relationships were detected for the bilinguals.

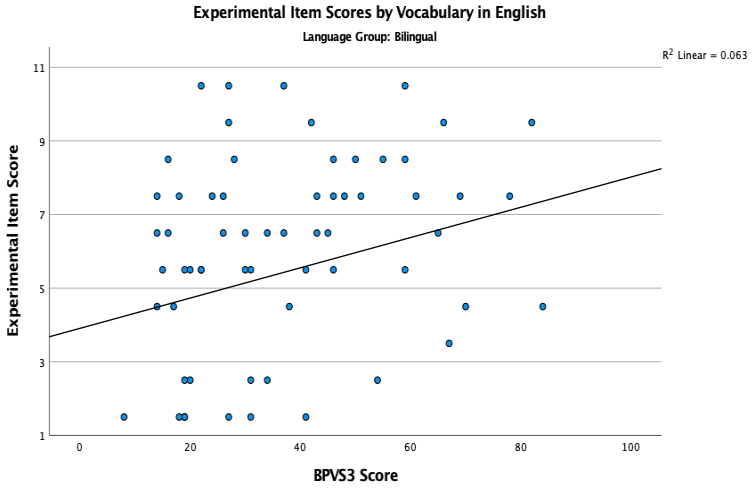


Figure 32. Correlation between the bilingual children's experimental item scores and their vocabulary scores in English.

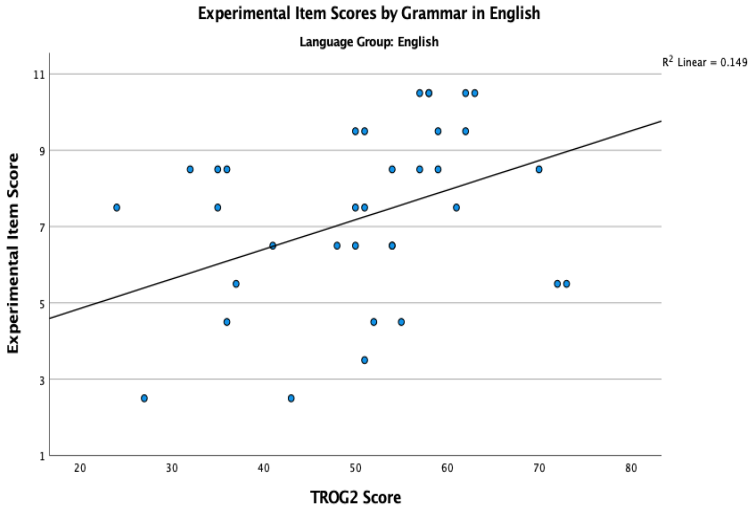


Figure 33. Correlation between the English children's experimental item scores and their grammar scores in English.

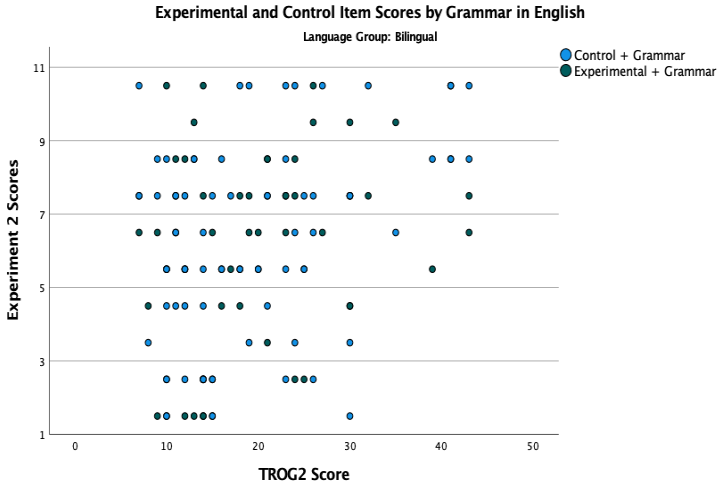


Figure 34. Correlation between the bilingual children's experimental and control item scores and their grammar scores in English.

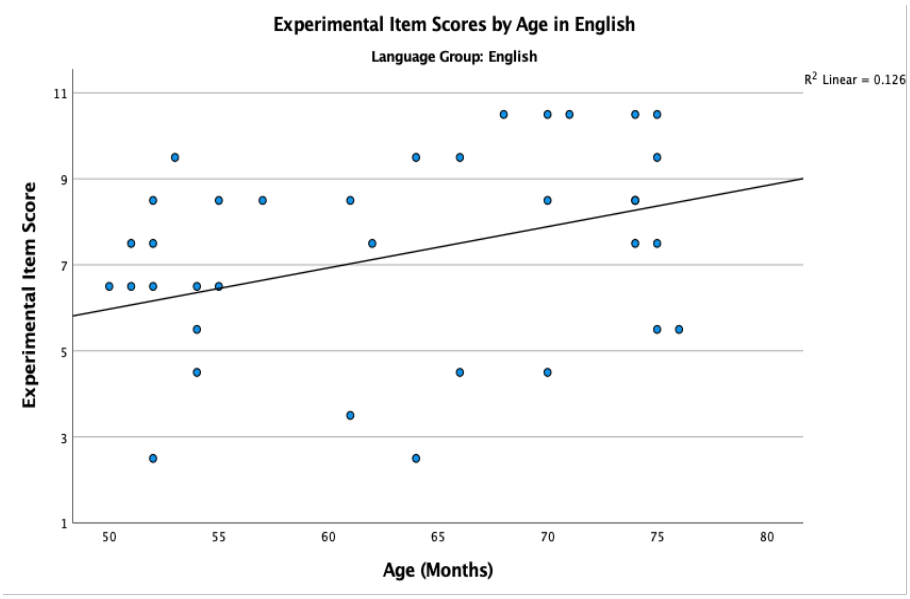


Figure 35. Correlation between the English children's age and experimental item scores in English.

Finally, there was a strong, positive and significant correlation between experimental and control item scores for the English group (Figure 36), suggesting that the more accurate children were in producing subordinate clause markers, the more accurate they were in producing verbs in the appropriate tense and/or mood. A similar relationship existed for the bilingual group, whose experimental and control scores were characterised by a moderate, positive and significant correlation (Figure 37).

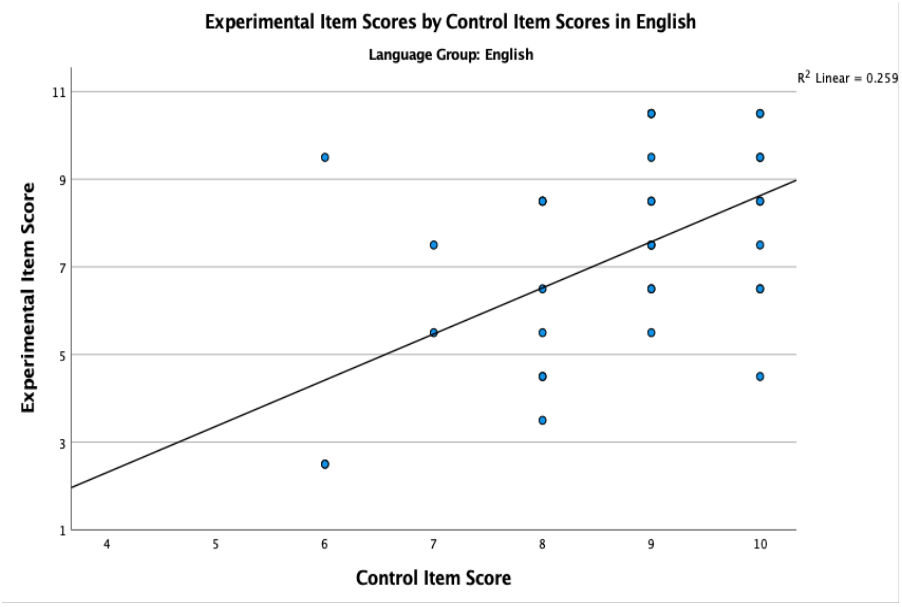


Figure 36. Correlation between the English children's experimental and control item scores in English.

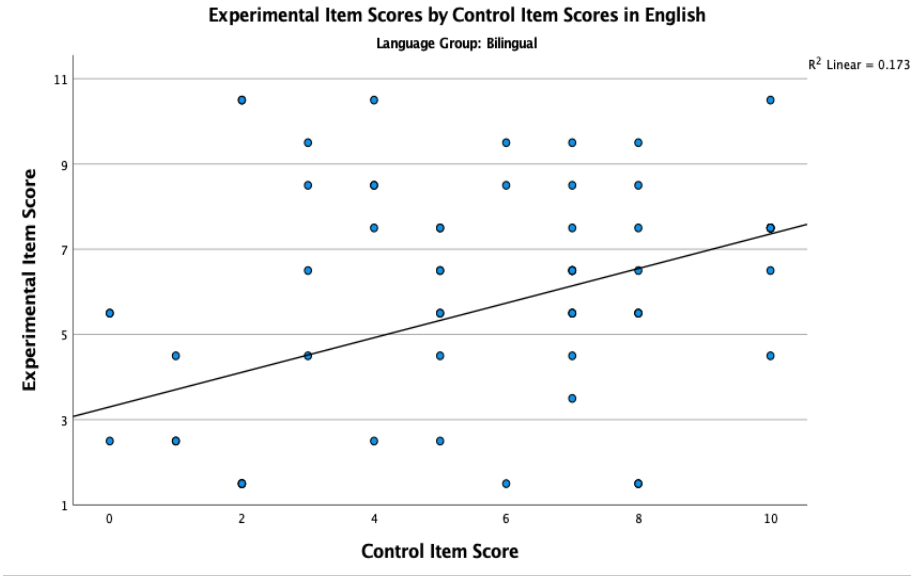


Figure 37. Correlation between the bilingual children's experimental and control item scores in English.

To explore the role that proficiency, as well as age, played in determining the bilingual children's experimental item scores in the English assessment (where there were significant correlations between the focal variables), a hierarchical linear regression was performed. Three models were put to the test: the first included vocabulary as the sole predictor; Model 2 also included grammar, while Model 3 added age to the previous two variables.

The results of the regression (which can be found in Table 14) suggest that the first two models were a reasonably good fit. In particular, Model 1 accounted for 6.3% of the variance [R^2=.06, F(1,68)=4.61, p=.04], with vocabulary being a significant predictor of children's scores. Model 2 accounted for an additional 2.4% of the variance (raising the total to 8.8%): R^2=.09, F(1,67)=3.21, p=.05. However, it did not reveal vocabulary or grammar to be significant predictors of children's experimental item scores. Finally, Model 3 accounted for 9.1% of the total variance in bilingual children's scores [R^2=.09, F(1,66)=2.19, p=.10]. None of the variables entered in the model were found to be significant predictors of performance.

Table 14. Predictors of Experiment 2 Scores in English.

	Model 1			Model 2			Model 3		
Variable	B	*SE* B	β	B	*SE* B	β	B	*SE* B	β
Vocabulary	.04	.02	.25*	.01	.03	.04	.01	.03	.03
Grammar				.09	.07	.26	.09	.07	.26
Age							-.04	.08	-.06
R^2		.06			.09			.09	
F Change in R^2		4.61*			1.77			.22	

3.3. Discussion

The data presented in the previous section can shed light on the questions that Experiment 2 sought to answer. The first question was: 'does the bilingual children's use of subordinate clause markers in English and Greek differ from that of their monolingual peers?'. It was predicted that the bilingual children would produce (and/or omit) the subordinate clause marker 'to' in English and 'na' in Greek at a similar rate to both their English and their Greek monolingual peers. This prediction was borne out in the Greek assessment, where the bilingual and the Greek monolingual children performed alike. Yet, it was not borne out in the English assessment, where the bilingual children omitted subordinate clause markers more often than English monolingual children.

The difference between the bilingual and the monolingual children cannot be attributed to CLI from the children's L1, Greek, to their L2, English, since both languages possess the target structure. According to theoretical accounts of CLI (e.g., Paradis & Genesee, 1996), structural similarity should prevent the manifestation of differences between bilinguals and monolinguals. At the same time, it might also facilitate the bilingual children's acquisition of this, arguably difficult, structure.

An alternative, and more plausible, explanation of the difference between Greek-English bilingual and English monolingual children in the English assessment could be that the

omissions of the subordinate clause marker 'to' by bilingual children result from processing limitations (Eilers & Oller, 1976). This idea could be supported by the observation that 4.7% of the monolingual English children also omitted the structure from their utterances, which indicates that, on the whole, the monolingual participants were aware of the fact that the subordinate clause marker should be present in order for the sentence to be grammatical, but omitted it in error. Completing the experiment in an L2 must have increased the level of the challenge for the bilingual children, who might have omitted the marker in an effort to produce the entire target sentence as the task required.

This suggestion was tested by Kirjavainen et al (2009), who compared children's utterances with and without the marker 'to' in order to see if the omissions resulted from performance limitations. The authors reported that the two types of utterances had the same mean length of utterance, and rejected the suggestion that 'to' omissions are (solely) attributable to performance limitations. Instead, Kirjavainen et al (2017) discuss the 'to' omissions under a usage-based framework of language acquisition, according to which children acquire linguistic structures through noticing their characteristics and calculating their distribution in the input to which they are exposed. Following Bybee (1985), they suggest that high frequency promotes lexical learning, while high variability promotes generalisation in acquisition (also see Goldberg, 2006; Wonnacott, Boyd, Thomson & Goldberg, 2012; Wonnacott, Brown & Nation, 2017). Under this framework, 'want+to+verb' structures are in competition with 'want+DP' structures, given their lexical and semantic overlap; however, as the latter are both more frequent and more abstract, they are likely to be acquired and used with greater ease and speed than the former.

It is worth noting that the bilingual children who are tested in this experiment have two sets of distributions to compute: one for their L1 and one for their L2. Since the 'want+to+verb' and the 'want+DP' contrast exists in Greek as well, it might be the case that 'want+DP'

acquires an even higher token frequency, whilst retaining its greater variability. These cross-linguistic characteristics, coupled with the fact that, in Greek, want+to+verb constructions are even more difficult than they are in English as they require the verb following the marker to appear in the subjunctive mood, might make its appearance scarce compared to its 'want+DP' rival at the earliest stages of L1 acquisition. Then, these use patterns might be transferred to the L2. As the children tested in the experiment are well past the earliest stages of L1 acquisition, it would make sense that their acquisition of the target structure is on par with the Greek monolinguals'. However, as their acquisition of the L2 is at its earliest stages (with most children having joined their English-medium preschools at or after the age of 3;0), it could be that structure competition could lead to the omissions observed in this experiment.

If this idea is on the right track, the bilingual children in total immersion, who receive more input in English (and, thus, have more opportunities to familiarise themselves with the target structure), should omit the marker 'to' less often in English than their partial immersion counterparts. This was the issue addressed in the second question of the experiment ('do the children in total immersion perform differently from children in partial immersion?'). The prediction associated with this question was also that children in total immersion would score higher than children in partial immersion with respect to the production of 'to' in the English assessment. Thus, it was borne out given the experiment's findings.

Since the children with the highest proficiency in English omit the target structure less, it would be logical for proficiency to affect their performance in the task. Indeed, in view of the third question that this experiment sought to answer ('does their proficiency in English affect their performance in the task and, if so, how?'), it was predicted that higher vocabulary and/or grammar would help reduce children's omissions. This prediction was confirmed by the data; with regards to the English assessment in particular, both vocabulary and grammar were correlated with the bilingual children's production of the marker 'to', such that bilingual

children with higher vocabulary and/or grammar also produced the marker 'to' more as part of their utterances. In the regression analyses that were performed, vocabulary was found to be a predictor of children's accurate use of the subordinate clause marker. This finding hints that children acquire 'verb+to+verb' as a lexically-specified structure. This was also argued with regards to sentential subjects in Chapter 2. Nevertheless, and as mentioned in Chapter 2, this idea is a mere speculation at present, since neither experiment was powered for regressions. Moreover, in this case, vocabulary did not explain a substantial proportion of the variance. In short, more research is required to assess if the speculation has merit.

In short, the findings of Experiment 2 suggest that the bilingual children omit the subordinate clause marker 'to' in English more often than English monolinguals (but produce the marker 'na' in Greek as often as the Greek monolinguals). Moreover, they highlight that the bilingual children perform according to the amount of input they receive in English and their proficiency in the language. However, their omissions of 'to' cannot be attributed to CLI, as the structure appears in a similar and function in English and Greek.

As Experiment 2 was conducted to further explore the nature of the omissions observed in Experiment 1, the findings of the second experiment also impact the conclusions of the first experiment. In view of the fact that omissions in child language are systematic, they should all have a common underlying cause. If the underlying cause of omissions in Experiment 1 was CLI, no omissions of subordinate clause markers would be expected in Experiment 2 since the latter experiment targeted a structure that is not believed to be affected by CLI. As such, the fact that the bilingual children omitted the subordinate clause markers renders the omissions of sentential subjects impossible to explain under traditional accounts of CLI.

The usage-based frameworks of language acquisition that were mentioned earlier in the thesis might offer a useful alternative explanation to this end. As discussed in Chapter 1, these accounts suggest that children acquire linguistic structures by taking into account their

characteristics and distribution in the L1 input (e.g., how frequent, salient and/or variable they are) (Ambridge et al., 2015). Following this logic, it can be argued that the children might then transfer their knowledge of the L1 distributional properties to the L2. This idea can provide a uniform explanation of children's omissions in both experiments. Moreover, it can speak to why the differences between the groups were more pronounced with regards to sentential subjects: one the one hand, sentential subjects are frequent and salient structures whose distributional properties differ between English and Greek; on the other hand, subordinate clause markers are infrequent and non-salient structures, whose distributional properties do not differ in the two languages. Given their characteristics, subordinate clause markers are more difficult to acquire than sentential subjects for monolingual and bilingual children alike. However, the bilingual children's exposure to the L1 (i.e., Greek) input might lead them to ignore sentential subjects but attain to subordinate clause markers in the L2 (i.e., English) (Ellis, 2019; Ellis et al., 2014; Ellis & Sagarra, 2010). Therefore, the cross-linguistic distributional properties of the structures under investigation, coupled with the selective attention mechanism developed as a result of prior experience with the L1, could explain the differences between the bilinguals and the monolinguals with regards to both structures. This novel usage-based account of CLI will be discussed in greater detail in Chapter 5.

Chapter 4: Adult L2 Learners' Use of Referential Expressions

The two experiments presented in Chapters 2 and 3 aimed to uncover the role of crosslinguistic influence (CLI) in Greek children's acquisition of English as a Foreign Language (EFL)[12]. The first experiment targeted children's use of sentential subjects, while the second explored their production of subordinate clause markers. In both investigations, the Greek-English sequential bilingual children appeared to omit the target structures more often than a group of age-matched English monolingual controls – thus behaving like adult L2 learners (and unlike simultaneous bilingual learners). However, the English monolingual children also appeared to omit sentential subjects (and subordinate clause markers) from their utterances. This finding could suggest that the bilingual children's omissions represent a point along a developmental trajectory that all children will outgrow, rather than a CLI effect.

To test this idea, adult learners were tested. Adults have a foundation in both their first (L1) and their second language (L2) and have received more cumulative input in the L2[13]. Despite the notable differences between the two groups, the comparison between the adult learners and the child learners was deemed worthwhile, as it would allow assessing if the same CLI patterns are evident in adult learners whose acquisition is further along the developmental continuum (due to the additional cumulative exposure they have amassed). In turn, this could shed more light on the phenomenon of CLI and its underlying causes: if the children's

[12] Note that the children who take part in the first two experiments learn EFL at preschool. The fact that the children's acquisition of an L2, English, starts around the age of 3;0 makes it possible to characterise them as (sequential) bilingual even though, in essence, they are EFL and/or L2 learners – not unlike the adults who took part in the third experiment (some, but not all, of whom are also sequential bilingual). Of course, as discussed elsewhere in the thesis, the boundaries that separate simultaneous, sequential and L2 learners are not clear-cut; this allows for some flexibility with regards to the use of the terms.

[13] It should be highlighted that the adult participants tested in Experiment 3 were different from the child participants tested in the previous two experiments. A major difference between the two groups was in terms of their cumulative input, as discussed in the paragraph; other differences include: the age at which they started learning English (with most adult learners having started to learn English at primary school); the educational context in which their L2 learning unfolded (with most adult learners having learned English in a Greek monolingual school) and the resulting quality of L2 input that the learners received (with most adult learners having been taught by L1 Greek teachers).

omissions in the first two experiments represent a phase, the adult participants should have outgrown it and, therefore, should not make (or at least should make fewer) mistakes in terms of the production of the target structures. After all, there is evidence that children opt for different strategies than adults in the course of processing and interpreting null and overt referents, but age minimises the differences between the two groups – such that older children behave more 'adult-like' (Papadopoulou et al., 2014). By contrast, if the sequential bilingual children's omissions result from CLI, the adult participants should not differ from their child counterparts and, thus, would be expected to omit the target structures to some extent.

4.1. Introduction

Experiment 3 focused on Greek adult EFL learners' use of sentential subjects and not subordinate clause markers, as this linguistic structure yielded greater differences between the bilingual and monolingual children, according to the results of Experiments 1 and 2. Moreover, this feature has received more attention in the (bilingual) language acquisition literature and can be tested through various experimental tasks. In addition to the experimental task itself, Experiment 3 also included measures of language use and linguistic proficiency. This chapter, therefore, explains the rationale behind experiment 3 and outlines its methods. It includes a detailed presentation of the results that were obtained and a brief discussion of the main issues that stem from the results (which will be explored further in the following chapter).

4.1.1. Theoretical Background

As noted throughout the thesis, Greek and English display similarities and differences in the use of sentential subjects. As a general rule, Greek prefers sentential subjects to be null, while English requires subjects to be overt. However, there are exceptions to the rule: indeed, Greek sometimes requires subjects to be overt, while English allows subjects to be null in (certain)

contexts[14]. In Greek, the introduction of a topic involves the use of a Determiner Phrase (DP), while the shift to an aforementioned topic involves the use of a pronoun or a determiner phrase. Topic maintenance across sentences involves the use of a null subject.

While English resembles Greek with regards to the referring expressions used for topic introduction and shift, it differs from Greek as far as the expressions used for topic maintenance across clauses are concerned: indeed, in English, referring to the same topic across sentences requires the use of an overt pronoun. These issues are illustrated in Examples (1) and (2), which were first presented in Chapter 1.

1. Alex is is in the garden. He is picking the apples. The apples are ripe.

2. O Alex ine ston kipo. [...]
   The.MASC.SG Alex is in-the.NEUT.SG garden
   [...]Mazevi ta mila. [...]
   Picks the.NEUT.PL apples
   [...] Ta mila ine orima.
   The.NEUT.PL apples are ripe

Adult speakers of a null-subject language (e.g. Greek) are found to accept and produce more (illicit) null subjects in a non-null-subject language (e.g. English) (Prenza, 2014; Prenza & Tsimpli, 2013; White, 1985). This finding is often attributed to CLI: the adults have acquired the rules that determine the distribution of null and overt subjects in their null-subject L1 and overgeneralise the rules to their non-null-subject L2. Yet, the undesirable effects of CLI from the L1 to the L2 can increasingly disappear with continued exposure to and increased proficiency in the L2: as speakers receive more exposure or increase their proficiency in the

[14] Indeed, particular contexts (e.g., diaries, informal speech, and particular registers) allow subjects to remain null (Wagner, 2018). In addition, English used to allow null subjects in the past – when its morphology was richer and its word order was flexible.

L2, they become better equipped to reject illicit null subjects in grammaticality judgement tasks as well as to produce them in language elicitation tasks.

There is less research on whether adult speakers (of a null-subject L1) are able to produce the right kind of referring expression in different discourse contexts according to the pragmatic considerations of their (non-null-subject) L2. The research that has been conducted suggests that, in general, adult speakers tend to be over-explicit in their L2 – opting for low accessibility markers, such as DPs, in contexts where high accessibility markers, such as pronouns, would be expected (Ahrenholz, 2005; Chini, 2005; Hendriks, 2003; Kang, 2004; Leclerq & Lenart, 2013; for an explanation of the terms 'high accessibility' and 'low accessibility' see Ariel, 1990). Some studies consider over-explicitness as a characteristic of intermediate learners, and argue that it decreases as a function of proficiency (e.g. Chini, 2005); however, other studies report that over-explicitness applies to all L2 learners, irrespective of their proficiency (e.g. Leclerq & Lenart, 2013).

4.1.2. Research Questions and Predictions

Taking these theoretical and empirical issues into account, Experiment 3 was designed to answer the following questions:

1. Do (adult) Greek EFL learners exhibit quantitative and/or qualitative differences in the use of sentential subjects in English relative to English L1 speakers?
    a. How often do they use sentential subjects?
    b. How often do they omit subject arguments?
    c. How often do they provide an inappropriate form of the subject argument?
    d. Do they have a preference for proper nouns or pronouns?
    e. Overall, are they as accurate in their production of sentential subjects?
2. What is the relationship between the adult Greek EFL learners':
    a. use of sentential subjects in English and their exposure to English?

b. use of sentential subjects in English and their linguistic proficiency in English?

c. exposure to and proficiency in English?

3. What are the factors that explain the adult Greek EFL learners' use of sentential subjects relative to English L1 speakers?

Starting from the first question, it is conceivable that the Greek EFL learners' production might differ from that of their English L1 counterparts from a quantitative perspective. For instance, the learners might use sentential subjects less often in their narrations (1a), omit and/or provide inappropriate forms of the subject argument more often (1b and 1c), and be less accurate in their production of sentential subjects overall (1e), relative to the English speakers. In addition, the two groups might differ in qualitative respects, with each group having a (contrasting) preference for a particular kind of subject (e.g. DPs over pronouns, or pronouns over DPs) (1d): taking into account previous studies on the topic, it is expected that the EFL learners will demonstrate a 'preference' for DPs, such that they will use this form even in discourse inappropriate contexts.

Turning to the other two questions, given the existing literature on the topic, exposure and proficiency are predicted to correlate with the learners' use of sentential subjects, such that learners with high exposure to and proficiency in English will be more accurate in the use of subjects. In addition, exposure and proficiency are expected to correlate with one another, such that learners with high exposure to English will also be more proficient in the language. Though the research on the factors that underlie sentential subject use is limited, exposure and proficiency are expected to predict the learners' use of sentential subjects in the present experiment. Proficiency, measured through a vocabulary test, is also expected to predict the English L1 speakers' use of sentential subjects.

4.2. Methods

This section presents the methods that were used to address the experiment's RQs. It describes the characteristics of the adult Greek EFL learners and English L1 speakers who took part, outlines the design and procedures involved in the experiment, and explains how the data obtained was prepared for analysis.

4.2.1. Participants

The participants were recruited using snowball sampling. This procedure involves the researcher(s) selecting a group of participants who meet the inclusion criteria and asking the participants to recommend other (groups of) participants who also meet the inclusion criteria (Rose et al., 2020). In the context of the experiment, potential participants were identified among Greek and British friends and acquaintances and were sent an invitation to take part in the study (containing the relevant link) via email, text or message on social media (Appendix O). They were then asked to pass on the invitation to others within their circles.

To be included in the sample the Greek EFL learners had to: (i) have Greek as their sole L1, (ii) have learned English after the age of 3;0 and (iii) understand English well enough in order to complete the questionnaire, test and task that the study involved. The only requirement for English monolinguals was that they had to have English as their sole L1 (even though they could have learned one or more L2s later in life). As the experiment took place online, there were some additional inclusion criteria that applied to both groups. In particular, all participants were asked to use a device with a microphone (so that their speech could be recorded) and a strong and stable connection to the internet (i.e., four MBPS or higher). The latter requirement was put in place after piloting, as it became clear that people with less strong and stable internet connections were unable to complete the experiment.

To prevent participants who were not eligible to take part from entering the online study, an information form outlining the inclusion criteria was presented before the start of the

experiment. Moreover, the demographics questionnaire that formed part of the experiment allowed identifying participants who did not meet the inclusion criteria despite having signed the consent form. It is worth noting that all of the participants who signed the consent form were indeed eligible to participate.

A power analysis was conducted on G*Power (Faul, et al., 2007) prior to the launch of the experiment, based on the average effect size obtained in the experimental and control condition of Experiment 1 (since Experiment 3 was conducted in order to explore adult learners' performance with regards to the same structure that was tested in Experiment 1). Focusing on main effects, the power analysis indicated that a total sample of 50 (with 25 participants in each group) would allow for 90% power at the α=.05 level.

The initial sample consisted of 90 adults (47 Greek EFL learners and 43 English L1 speakers). However, a substantial number of participants had to be excluded for numerous reasons. In particular: five Greek and two English participants dropped out right after providing their consent to take part in the online experiment; two Greek and two English participants did not answer one or more of the sets that formed part of the vocabulary tests and, thus, their data could not be used; two Greek and nine English participants were prevented from taking part as their devices and/or browsers did not support audio-recording; and seven Greek participants were excluded because some or all of their transcriptions were not uploaded on the platform due to fluctuations in their internet connection during the experiment. Thus, the final sample consisted of 61 participants (31 Greek EFL learners and 43 English L1 speakers), whose sex and age characteristics are presented in Table 15 below:

Table 15. Age and gender characteristics of the adult participants.

	Male	Female	Mean Age (Years)	SD Age (Years)	Total
Greek EFL	12	19	31.35	11.89	31
English L1	12	18	31.65	9.64	30
Total	25	36	31.45	10.69	61

A dimension that the Greek EFL learners differed on was their Age of Onset (AoO) – or, in other words, the age at which they started learning English. While most of them (i.e., 24 out of 31) reported that they started learning English at primary school, there were five participants whose acquisition began at preschool and two (older) participants whose acquisition began at secondary school. The distribution of AoOs in the sample reflects Greece's foreign language provision: while in the past foreign languages were introduced at secondary school, they now form part of primary school curricula. To this effect, Greek children start to learn an L2 at the age of 8;0 (Grade 2) and an L3 at the age of 10;0 (Grade 4) (Eurydice, 2017). At the same time, children can attend private preschools that include L2 lessons in their curriculum (Dendrinos et al., 2013).

Moreover, 18 of the EFL learners claimed to know other languages in addition to Greek and English. The most common third languages (L3s) in the Greek group were French (39.5%), German (23.7%) and Spanish (23.7%); followed by Italian (10.5%) and Catalan (2.6%). The fact that most of the Greek EFL participants knew an L3, as well as the distribution of L3s in the sample, can once again be explained by the observation that, according to Greece's foreign language educational provision, all Greek pupils must be taught English as their L2 and can choose between French, German or Spanish as their L3 (Eurydice, 2017). It should be noted that eight of the participants who claimed to know other languages were, in fact, able to speak more than one L3.

Half of the 30 English L1 participants were found to speak a language in addition to English. Similar to the Greek group, the most common L2s among participants in the English sample were French (27.5%), German (17.5%) and Spanish (12.5%) – followed by Czech, Dutch, Hindi, Japanese, Korean, Portuguese, Russian, Vietnamese (5% each), and Italian (2.5%). Most of the participants who reported knowing a language in addition to their L1 (i.e., 13 out of 15) indicated that they were able to speak multiple languages.

4.2.2. Materials

The experimental battery was created using Gorilla Experiment Builder. It included two questionnaires asking participants about their characteristics and exposure to English, a Yes/No vocabulary test measuring participants' knowledge of receptive vocabulary in English, and an experimental task investigating the use and distribution of referential expressions in English. The method used in the latter task was narration. In narration tasks participants are asked to describe a visual prompt. The minimal control that researchers exert over participants' responses during these tasks makes them ideal for investigating the use of frequent structures that appear in alternating manifestations in the discourse, such as subject referents (Faitaki & Murphy, 2020). Moreover, narration tasks provide an authentic communication context and are suitable for testing adults who process and retain information faster than children and, thus, would likely grow tired of a more restrained oral elicitation task (e.g. sentence completion).

4.2.2.1. Demographics Questionnaire

The first of the experimental materials that the participants saw was the demographics questionnaire (Appendix P), which was created to collect some important information about their characteristics and language acquisition process. The Greek version of the questionnaire contained questions about the participants' (1) gender (2) age (3) AoO and (4) knowledge of languages other than English and Greek. To answer the first question, participants could select one of the three options on a drop-down menu (i.e., 'male', 'female', 'other'). The second question required participants to input their age as a number. The third question allowed participants to choose among four options (i.e., 'at preschool', 'at primary school', 'at secondary school', 'at/after university'), while the fourth question contained two possible answers: 'yes' or 'no'. If the participants selected 'yes', they were prompted to answer an additional question about which languages other than English and Greek they know. The English version of the questionnaire contained the same questions as the Greek version, except

for the question on AoA, which was redundant. Moreover, the question on knowledge of other languages was rephrased to exclude Greek so as to reflect the English L1 speakers' experience.

4.2.2.2. Exposure Questionnaire

An extensive search on the IRIS database (Marsden, Mackey & Plonsky, 2016) and individual websites revealed that the existing questionnaires on adult learners' L2 exposure seem to be limited in both number and scope. Thus, an exposure questionnaire was created from scratch to meet the needs of the context under investigation.

The questionnaire contained nine questions: the first five aimed to capture the participants' linguistic environment; the latter four aimed to quantify their language use. The 'linguistic environment' questions asked about the participants' country of residence, home and work language, and preferred language for communicating with friends and relatives. To answer these questions, the participants needed to write their response in the box. The 'language use' questions asked about the time that participants spend engaging with the L2 relative to the four skills (listening, reading, speaking and writing) on a normal day (see Paradis, 2011). To answer these questions, the participants were expected to slide a button along a scale. The scale ranged from 'I never hear/read/speak/write (in) English' (which corresponded to 0% of the time) to 'I only hear/read/speak/write (in) English' (which corresponded to 100% of the time). The participants were not able to see the percentages that were associated with the statements or with the slider in order to prevent them from answering based on pre-conceived notions of what each percentage means.

The questionnaire was presented in two versions (which can both be found in Appendix Q): the Greek version was aimed for Greek EFL learners, while the English version was aimed for the L1 English speakers. The two versions asked the same questions but used different phrasing. In particular, the word 'English' in the Greek version was substituted with the words 'other languages' in the English version.

4.2.2.3. Yes/No Vocabulary Test

An adaptation of the classic Yes/No Vocabulary Test (Meara & Buxton, 1987), a measure of vocabulary size, was used to assess the adult participants' L2 proficiency. The use of the measure was justified by the fact that vocabulary size is known to be a reliable predictor of linguistic proficiency, especially among intermediate and advanced learners (Miralpeix & Muñoz, 2018; Staehr, 2008). In general, Yes/No tests constitute a combination of semantic judgement and meaning recall tasks: the participants see a list of words and tick the ones they know ('Yes') while leaving blank the ones they do not know ('No'). Due to their simple format, Yes/No tests allow for the assessment of a large number of words with a large number of people (Pellicer-Sánchez & Schmitt, 2012). As such, Yes/No tests are considered to be time- and resource-efficient for participants and researchers alike (Mochida & Harrington, 2006). Moreover, they have been found to correlate with other tests of vocabulary knowledge, such as the Vocabulary Size Test (Nation & Beglar, 2007), and to be valid and reliable (Beeckmans, Eyckmans & Janssens, 2001; Huibregtse, Admiraal & Meara, 2002).

The version of the test used in the present study consisted of three sets of items. Each set contained 32 items, of which 20 were real words and 12 were pseudowords (i.e., meaningless words that meet the phonological constraints of the language) (see Appendix R for the pseudowords, and Appendix S for the test lists). The sets are based on word frequency levels according to the British National Corpus: the first set was composed using some of the 2000 most frequent words; the second set consisted some of the 3000 most frequent words; and the third set contained some of the 4000 most frequent words. These levels are believed to be appropriate for testing intermediate learners of English and, therefore, were considered appropriate for capturing the participants' (various) proficiency level(s) (Meara, 1992). The pseudowords were created by putting together syllables of real words within the relevant

frequency ranges and judged by native speakers for their consistency with English phonology[15].

The presentation of the items matched the traditional Yes/No Vocabulary Test as much as possible: the Greek EFL and the English L1 participants saw each set as a list and were told to tick off the words they knew from the list. They could select as many of the items as they wanted (the minimum was one, and there was no maximum), and they could also untick items if they changed their mind before moving to the next screen (after moving to the next screen, it was not possible to change their answers).

4.2.2.4. Narration Task

The experimental task focused on the use of referential expressions. To test the alternating forms of referential expressions in one task (and assess whether EFL learners differ from monolinguals in their preference for certain forms in certain contexts), participants were asked to narrate what happens in a series of pictures in a Narration task.

Four comic strips from the series 'Peanuts', written and illustrated by Charles M. Schultz (which can be found in Appendix T), were used for this purpose. As Figure 38 illustrates, the strips were composed of eight to ten panels and had no words (except for sounds and exclamations that were not crucial to the meaning). The strips were self-contained and not connected to one another: each depicted a pair of characters and different events (although they all involved a fight between the characters and its eventual resolution). They were chosen as their character pairs could give rise to varied referring expressions and their plot could be described using vocabulary that the participants would undoubtably be familiar with.

[15] Both the words and the pseudowords that were used in this experiment were generated using the randomizers that are embedded in the Lex Tutor website (https://www.lextutor.ca/rand).

Figure 38. Example picture for Experiment 3.

4.2.3. Procedure

Experiment 3 (and the DPhil as a whole) was approved by Central University Research Ethics Committee (CUREC) of the University of Oxford. However, the participants had to give their consent to participate. Thus, an information sheet and consent form were embedded in the experiment and presented as the first screen that the participants saw upon clicking on the experiment link (Appendix U). To proceed to the experiment, potential participants had to tick a box in order to confirm that they had read the information, met the inclusion criteria, and agreed to take part.

4.2.3.1. Piloting

The experiment involved three rounds of piloting in order to safeguard that the materials and procedures were appropriate. The first round was pivotal for gauging the appropriateness of the materials and was conducted on Google Forms. It involved 12 participants from the target populations (i.e., ten Greek EFL learners and two L1 English speakers). The participants completed the Exposure Questionnaire, a longer version of the Yes/No Vocabulary Test that involved six sets of questions, and a different version of the Narration Task during which they had to watch a short video and narrate the events that took place in the video in a recording (that could last between three and four minutes).

Looking at the participants' responses, it was possible to determine that the measures were well-suited for the purpose of the study: the Exposure Questionnaire provided a good quantitative estimate of participants' experience with English, as well as valuable qualitative information to complement the quantitative estimate; the Yes/No Vocabulary Test was considered to be rather repetitive (and thus, tiring) for most participants; finally, the Narration Task gave rise to substantial instances of referential expressions. Taking these observations forward, the Yes/No test was shortened to three sets (which is considered to be the minimum length for obtaining accurate results according to Meara, 1992).

After the materials were finalised, a second round of piloting was conducted in order to provide information about the technical aspects of the experiment. This round took place on the same platform as the experiment itself (i.e., Gorilla Experiment Builder), and involved another ten participants, of whom seven were Greek EFL learners and three were English L1 speakers. The first part of the second round of piloting revealed that the Narration Task did not work as well on the platform: due to internet connection issues, the video would not load and the audio-recordings would not be uploaded on the server for some participants.

To solve these issues, the video was replaced with the aforementioned comic strips, and the audio-recordings were shortened to one minute per picture. Moreover, participants with connections slower than four MBPS (the minimum connection that allowed audio-recordings to be uploaded on the Gorilla server) were excluded from taking part. With these changes in place, a third round of piloting was conducted with (a different) ten participants, eight of whom were Greek EFL learners and two English L1 speakers. This time, none of the participants encountered technical problems, and thus the experiment was deemed ready for administration.

4.2.3.2. Administration

To administer the adult battery, two web links were generated: one for the 'English' version of the experiment (for the English monolingual controls) and one for the 'Greek' version of the

experiment (for the Greek EFL learners). In both versions, the order of the tasks was the same: the Demographics Questionnaire was administered first, followed by the Exposure Questionnaire; the Yes/No Vocabulary Test was administered after, and the procedure concluded with the Narration Task. The latter involved a microphone check (to ensure that participants' audio-recording devices were operating properly) as well as the experiment itself.

The administration of the questionnaires was straightforward, so no particular instructions were given; this was not the case for the Yes/No Vocabulary Test and the Narration Task. In the Yes/No Vocabulary Test, participants were told to 'check the box next to the word' if they knew what the word meant (and leave it blank if they did not). An additional instruction was that participants should answer honestly and not look up the words while doing the task. In the Narration Task, participants were informed that they would see four comic strips and would be asked to describe what happens in each comic strip in English. They were advised to pay particular attention to the characters and the events that appeared in the strips, and were reminded that the descriptions would be audio-recorded. A more detailed explanation of how the audio-recording would take place was also included for participants' information.

The protocol was short: it took a minimum of 20 and a maximum of 30 minutes to complete. In particular, the Demographics Questionnaire took between one and three minutes, the Exposure Questionnaire between two and five minutes, the Yes/No Vocabulary Test between five and ten minutes and the Narration Task between 12 and 20 minutes. However, to make sure that participants' answers were uploaded on the server one hour was set as the cut-off limit after which participants were excluded from the procedure.

As a token of appreciation, participants were rewarded for their participation. Due to the limited funds that were available for this purpose, it was not possible to reimburse each participant; rather, the total funding was split into three vouchers (worth £50 each). To enter the competition for the vouchers, the participants had to put their email address at the end of

the experiment. The three winners were selected randomly, and were contacted on the email address they provided with information about how to receive their vouchers. Entering an email address threatens the anonymity of the data; thus, to ensure that the responses were not associated with the email addresses (and, thus, the person who provided them), the file containing the email addresses was downloaded separately from the other data files from the server at the end of data collection.

4.2.4. Preparation for Data Analysis

After data collection, the data was prepared for statistical analysis. The information that the participants provided in the questionnaires, test and task was saved on a password-protected folder, according to the CUREC's guidelines. Then, the Demographics Questionnaire, the Exposure Questionnaire, the Yes/No Test and the Narration Task were scored. Prior to scoring, the participants' narrations in the respective task had to be transcribed and coded. These operations are discussed in the following section.

4.2.4.1. Coding & Scoring the Background Measures

As the Demographics Questionnaire contained questions that corresponded to different kinds of variables, it required various kinds of coding. The question on gender gave rise to categorical data was coded using the following scheme: '1' for 'male', '2' for 'female', '3' for 'other' (although no participants identified as 'other'). The question on age gave rise to numerical data and its logging was straightforward. By contrast, answers to the question on AoO (applicable to the Greek EFL participants) were scored as '1' for 'preschool', '2' if 'primary school', '3' if 'secondary school' and '4' for 'university' (however, no participants claimed to have started learning English at university). Answers to the question on knowledge of other languages (i.e., besides Greek and English for Greek EFL learners, and English for English L1 learners) were coded as '1' if 'Yes' was selected, and '2' if 'No' was selected. The information about the

learners' additional FLs was saved to monitor potential CLI effects among the Greek EFL and English L1 speakers alike, but not coded as it was not pivotal for data analysis.

The Exposure Questionnaire contained two parts: its 'linguistic environment' component provided categorical data, while its 'language use' component gave rise to numerical data. The categorical data was coded using the following scheme: '1' for 'same language as group' (i.e., Greek for the Greek EFL learners and English for the English L1 speakers), '2' for 'opposite language than group' (i.e., English for the Greek EFL learners and Greek for the English L1 speakers) and '3' for 'other' (e.g., a language other than Greek or English). The numerical data were logged under four individual variables ('%listening', '%reading', '%speaking', '%writing'), while aggregate scores for '%receptive' (encompassing the listening and reading estimates), '%productive' (encompassing the speaking and writing estimates), and '%exposure' (containing an average of all four variables) were calculated based on the scores of the individual variables.

Deciding how to score the Yes/No Vocabulary Test proved to be challenging, because scoring this particular test is a contentious topic in and of itself (Pellicer-Sánchez & Schmitt, 2012). The 'hits' (i.e., 'Yes' answers to a word) and the 'false alarms' (i.e. 'Yes' answers to a pseudoword) were calculated. Then, a simple formula was used to produce a total score for the test: 'hits – false alarms'. Although more complex formulas have been proposed in the literature, research suggests that the 'hits – false alarms' formula provides a sufficiently valid and reliable estimate of learners' performance in the Yes/No Test: it produces accurate results whether the false alarm rate is high or low, and has strong correlations with both vocabulary tests and reaction time measures that have been used as reference for the Yes/No Test (Pellicer-Sánchez & Schmitt, 2012).

4.2.4.2. Transcribing, Coding and Scoring the Narration Task

As the Narration Task required participants to describe a set of pictures, it yielded a set of audio-recordings that had to be transcribed. The format of the transcriptions was similar to the previous two experiments in that the transcripts marked unintelligible words, short and long pauses, and code-switching (although there were no instances of code-switching in the participants' narrations).

After the audio-recordings were transcribed, they were coded. The process of coding was complex: it involved the identification of obligatory contexts for referential expressions in each narrative, and of the most suitable referring expression for each obligatory context in relation to the rules of the language. According to Brown (1973) obligatory contexts constitute "an acquisition criterion not simply in terms of output but in terms of output-where-required" (296). In effect, each obligatory context that a speaker creates is a test item that he/she can pass or fail (by providing or omitting the required linguistic structure respectively).

The participants were awarded one point for producing a referring expression in an obligatory context, and zero points for omitting it; this procedure allowed for the calculation of omission errors. Then, all produced referential expressions were marked as appropriate or inappropriate given the pragmatic demands of the context; this (qualitative) procedure allowed for the calculation of 'commission' errors or, in other words, the erroneous use of a grammatically appropriate form in a pragmatically inappropriate fashion (e.g., using an overt third person singular pronoun to introduce a new referent to the discourse; for the term see Pierce, Genesse & Paradis, 2012).

To ensure the reliability of the coding, 10% of the transcriptions were coded by an independent rater, who was an English L1 PhD Candidate at the University of Oxford, using the aforementioned coding scheme. Interrater agreement was 79.5%. The score was lower than expected and desired, because of the fact that a single obligatory context often allows for

various referring expressions and, thus, different raters might consider different referring expressions as the most appropriate. All noted discrepancies in the coding were discussed and resolved between the raters, most often by adopting the independent rater's scoring. This decision was motivated by the fact that she was an English L1 speaker (unlike the researcher), and following her intuitions regarding subject use was considered to be more appropriate.

After reaching an agreement about the coding, six scores were generated for each participant: (1) the total number of obligatory contexts (2) the percentage of subject omission (3) the percentage of subject commission (4) the percentage of DPs used in the narratives (5) the percentage of pronouns used in the narratives (6) the percentage of correct subject use across narratives. To calculate the first four percentages, the total number of omissions, commissions, DPs and pronouns was calculated, divided by the total number of obligatory contexts and multiplied by 100. The final percentage (of correct subject use across narratives) was calculated by subtracting the total number of omissions and commissions from the total number of obligatory contexts and multiplying the end-product by 100.

4.3. Results

After the data were prepared, they were analysed so as to provide answers to the RQs. This section of the chapter reports on these conducted analyses. The results obtained for the background measures (i.e., the Exposure Questionnaire and the Yes/No Vocabulary Test) will be presented first followed by the results of the focal measure (i.e. Narration Task) for the Greek EFL and the English L1 participants. The final part of this section will focus on the predictors of subject use for the two groups.

4.3.1. Background Measures

The mean percentages for each of the four skills, as well as the receptive, productive and overall exposure aggregates, for the two groups are presented in Table 16.

Table 16. Mean exposure percentages per group, presented alongside skewness and kurtosis values, and independent samples t-test results comparing the two groups' English use per skill.

	Mean (SE)		SD		Skewness		Kurtosis		Significance (p-value)	Effect size (Cohen's d)
	Greek	English	Greek	English	Greek	English	Greek	English		
% listening	46.90 (4.80)	84.17 (3.33)	25.80	18.23	.33	-.99	-.81	-.07	<.001	1.71
% reading	51.28 (5.80)	92.33 (1.60)	31.22	8.77	-.10	-.61	-1.47	-1.25	<.001	1.66
% speaking	31.28 (5.57)	87.93 (3.28)	30.01	17.94	.73	-1.34	-.74	.22	<.001	2.37
% writing	42.72 (7.13)	97.27 (.78)	38.42	4.26	.39	-1.22	-1.65	.04	<.001	1.95
% receptive	49.09 (4.65)	88.25 (2.30)	25.06	12.58	.33	-.96	-1.20	-.09	<.001	1.94
% productive	37.00 (6.06)	92.60 (1.94)	32.62	10.61	.52	-1.29	-1.28	.21	<.001	2.33
% exposure	43.04 (5.12)	90.43 (2.06)	27.58	11.30	.47	-1.14	-1.11	.05	<.001	2.25

While the categorical data that participants provided in the Exposure Questionnaire (e.g., what language they preferred to use at different environments and/or with different people) was used to safeguard data quality, the numerical data was used for data analysis. As expected, Greek EFL learners and English L1 speakers differed from each other in terms of their (receptive and productive) exposure to English. The mean receptive English score was lower for the Greek EFL learners relative to the English L1 speakers. Moreover, the productive English percentage was lower than its receptive English equivalent for the Greek EFL learners, but higher for the English L1 speakers.

Since the mean percentages for each of the four skills and the aggregate scores were normally distributed per group (as evident by the skewness and kurtosis values), a series of independent samples t-tests was conducted in order to determine whether the differences in the exposure of the Greek EFL learners and their English L1 counterparts were of statistical significance. To account for the numerous t-tests being conducted, a Bonferroni correction was manually applied; as such, the significance threshold was lowered to α=.007. The results of the

independent samples t-tests (also presented in Table 16) highlight that the differences between the groups were indeed significant and represented very large effects.

The Greek EFL learners were also found to differ from L1 English speakers in terms of their English linguistic proficiency. Their scores on the Yes/No Vocabulary Test were lower on each of the three sets and in total, with the largest absolute difference between the groups noted in the third and most difficult set. The groups' scores on the Yes/No Vocabulary Test, based on the 'hits – false alarms' derivation, are presented in detail in Table 17.

Table 17. Mean Yes/No Test scores per group, presented alongside skewness and kurtosis values, and independent samples t-test results comparing the two groups' performance per set and in total.

	Mean (SE)		SD		Skewness		Kurtosis		Significance (p-value)	Effect size (Cohen's d)
	Greek	English	Greek	English	Greek	English	Greek	English		
Set 1 Score	16.00 (.31)	19.07 (.20)	1.65	1.11	-.99	-1.11	.59	.60	<.001	2.17
Set 2 Score	15.55 (.58)	18.20 (.30)	3.06	1.54	-1.16	-.84	.80	-.08	<.001	1.09
Set 3 Score	13.81 (.51)	18.57 (.25)	273	1.19	-.48	-.88	.23	.18	<.001	1.60
Total Score	45.31 (1.14)	55.83 (.66)	6.08	3.30	-.86	-1.08	-.40	.41	<.001	1.49

After dealing with an outlier in the English L1 group (whose scores in the second and third set of the Yes/No Vocabulary Test were over three SDs below the mean and were thus replaced with the minimum score that corresponded to three SDs below the mean), the scores for each set and in total became normally distributed for both groups. Thus, a series of independent samples t-tests was conducted, with a Bonferroni correction lowering the significance threshold to α=.01. As shown in Table 17, the two groups' scores were significantly different and represented large effects.

4.3.2. Subject Use

The Narration Task involved the calculation of six individual scores: (1) the total number of obligatory contexts ('total contexts') (2) the percentage of subject omission ('%omission') (3)

the percentage of subject commission ('%commission') (4) the percentage of DPs used in the narratives ('%DP') (5) the percentage of pronouns used in the narratives ('%pronoun') (6) the percentage of correct subject use across narratives ('%correct'). The scores for each group are presented in the following paragraphs, alongside a comparison between the groups.

4.3.2.1. Greek EFL Learners

In narrating the stories depicted in the four comic strips, the Greek EFL learners created a mean of 36.97 obligatory contexts (SD=10.04). On average, they were 86.90% accurate (SD=6.71) in terms of their subject use. Their mistakes were due to the omission of sentential subjects and to the provision of a subject that was inappropriate for the context. Indeed, the Greek EFL learners had a mean of 3.67% omissions (SD=3.07) and a mean of 9.43% commissions (SD=5.77). A more qualitative look at the learners' commission errors revealed a general tendency to extend the use of DPs in contexts where pronouns would be required.

This tendency is reflected in the Greek EFL learners' use of DPs and pronouns. On average, their narratives consisted of 58.20% DPs (SD=14.62) and of 37.80 pronouns (SD=15.86). In fact, a dependent samples t-test suggested that the mean difference between the percentage of DPs and percentage of pronouns that the learners used, 22.30, was statistically significant and represented a large effect: $t(30)=4.01$, $p<.001$, $d=.73$. The mean scores for each of the Narration Task measures are exhibited in Table 18.

Table 18. Means for Narration Task measures, presented alongside skewness and kurtosis values.

	Mean (SE)	**SD**	**Skewness**	**Kurtosis**
total contexts	39.97 (1.87)	10.04	-.12	-.43
% omission	3.67 (.57)	3.07	.55	-.53
% commission	9.43 (1.07)	5.77	.55	-.70
% DP	58.20 (2.72)	14.62	-.44	.37
% pronoun	37.80 (2.95)	15.86	.47	-.19

% correct	86.90 (1.25)	6.71	-.52	-.79

4.3.2.2. English L1 Speakers

The English L1 speakers created a mean of 44.67 obligatory contexts (SD=16.30) in their narratives. On average, they used the correct subject 90.45% (SD=4.64) of the time. Like their Greek EFL counterparts, the English L1 speakers also made omission and commission errors: their narrations were characterised by a mean of 1.24% omissions (SD=1.76) and a mean of 8.31% commissions (SD=5.05). In this group, the commission errors were not as uniform: while some learners overextended the use of DPs (e.g. for topic maintenance), others overextended the use of pronouns (e.g., for topic shift).

However, there was a clear difference in the average use of DPs and pronouns in the English L1 speakers' narratives: while DPs were used 63.02% of the time (SD=13.88), pronouns were used 35.60% of the time (SD=14.68). In fact, according to the result of a dependent samples t-test, the mean difference between the percentage of DPs and percentage of pronouns, 27.42, was statistically significant and represented a large effect: $t(29)=5.27$, $p<.001$, $d=.96$. The mean scores that English speakers achieved on each of the Narration Task measures are displayed in Table 19.

Table 19. Means for Narration Task measures, presented alongside skewness and kurtosis values.

	Mean (SE)	**SD**	**Skewness**	**Kurtosis**
total contexts	44.67 (2.98)	16.30	-.08	-1.00
% omission	1.24 (.32)	1.76	1.51	1.92
% commission	8.31 (.92)	5.05	.08	-.37
% DP	63.02 (2.53)	13.88	-.86	.27
% pronoun	35.60 (2.68)	14.68	.80	.13
% correct	90.45 (.85)	4.64	-.56	-.01

4.3.2.3. Group Comparisons

To determine whether the potential quantitative differences between the groups were of statistical significance, and thus answer RQ1, a series of statistical analyses were conducted. First of all, there was a statistically significant difference as far the total number of obligatory contexts was concerned: t(59)=2.44, p=.02, d=.63, as presented in Figure 39. This difference, which represented a large effect, indicates that the English L1 participants allowed themselves more opportunities to produce referential expressions in the Narration Task.

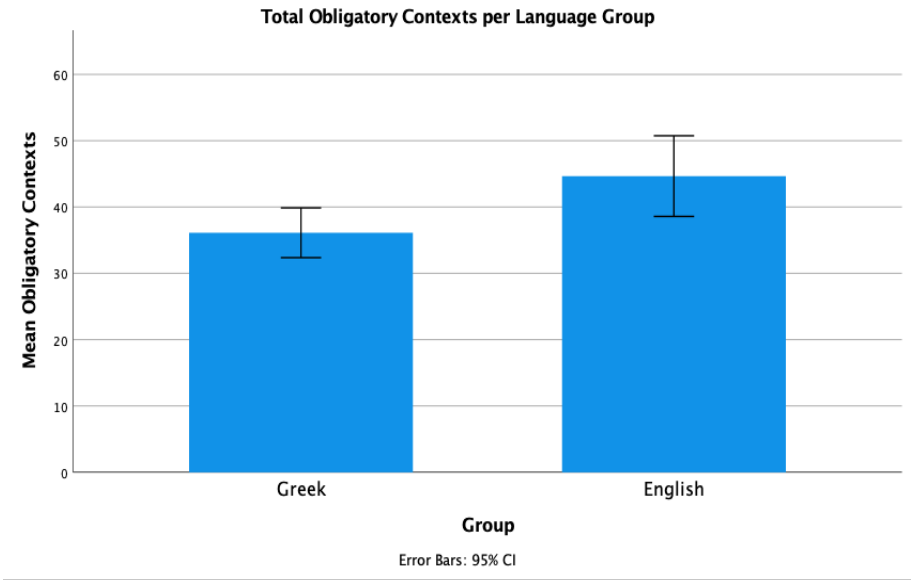


Figure 39. Obligatory contexts provided by Greek EFL learners and English L1 speakers in the task.

A second t-test revealed that there was also a significant difference between the groups' correct uses of referential expressions: t(59)=2.36, p=.02, d=.60, as illustrated in Figure 40. The difference between the groups highlights that the English L1 speakers were more accurate in the realisation and/or the choice of appropriate referential expressions.

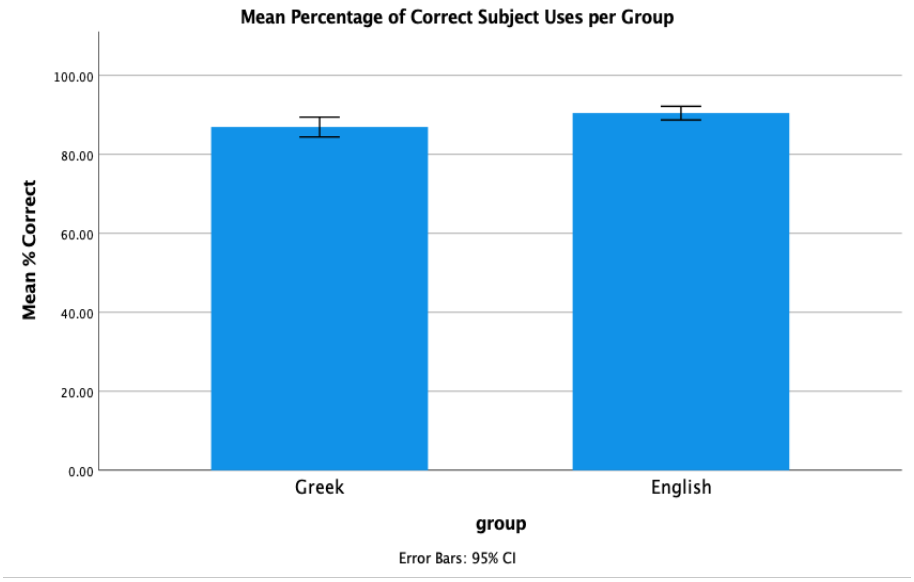


Figure 40. Percentage of correct subject uses in the Greek EFL learners and the English L1 speakers' narratives.

Indeed, to explore whether the difference in the learners' correct uses corresponded to omission or commission errors, two chi-squared tests were conducted. Figure 41 visualises the significant effect that was observed in the case of omissions [$\chi^2(1,61)=4.76$, $p=.03$], such that Greek EFL learners produced fewer overt subjects than their English L1 speaker counterparts. With regards to commissions, no effects were observed: $\chi^2(1,61)=2.07$, $p=.15$.

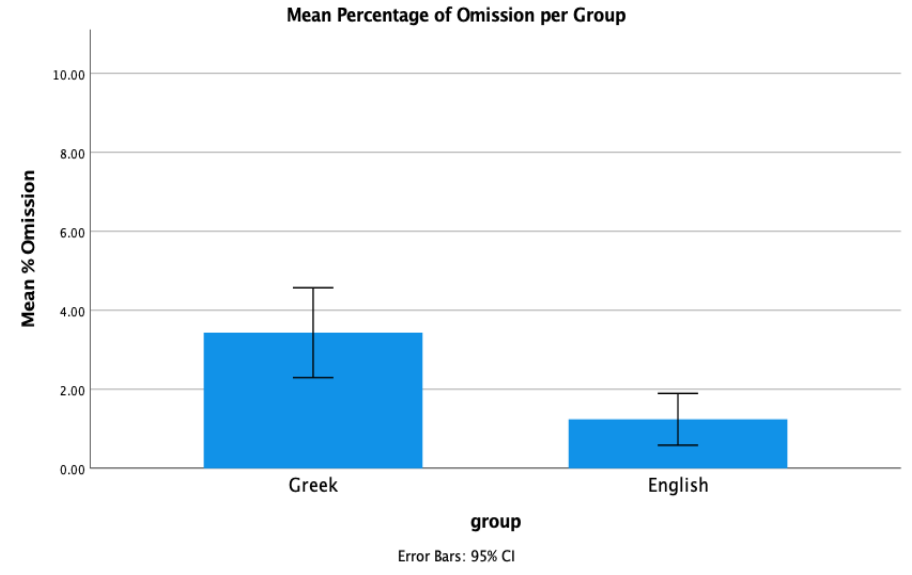


Figure 41. Percentage of omissions in the Greek EFL learners and the English L1 speakers' narratives.

Turning to groups' preference (i.e., most frequent use) of DPs and/or pronouns, there appeared to be no 'qualitative' difference between them [$\chi^2(1,61)=.17$, $p=.68$], since they both had a preference for DPs over pronouns (as previously discussed).

4.3.3. Predictors of Subject (Mis-)Use

Having observed the performance of Greek EFL and English L1 participants in the Narration Task and noted group differences therein, it is appropriate to investigate which factors underlie their use of subjects. The potential factors are explored in the following paragraphs.

4.3.3.1. Correlations Between Variables

To investigate the relationship between the Greek EFL learners' exposure to and proficiency in English and their use of sentential subjects in the task, and thus answer the second research question, a bivariate Pearson's correlation analysis was conducted, the results of which are presented in Table 20. The analysis made use of the learners' receptive and productive exposure estimates in order to explore whether there were differences between the two sets of skills. Moreover, the total vocabulary score was included instead of the individual vocabulary set scores in order to limit the number of variables and, therefore, reduce the amount of bias.

Table 20. Correlations between the variables for the Greek EFL learners.

	% receptive	% productive	total vocab score	total contexts	% omission	% commission	% DP	% pronoun	% correct
% receptive	-	.82***	.63***	-.20	-.07	-.001	.09	-.07	.03
% productive	-	-	.49***	-.15	-.07	-.06	-.03	.04	.08
total vocab score	-	-	-	-.14	.05	.26	.003	-.01	-.25
total contexts	-	-	-	-	-.36**	-.21	-.52***	.60***	.35
% omission	-	-	-	-	-	.02	.09	-.32	-.47***
% commission	-	-	-	-	-	-	.11	-.12	-.89***
% DP	-	-	-	-	-	-	-	-.97***	-.14
% pronoun	-	-	-	-	-	-	-	-	.25
% correct	-	-	-	-	-	-	-	-	-

As expected, there was a large, positive and significant correlation between receptive use and productive exposure, emphasising a connection between how much the learners hear and/or read and how much they speak and/or write in English (see Figure 42). In addition, there were large positive and significant correlations between the learners' receptive and productive exposure and their vocabulary score (see Figures 43 and 44)

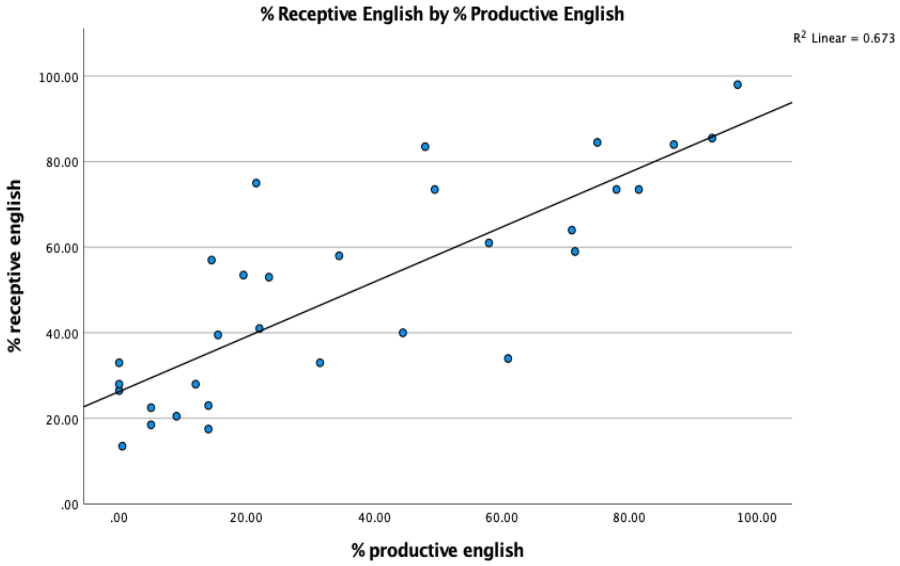


Figure 42. Correlation between the Greek EFL learners' receptive and productive exposure.

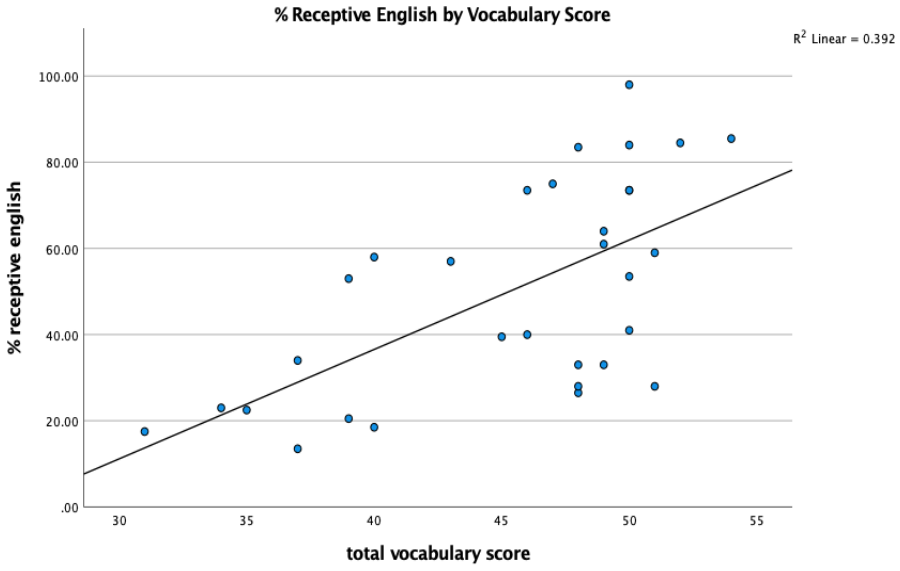


Figure 43. Correlation between the learners' receptive exposure and their vocabulary test scores.

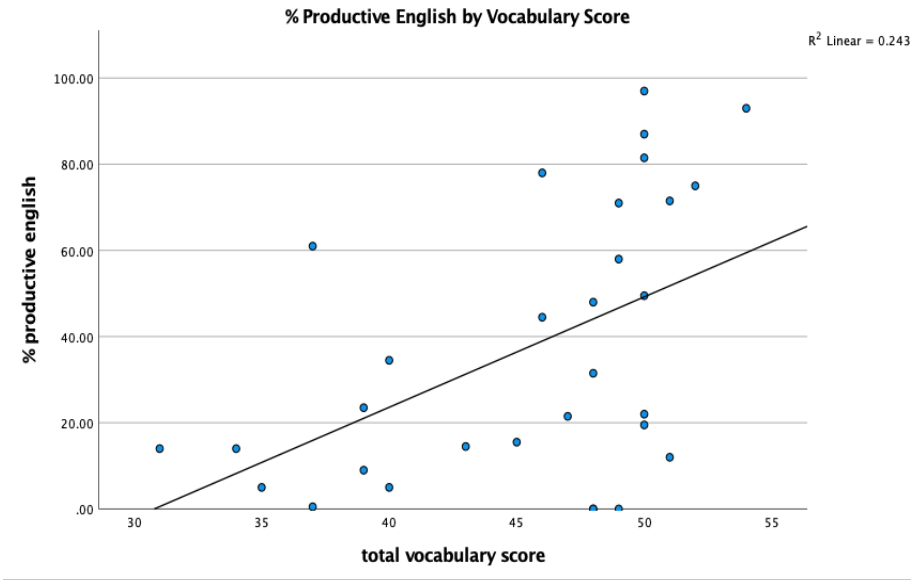


Figure 44. Correlation between the learners' receptive exposure and vocabulary test scores.

Turning to the performance measures themselves: as Figure 45 shows, there was a moderate, negative but significant correlation between the total number of obligatory contexts that the learners provided in their narrations and the percentage of omissions, such that learners whose narrations gave rise to more obligatory contexts for subjects to appear also contained fewer subject omissions. The total number of obligatory contexts was also significantly correlated with the percentage of proper nouns and the percentage of pronouns (see Figures 46 and 47): the former correlation was negative, suggesting that the more obligatory contexts the learners provided, the fewer proper nouns they used; rather, the latter correlation was positive, hinting that the more obligatory contexts the learners provided, the more pronouns they used.

In addition, as Figure 48 illustrates, there was a very large, negative and significant correlation between the percentage of DPs and the percentage of pronouns. This correlation highlights that the two linguistic structures were in complementary distribution in the learners' narratives. In other words, learners who used more DPs did not tend to use pronouns (and vice-versa), often at the cost of making a commission error).

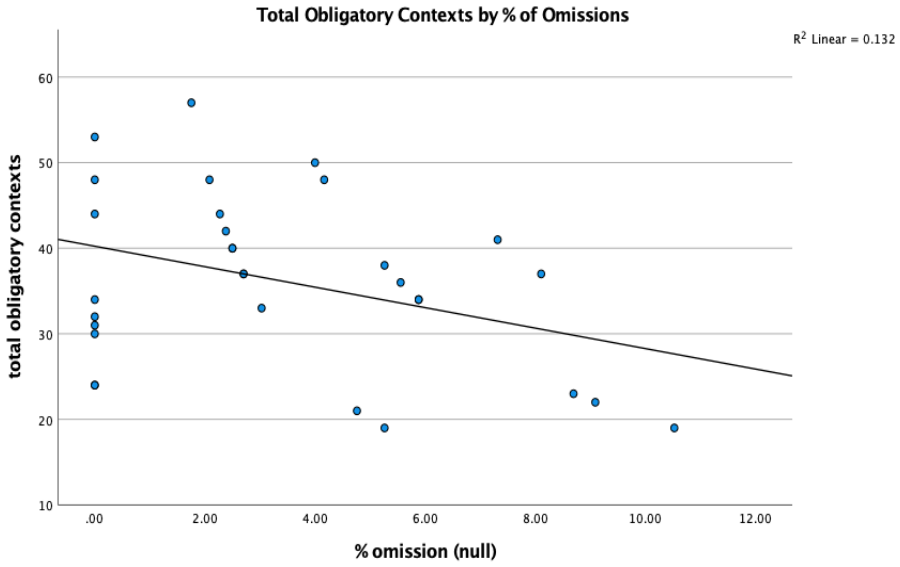


Figure 45. Correlation between the learners' total obligatory contexts and percentage of omissions.

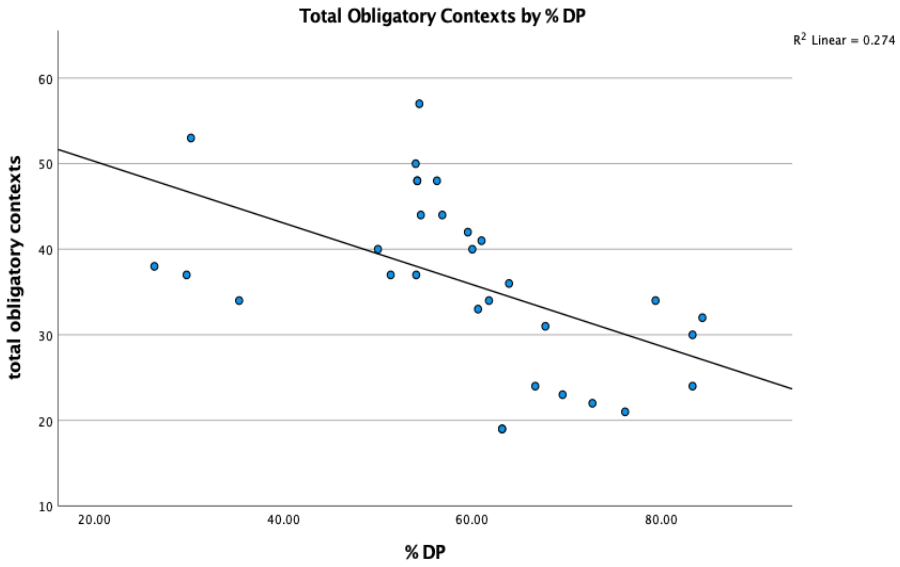


Figure 46. Correlation between the learners' total obligatory contexts and the percentage of DPs in their narratives.

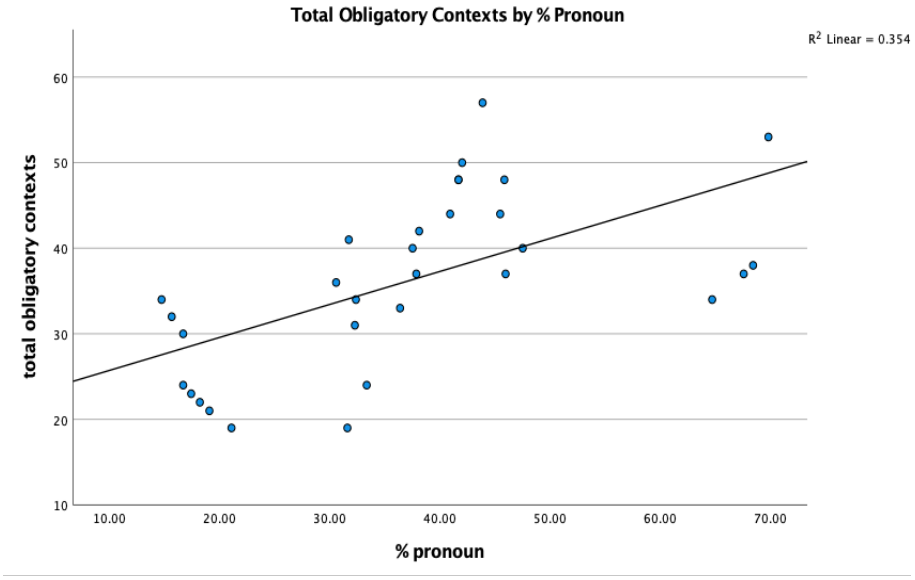


Figure 47. Correlation between the learners' total obligatory contexts and the percentage of pronouns in their narratives.

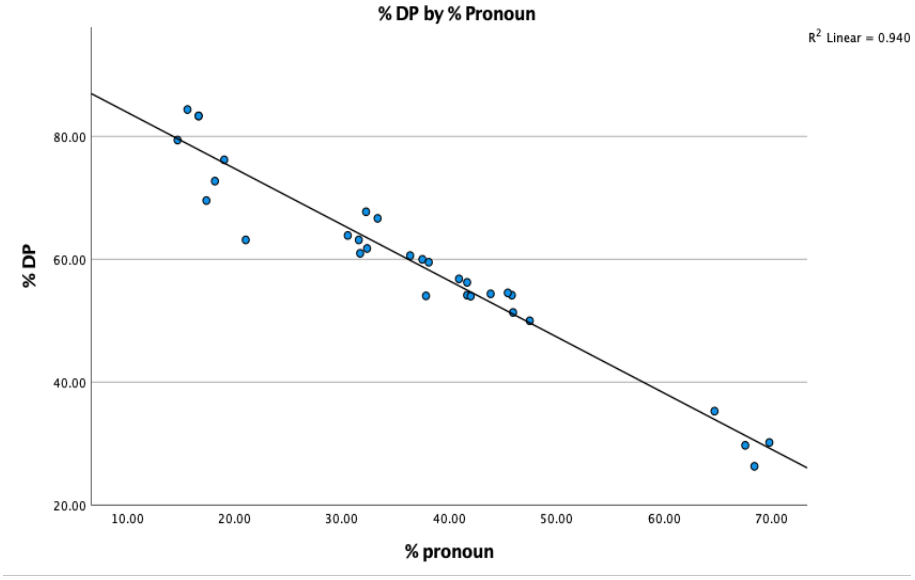


Figure 48. Correlation between the percentage of DPs and pronouns in the learners' narratives.

The final two significant correlations were between the percentage of omission and the percentage of correct subjects, as well as between the percentage of commission and the percentage of correct subjects (presented in Figures 49 and 50 respectively). The former was a moderate negative correlation, while the latter was a large negative correlation. Both indicate the more omissions or commissions the learners' narrations contained, the less correct their subject use was deemed to be. The magnitude of the correlation between commissions and correct use was greater than that between omissions and correct use: this finding highlights that

the Greek EFL learners were more likely to make commission errors than omission errors despite the fact that, as discussed earlier in the chapter, omission errors were the ones that set them apart from English L1 speakers. It should be reminded that the latter were also prone to commission errors.

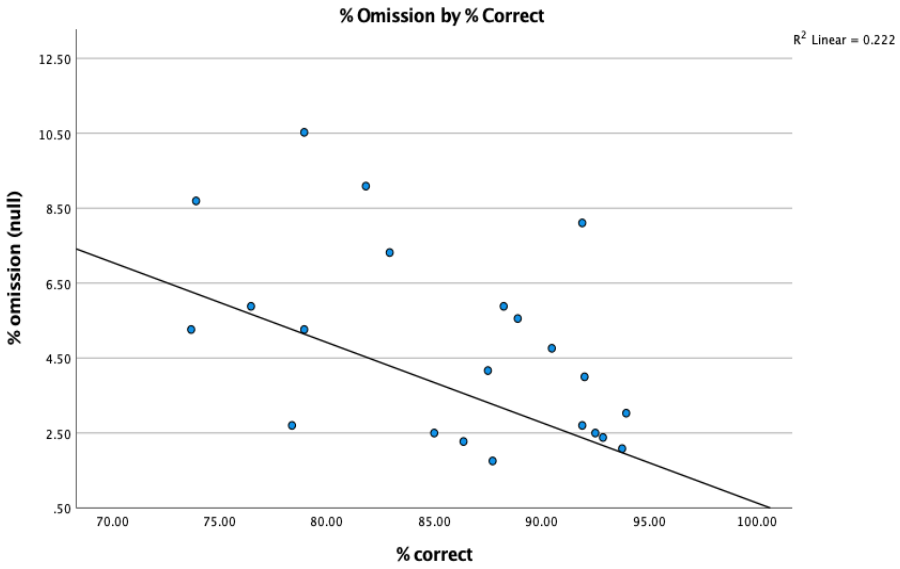


Figure 49. Correlation between the percentage of the learners' omission errors and the percentage of their correct subject uses in the task.

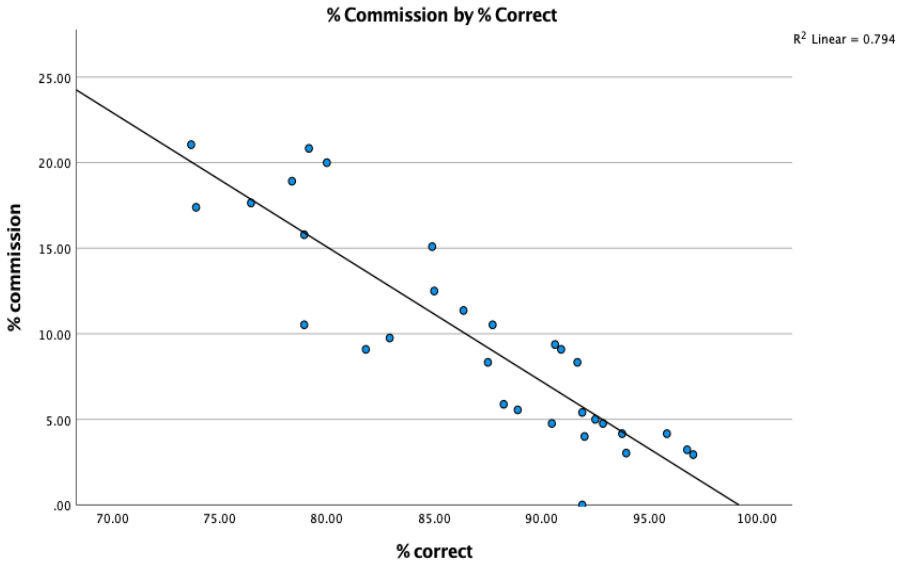


Figure 50. Correlation between the percentage of the learners' commission errors and the percentage of their correct subject uses in the task.

4.3.3.2. Determinants of Performance

Due to the fact that neither exposure nor proficiency were correlated with the Narration Task measures for the Greek EFL learners, alternative (theoretically-motivated) predictors of performance were considered among the variables of the experiment. These included the participants' AoO and L3 knowledge. Yet, AoO could not be used as a predictor since the AoO sub-groups (created based on the participants' responses) did not contain an equal (or, indeed, a reliable) number of observations and, as such, would give rise to questionable results. This was not a problem for L3 knowledge, as the participants were more or less equally divided between those who spoke an L3 and those who did not. Nevertheless, adding L3 knowledge as a predictor was not considered worthwhile, as a correlation analysis between L3 knowledge and the various subject use measures did not detect a relationship between them. In view of these issues, the reasons behind the Greek EFL learners' performance in the Narration Task require further investigation – using predictors that were not included in Experiment 3.

Turning to the English L1 speakers, proficiency (as measured through the Yes/No Vocabulary Test) was the only potential predictor that was considered capable of accounting for the participants' subject omission, subject commission and overall accuracy in the Narration task. The regression model produced to account for the English L1 speakers' omission errors, with vocabulary as the sole predictor, was a good fit; it explained 36.4% of the variance [R^2=.36, $F(1,29)=16.06$, $p<.001$] and led to vocabulary being a significant predictor of subject omission. Likewise, the regression model produced to account for the learners' commission errors was a good fit, explaining 16.3% of the variance (R^2=.16, $F(1,29)=5.44$, $p=.03$) and, as a result, vocabulary was also a significant predictor of subject commission. This is illustrated in Table 21 below.

Table 21. Predictors of omission and commission among English L1 speakers.

	Omission			Commission		
	B	SE B	β	B	SE B	β
Vocabulary	-.32	.08	-.60***	.62	.26	.40*
R^2		.36			.16	

However, the same did not hold for English L1 speakers' overall accuracy. The model was not deemed to be a good fit; it explained 4.4% of the variance and was not statistically significant: R^2=.04, F(1,29)=1.29, p=2.66. Thus, to help explain the learners' overall accuracy, a second model was created. The model included the learners' aggregate exposure score as a predictor (so as to circumvent the issue of collinearity which would arise if receptive and exposure scores were used in the model) in addition to vocabulary. Despite being better, the second model was not a perfect fit either; it explained 18.3% of the variance but was not statistically significant: R^2=.18, F(2,29)=3.02, p=.07. In this model, exposure was a significant predictor[16]. The two models for correct subject use are presented in Table 22.

Table 22. Predictors of correct subject use among English L1 speakers.

	Model 1			Model 2		
	B	SE B	β	B	SE B	β
Vocabulary	-.30	.26	-.21	-.22	.25	-.20
Exposure	-	-	-	.16	.07	.38*
R^2		.04			.14	
F Change in R^2		.14			4.59*	

4.4. Discussion

Experiment 3 was conducted in order to examine whether adult L2 learners use sentential subjects like sequential bilingual children. To this end, the experiment's first question asked whether Greek EFL learners exhibit quantitative and/or qualitative differences in the use of

[16] The VIF value for the second model was <5 (1.02). As such, collinearity was not an issue in the model.

sentential subjects in English relative to English L1 speakers. The prediction associated with this question was that Greek EFL learners would use fewer sentential subjects, make more omission and commission errors and, therefore, be less accurate in their production of sentential subjects overall as compared to the English L1 speakers. It was also anticipated that the EFL learners will over-extend DPs and use them even in inappropriate contexts.

These predictions were in part borne out: there were (quantitative) differences between the two groups in terms of the total number of obligatory contexts produced, the percentage of omission errors and the groups' overall accuracy. In other words, the Greek EFL learners created fewer opportunities to produce subjects in their narratives, and omitted the subject in some of the opportunities that they did create. Consequently, their overall accuracy in terms of subject use was also decreased. However, no differences between the groups with regards to the percentage of commission errors were detected.

Furthermore, the Greek EFL learners were found to have a preference for DPs over pronouns, and overextended the use of the former structure in inappropriate contexts (indeed, the most common commission error in the learners' narratives was the use of a DP in lieu of a pronoun). Yet, the English L1 speakers also displayed a preference for DPs over pronouns and, to a lesser extent, a tendency to overextend the use of DPs in discourse-inappropriate contexts. In fact, the English L1 speakers used even more DPs and fewer pronouns in their narratives relative to their Greek EFL counterparts (albeit this difference is descriptive and not statistical). These findings are in line with previous literature on the use of referential expressions, according to which L2 learners have a tendency for over-explicitness in their narratives and tend to opt for low accessibility markers (i.e., DPs) to refer to entities that are highly accessible and would require high accessibility markers (i.e., pronouns) instead (Chini, 2005; Hendriks, 2003; Leclerq & Lenart, 2013).

Turing to the experiment's second question ('what is the relationship between the Greek EFL learners' (a) use of sentential subjects in English and their exposure to the language, (b) use of sentential subjects in English and their linguistic proficiency and (c) exposure to the language and linguistic proficiency?), the answer was once again mixed. Contrary to the prediction that exposure and proficiency would be correlated with the Greek EFL learners' subject use patterns in the Narration Task, the analysis revealed no significant relationships. Nevertheless, as predicted, receptive and productive exposure to English were both associated with proficiency.

Notably, the finding that exposure and proficiency are not related to subject use goes against numerous studies according to which increases in exposure and/or proficiency lead to improvements in adult learners' performance with respect to the target structure (e.g., Prenza, 2014; Prenza & Tsimpli, 2013; White, 1985). Most of these studies test the comprehension (rather than the production) of sentential subjects. Moreover, they divide learners into proficiency groups (e.g., based on their scores on a test), thus treating the variable as a categorical predictor. Instead, the present investigation treated exposure and proficiency as continuous predictor. Nevertheless, many of the aforementioned studies use more holistic estimates of proficiency, that capture lexical as well as grammatical dimensions of the phenomenon. Rather, the present investigation only targeted vocabulary. Though this decision was underpinned by theoretical and practical reasons, it makes it difficult to assess the true role of proficiency in the learners' performance. Indeed, replications of this experiment should, at a minimum, include the administration of a grammar test.

Turning to the third and final question of the experiment ('what are the factors that explain the Greek EFL learners' use of sentential subjects relative to English L1 speakers?'), the answer differs for Greek EFL learners and English L1 speakers. In short, it was impossible to determine the factors that underlie the Greek EFL learners' performance: the lack of

correlations between linguistic proficiency and/or exposure, and subject use rendered any regression analyses moot. In contrast, the correlations between linguistic proficiency and/or exposure, and subject use for English L1 speakers allowed conducting regression analyses. These revealed that the speakers' vocabulary explained a substantial percentage of the variance in the participants' omission and commission percentages, while receptive and productive exposure were reported to account for the variance in the participants' overall accuracy percentage. This finding suggests that L1 speakers differ in their proficiency but, also, in their exposure to their own mother tongue. This observation could give rise to interesting theoretical and practical discussions about the nature of 'native' and/or 'monolingual' speakers.

This experiment was conducted in order to safeguard that the children's sentential subject omissions in Experiment 1 were not the mere consequence of a developmental phase but, rather, due to CLI. The idea behind Experiment 3 was that, if children's omissions constitute a performance 'glitch' that can be overcome as children grow older (and accumulate exposure to the language), they should not be present in adult L2 learners' speech, as L2 learners are further along the developmental continuum and should have overcome such 'glitches'. The fact that the adult Greek EFL learners did omit the sentential subjects more often than English L1 speakers suggests that the reason behind the omissions is more likely to be CLI from the learners' L1 to their L2.

Yet, traditional accounts of CLI –purporting linguistic structure and/or input quantity to affect the appearance of the phenomenon– would predict the learners' omissions of sentential subjects to decrease as a function of proficiency. This expectation was not met in Experiment 3 (although the reason behind this finding, or lack thereof, could lie in the experiment's methodological limitations, discussed in greater detail in the following chapter). An alternative to traditional accounts would be the usage-based account of CLI sketched out in the previous chapter. This account follows associative learning theories, according to which

the learners' attention to the L2 input is calibrated as a result of their exposure to the L1 (see Ellis, 2019). In this case, it could be argued that Greek EFL learners have learned not to rely on the subject referent in their L1 (due to the fact that overt subjects, especially in their pronominal form, are infrequent and ambiguous in Greek). The learners transfer their L1 knowledge (and attentional cues) to their L2 and, as a result, end up producing fewer overt subjects in English than grammatically appropriate and relative to English L1 speakers. It is only through carefully noticing the L2 distribution and/or explicitly learning the rules that underpin the use of referential expressions that learners can realise and correct their (mis)use of the structure (Schmidt, 1990). This speculative account will be further discussed in the following, and final, chapter.

Chapter 5: Discussion

The thesis focused on the role of crosslinguistic influence (CLI) in Greek preschool children's acquisition of English as a Foreign Language (EFL). CLI, or the influence that one of the bilingual's languages might have on the other, has been under investigation for over three decades (Serratrice, 2013). Research on the phenomenon has revealed that the relative structure of the languages that the bilingual is learning, the amount of input that the bilingual receives in either language, and qualitative features of said input might all be determinants of CLI. However, most investigations focus on either simultaneous bilinguals or second language (L2) learners. It is unclear whether CLI also affects the acquisition of sequential bilinguals and, if so, what are the reasons that underpin the presence of the phenomenon in this population. Answering these questions was the overarching goal of the thesis, which constituted one of the first attempts to explore the phenomenon of CLI from a holistic perspective, in an under-investigated (but ever-growing) population of Greek-English sequential bilingual learners, and by offering a direct comparison between the child and adult learners.

Experiment 1 targeted sequential bilingual children's use of sentential subjects. Sentential subjects are believed to be regulated in the mental interface between syntax and pragmatics, a known locus of CLI effects (Sorace, 2011). Moreover, they are ambiguous in Greek (which, as a null-subject language, allows subjects to remain unarticulated) but unambiguous in English (which, as a non-null-subject language, requires subjects to be articulated). Using an oral language elicitation task, I tested two groups of Greek-English sequential bilingual children in Greece, one attending a total immersion programme (and, thus, receiving around 30 hours of English language input per week) and one attending a partial immersion programme (receiving around 15 hours of English language input per week).

The results of the experiment (presented alongside the research questions and predictions in Table 23) suggested that both groups of bilingual children omitted subjects from

their utterances more often than a group of age-matched English monolingual children, but as often as a group of age-matched Greek monolingual children. Moreover, the total immersion group was more accurate in the production of sentential subjects relative to the partial immersion group in English. These results do not fully support structural linguistic accounts of CLI, but they do underscore the role of input quantity in shaping the bilingual children's performance in the experiment, as the children seemed to perform according to the amount of input they receive in their L2 as a result of the educational programme they are enrolled in.

Table 23. Research Questions, Predictions and Findings Summary of Experiment 1.
Key: confirmed predictions, unconfirmed predictions

	Research Question	Predictions	Greek Assessment Findings	English Assessment Findings
Experiment 1	How does the bilingual children's use of subject pronouns in their two languages differ from that of their monolingual peers?	• Bilinguals expected to produce fewer overt subjects in English than English monolinguals. • No differences expected between bilinguals and Greek monolinguals.	• Bilinguals produced a high number of null subjects, like Greek monolinguals. • Bilinguals produced a low number of overt subjects, like Greek monolinguals.	• Bilinguals omitted more subject pronouns relative to English monolinguals. • Bilinguals omitted more subject DPs relative to English monolinguals.
	How does the bilingual children's use of subject pronouns differ depending on whether they attend the total immersion or the partial immersion preschool programme?	• Children in total immersion expected to differ from children in partial immersion in English (and perform more like the English monolinguals). • No differences between the groups in Greek.	• The bilingual groups produced a similar number of null subjects. • The bilingual groups produced a similar number of overt subjects.	• Children in total immersion produced more subject pronouns than children in partial immersion. • The bilingual groups produced a similar number of subject DPs.
	To what extent does the bilingual children's linguistic proficiency in Greek and English affect their use of the target structure in the two languages?	• Proficiency (vocabulary and grammar) expected to improve the children's use of subject pronouns in English. • No relationship between proficiency and children's use of null subjects expected in Greek.	• No correlation between subject use and vocabulary or grammar. • No correlation between subject use and age. • No significant models and/or predictors of (null) subject use.	• Vocabulary and grammar were both positively correlated with subject use. • No correlation between subject use and age. • Vocabulary a significant predictor of subject pronoun use.

To ascertain whether the children's omissions of sentential subjects were indeed the result of CLI, I designed a second experiment, which focused on the subordinate clause marker 'to'. Although this structure has not been investigated in the bilingual language acquisition literature, it would not be deemed as a locus of CLI, as it appears in a similar form in both English and Greek (Christidis, 2002). Thus, it could function as a 'control' for sentential subjects: if the bilingual children's use of subordinate clause markers did not differ from that of the monolingual children's, the role of CLI (and the importance of linguistic interfaces) in

Experiment 1 would be underscored; if the bilingual children also omitted subordinate clause markers more often than their English monolingual peers, and given the importance of providing a uniform account of omissions in (bilingual) language acquisition, the traditional determinants of the phenomenon (i.e., the structural linguistic and input quantity accounts) would be called into question.

As Table 24 illustrates, testing the same bilingual and monolingual children using a different oral language elicitation task (that involved structural priming) revealed that the bilingual groups also omitted more subordinate clause markers from their utterances relative to the English monolinguals. The difference in omission rates was more pronounced in the partial immersion group, who achieved lower scores than the total immersion group in the English assessment. These omissions cannot be attributed to structural linguistic or input quantity accounts of CLI: since both maintain that CLI arises in the case of syntax-pragmatics interface structures that differ in their cross-linguistic manifestation, neither would deem the structure susceptible to CLI. To maintain the idea that CLI does affect the children's use of subordinate clause markers and, thus, sentential subjects, it can be argued that the omissions result from the transfer of the structures' (cross-)linguistic distributional properties.

Table 24. Research Questions, Predictions and Findings Summary of Experiment 2.
Key: confirmed predictions, unconfirmed predictions

	Research Question	Predictions	Greek Assessment Findings	English Assessment Findings
Experiment 2	Does the bilingual children's use of subordinate clause markers in English and Greek differ from that of their monolingual peers?	• Bilinguals expected to produce (or omit) clause markers at a similar rate as English monolinguals. • Bilinguals expected to produce (or omit) clause markers at a similar rate as Greek monolinguals.	• Bilinguals produced (and omitted) clause markers as often as Greek monolinguals.	• Bilinguals omitted clause markers more often than English monolinguals.
	Do the children in total immersion perform differently from children in partial immersion?	• Children in total immersion expected to differ from children in partial immersion in English (and perform more like the English monolinguals). • No differences between the groups in Greek.	• The bilingual groups produced a similar number of clause markers.	• Children in total immersion omitted clause markers less often (and, thus were more target-like) than children in partial immersion.
	Does the proficiency in English influence their performance and, if so, in what way?	• Proficiency (vocabulary and grammar) expected to improve the children's production of clause markers in English. • Proficiency (vocabulary and grammar) also expected to improve the children's production of clause markers in Greek.	• No correlation between clause marker use and vocabulary or grammar. • No correlation between clause marker use and age.	• Vocabulary and grammar were both positively correlated with clause marker use. • No correlation between clause marker use and age. • Vocabulary a significant predictor of clause marker use.

Furthermore, and given the fact that some of the English monolingual children also omitted sentential subjects from their utterances, it was important to investigate whether the omissions reported in Experiment 1 represent a developmental stage (as opposed to a CLI effect). To this end, I decided to investigate the use of sentential subjects in a group of Greek L1 adults who were EFL learners in order to see if the same patterns emerge at a more progressed stage of L2 acquisition; if the same patterns emerged, the omissions in children's speech could then be argued to represent an effect of CLI, rather than a performance error. Thus, I designed a third experiment which involved a narration task. The results of Experiment 3 (visualised in Table 25) indicated that, much like their child counterparts, the Greek EFL learners omitted the subject argument from their utterances more often than their English L1 counterparts. Moreover, and surprisingly, this effect was not (cor)related to the learners' exposure to and/or lexical knowledge in English.

Table 25. Research Questions, Predictions and Findings Summary of Experiment 3.
Key: confirmed predictions, unconfirmed predictions

	Research Question	Predictions	Findings
Experiment 3	Do (adult?) Greek EFL learners exhibit quantitative differences (in terms of the total number of obligatory contexts, accuracy, omission errors, commission errors) and/or qualitative differences (in terms of a preference for DPs or pronouns) in the use of sentential subjects in English relative to English L1 speakers?	• Greek EFL learners expected to differ from English L1 speakers from a quantitative perspective (i.e., using subjects less frequently, having lower accuracy, making more omission and more commission errors). • Greek EFL learners also expected to differ from English L1 speakers from a qualitative perspective (i.e., overextending the use of DPs).	• Greek EFL learners differed from English L1 speakers in most quantitative respects (i.e., total number of obligatory contexts produced, overall accuracy, percentage of omission errors). • The learners did not differ from English L1 speakers in qualitative respects.
	What is the relationship between the adult Greek EFL learners' (a) use of sentential subjects in English and their exposure to the language? (b) use of sentential subjects in English and their linguistic proficiency? (c) exposure to the language and linguistic proficiency?	• Exposure expected to improve the learners' use of sentential subjects. • Proficiency expected to improve the learners' use of sentential subjects. • Exposure and proficiency expected to be (cor)related.	• No correlation between exposure and subject use. • No correlation between proficiency and subject use. • Positive correlation between exposure and proficiency.
	What are the factors that explain the adult Greek EFL learners' use of sentential subjects relative to English L1 speakers?	• Exposure and vocabulary knowledge expected to affect the Greek EFL learners' sentential subject use. • Vocabulary knowledge also expected to affect the English L1 speakers' performance.	• No significant relationships between exposure/proficiency and subject use for the Greek EFL learners. • Vocabulary a significant predictor of subject use for the English L1 speakers.

In this chapter, I will discuss the results of the three experiments and the tentative conclusions they give rise to in greater detail. In Section 1, I will address the overarching question of the thesis (what is the role of CLI in Greek child and adult learners' acquisition of EFL?) by discussing possible (or, rather, impossible) explanations of each experiment's findings. In Section 2, I will provide an alternative (usage-based) account of the findings that focuses on the distributional properties of the structures under investigation and might be able to account for the observed CLI effects better than traditional accounts of the phenomenon. In Section 3, I will delve into the intricacies of sequential bilingualism while, in Section 4, I will tackle the methodological limitations and shortcomings of the thesis. Section 5 will present the implications and applications that the findings of the thesis have for practice, and Section 6 will summarise the chapter and conclude the thesis.

5.1. The Role of Crosslinguistic Influence

Since the idea that a bilingual's two languages might be separate but systematically interacting became established (Hulk & Müller, 2000), numerous investigations were conducted to

pinpoint the principles that underpin their interaction (Serratrice, 2013). These identified two potential causes of CLI: linguistic structure and input quantity. Each of these accounts comes with its own predictions about what are the conditions that will give rise to CLI, and what is the precise form that CLI will take.

Overall, structural linguistic accounts suggest that learners experience CLI when their languages differ with regards to a linguistic structure, especially when the structure in question is regulated in specific linguistic domains and/or their interfaces. These accounts include the Subset-Superset Condition (Hulk & Müller, 2000), and the Interface Hypothesis (Sorace, 2006; Sorace & Filiaci, 2006; Tsimpli & Sorace, 2006). The Interface Hypothesis further consists of representational accounts (e.g., Serratrice et al., 2009) and processing resources accounts (e.g., Hopp, 2009). The former suggest that bilinguals differ from monolinguals in terms of their mental representations of linguistic structures, while the latter purport that bilinguals have the same mental representations as monolinguals but are not as efficient in accessing them.

Like structural linguistic accounts, input quantity accounts suggest that bilinguals experience CLI when their languages differ in terms of the manifestation of a linguistic structure; but input quantity accounts consider that learners are influenced by the language to which they receive higher exposure, rather than by the language that is necessarily less complex and/or more economical (e.g., Yip & Matthews, 2000).

5.1.1. Experiment 1

The findings of Experiment 1 are in line with input quantity accounts of CLI, which predicted that Greek-English sequential bilingual children would overuse null subjects in English relative to English monolinguals. In addition to the differences between the bilinguals and the monolinguals, the two bilingual groups themselves were found to perform in an input-based gradient: children in total immersion (who receive more input in English) were more accurate in the use of sentential subjects than children in immersion. The role of input quantity in

determining the magnitude of CLI effects has been emphasised in the literature over the years (e.g., Argyri & Sorace, 2007; Serratrice et al., 2009; Sorace et al., 2009). The findings of the present experiment confirm and extend the literature, as they highlight that input quantity determines not only the magnitude but, also, the direction of CLI effects in sequential bilingual children (Daskalaki et al., 2019; Paradis & Blom, 2015; Zdorenko & Paradis, 2011).

By contrast, the data seem to be inconsistent with the predictions of (most) structural linguistic accounts of CLI. For instance, the Subset-Superset Condition (Hulk & Müller, 2000) would predict that the bilingual children use more overt subjects in Greek relative to Greek monolingual children. This prediction has been confirmed in numerous studies carried out across different languages over the years (e.g., Hacohen & Schaeffer, 2007; Serratrice et al., 2004; Paradis & Navarro, 2003), but was not borne out in this one. The reason behind this discrepancy might be related to the fact that the previous studies targeted simultaneous bilingual children, while the present one focused on sequential bilingual children. Indeed, given the results of the experiment, it appears that sequential bilingual children do not behave like simultaneous bilingual children with regards to CLI. This finding is (somewhat) expected, since the sequential bilingual children who participated in the thesis had started to learn English at 3;0 but remained more dominant in Greek, as English was limited to the school context.

The Subset-Superset Condition is not the only structural linguistic account of CLI that has been proposed in the literature. Working under a representational account of CLI, Prenza and Tsimpli (2013) investigated the use of sentential subjects by adult Greek EFL learners. The authors found that the learners accepted and produced more null subjects in English relative to English L1 speakers. They accounted for their findings by explaining that some features of the L2 are inaccessible to the Greek EFL learners, who have acquired the L1 parameters of these features and cannot acquire the (contrasting) L2 parameters after a certain point in time. Given the similarities between the present findings and those obtained by Prenza and Tsimpli (2013),

it could be argued that the sequential bilingual participants cannot acquire the rules that underpin the use of sentential subjects in English because they passed the critical timepoint.

However, reaching this conclusion would disregard the fact that, in the English assessment, there was a large, positive and significant correlation between experimental items (in which children had to produce a pronoun, e.g., 'she') and control items (in which they had to produce a noun phrase, e.g., 'the teacher'). This correlation hints that bilingual children do have a knowledge of the rules that underpin the use of referential expressions in English, and contradicts the idea that the bilingual children's mental representations of referential expressions are different from those of monolingual children.

An explanation along the lines of the processing resources account of CLI (that falls under the structural linguistic umbrella) could be more suitable. Indeed, it is possible that the bilingual children's omissions result from their difficulty to access their mental knowledge representations of the target structure during the task (Sorace, 2006). Of course, the offline nature of the task that Experiment 1 employed does not allow for an accurate assessment of this account's validity in this context. Moreover, the processing resources account has been proposed in the context of advanced adult speakers (Sorace, 2011); thus, it is uncertain whether it would also apply to children.

Another possible explanation, not related to traditional accounts of CLI, is that the children's performance is 'skewed' by performance limitations related to the task's cognitive and production demands (Valian, Hoeffner & Aubry, 1996). For instance, according to Bloom (1990), longer utterances have higher production demands. Hence, an utterance that only contains a subject and a verb will be less difficult to produce than a structure that requires a subject, a verb and an object. It can be assumed that the demands become greater in the case of bilingual children who are completing the task in their L2 and, therefore, also have to inhibit their L1 during speech production (Sorace, 2011). Under this light, the children who took part

in the task might have omitted the subjects in English in performance error, whilst attempting to compute, retrieve and produce the relevant structure.

Furthermore, the children might have considered producing the sentential verb more important than producing the subject, since the verb answered the experimental prompt ('what is happening today?'). This idea is echoed in Allen (2000), who argues that subject realisation is partly determined by the subject's informativeness and accessibility (from the hearer's perspective). Thus, more informative and less accessible items, like topics that are new to the discourse, tend to be realised more often than less informative and more accessible ones in children's speech (see also Allen et al., 2008; Matthews, Lieven & Tomasello, 2006).

Shin and Cairns' (2012) study speaks to this issue. It aimed to explore how children select appropriate (null or overt) referents in accord with the discourse context and, thus, involved 169 L1 Spanish speakers (30 adults and 139 children between the ages of 6;8 and 14;7) who completed an acceptability judgement task. The task included two conditions: the 'same-reference' condition, where the correct referential term would be a null subject, and the 'switch-reference' condition, where an overt subject would have to be used instead. Shin and Cairns (2012) reported that all of their child participants accepted more overt subjects in the 'same-reference' condition as well as more null subjects in the 'switch-reference' condition (also see Sorace et al., 2009). To explain these findings, the authors stipulated that children are not capable of taking their interlocutors' perspective into account in the course of complex tasks, like the interpretation of referential expressions.

Shin and Cairn's (2012) study highlighted that monolingual children also allow the use of null subjects in ungrammatical contexts (also see: Sorace et al., 2009). This finding is replicated in the present experiment, as the data indicated that the English monolingual children omitted sentential subjects from their utterances. Research on the acquisition of sentential subjects has demonstrated that children undergo a phase of subject omissions during L1

acquisition (Bloom, 1990; Valian, 1991). This phase comes to pass, but tends to last longer for children whose L1 allows null subjects than for children whose L1 does not (Valian & Eisenberg, 1996). Then, given that monolingual children also omit sentential subjects, the omissions observed in the sequential bilingual children's speech could represent a developmental trend. This possibility has been suggested in previous studies (e.g., Serratrice et al., 2004; Sorace et al., 2009; Zdorenko & Paradis, 2011). For example, Zdorenko and Paradis (2011) investigated the production of articles in English by sequential bilingual children with different L1s, some of which possessed an article system and others that did not. The authors found that all bilingual children acquired the articles and the rules that constrain their use in the same order as English monolingual children. Moreover, they reported that the effects of L1 transfer were short-lived and limited to children with article-less L1s.

Following Zdorenko and Paradis (2011), if the omissions observed in the bilingual children's productions in Experiment 1 are purely developmental in nature (and not the result of CLI), they should stop naturally for bilingual children in their L2 – even if their L1 provides contrasting and/or ambiguous evidence. To this end, the omissions should be reduced as a function of age. In the present experiment, it was only (lexical) proficiency that significantly predicted children's performance, while age did not play a role. The lack of an effect of age might be due to the restricted age-range of the children who took part in the experiment. Yet, this finding seems to go against the idea that the omissions are due to a performance error or developmental in nature.

5.1.2. Experiment 2

To further explore the reason(s) behind the omissions observed in Experiment 1, and to ascertain whether they stem from the transfer of grammatical properties from the children's L1 to their L2, Experiment 2 tested the same bilingual and monolingual children's use of subordinate clause markers. This structure was used as a control: structural linguistic accounts

would not predict transfer as the structure is not regulated in a mental interface and its form (and function) overlaps in the two languages; since structural overlap is also a prerequisite for input quantity accounts of CLI, they would not predict transfer either. As a result, the omissions that were indeed found in Experiment 2 cannot be explained by traditional accounts of CLI.

So, what is the reason behind the omissions in Experiment 2 (and, by extension, in Experiment 1)? The performance-based explanation proposed earlier could, prima facie, be a suitable candidate. It is possible that children omitted the structure due to the fact that it is not essential to the meaning of the sentence; in other words, children could get their message across even without producing the marker 'to'. In fact, this is understood to be the reason why functional categories, like subordinate clause markers, are often omitted from L2 learners' utterances. For instance, Antonijević et al (2017) used the Sentence Repetition Task (Marinis & Armon-Lotem, 2015) to investigate the linguistic development of 28 Year 3 English-Irish sequential bilingual children in English and Irish and found that participants produced fewer function words in Irish (their L2), and frequently omitted function words from their utterances. Antonijević et al (2017) did not find an effect of age or length of exposure to Irish. However, the reason behind the absence of significant effects might be due to lack of variability in the measures themselves, as the children were of a similar age and had a similar length of exposure to Irish (having all been pooled from the same school). These findings (and issues with variability) are echoed in the present experiment.

Yet, in the present experiment, some of the monolingual English children who took part in the task also omitted subordinate clause markers from their utterances from time to time. This is not a strange finding: L1 speakers are known to sometimes omit function words from their utterances – not unlike L2 learners (Brown, 1973). Specifically with regards to the subordinate clause marker 'to', Kirjavainen et al's (2017) experimental study indicated that L1 English children under the age of 3;0 tend to omit the structure from their utterances. However,

Kirjavainen et al's (2017) results reject the idea that the children's omissions of 'to' are the result of a performance limitation like the one described in the previous paragraphs: indeed, comparing utterances with or without the subordinate clause marker in the speech of their participants, the authors argued that the presence (or absence) of 'to' did not change the mean length of the children's utterance (also see: Kirjavainen & Theakston, 2011). Thus, following Kirjavainen et al (2017), it seems unlikely that performance limitations underlie the children's omissions of subordinate clause markers (of course, to evaluate this point and its validity in the context of the present investigation, it would be appropriate to calculate and compare the children's mean length of utterance).

An alternative explanation could, once again, be sought in developmental grounds. However, for such an explanation to hold, an effect of age on the children's use of the structure would be required. However, age was not found to be related to or predictive of subordinate clause marker (mis)uses in Experiment 2: instead, it was children's lexical proficiency that was found to predict the children's performance (as was the case in Experiment 1). This finding might suggest that the children's experience with the language shapes the appearance of omissions, such that children who receive more input (and/or produce more output, see: Swain & Lapkin, 1995) might develop their proficiency faster and, thus, become more efficient in noticing and producing subordinate clause markers. Indeed, children attending the total immersion programme did not omit the marker from their utterances as often as their peers in the partial immersion programme.

5.1.3. Experiment 3

Experiment 3 focused on a different population, adult Greek EFL learners, and explored their use of sentential subjects further in order to attest whether the same patterns that were observed in the child population emerge. At first glance, the experiments' findings could support representational accounts of CLI, as they tie in with those presented in previous research on

Greek EFL learners (e.g., Prenza & Tsimpli, 2013; Tsimpli & Roussou, 1991; White, 1985). However, adopting these accounts would not capture the EFL learners' outstanding overall performance in the task, as well as the fact that their performance could not be predicted based on their lexical proficiency and/or exposure to the language. Indeed, in previous research, the EFL learners' performance was found to be poor but improvable as a function of proficiency. For instance, Prenza (2014) suggested that all of the Greek EFL learners she tested accepted and produced more illicit null subjects in English, but advanced learners were more target-like than their beginner and intermediate counterparts. Similar observations were made by Hopp (2009), whose investigation of near-native and advanced German L2 learners revealed that proficiency did dampen some of the adverse effects of transfer, such that the near-native speakers tended to be accurate than the advanced learners. Thus, although a processing resources account of CLI can capture the learners' excellent performance in the task, it does not capture the full extent of the present experiment's results. Moreover, the methods used in Experiment 3 were offline; thus, they are unable to speak to the presence of processing effects.

Yet, the results of Experiment 3 do not exclude the possibility that the learners' performance was affected by the limitations of the task itself. In addition to the demands of producing subjects that were discussed earlier, Experiment 3 posited additional difficulties due to its web-based administration. Using a (discourse-appropriate) referential expression entails taking into account the interlocutor's perspective and keeping track of the entities that were referred to in previous discourse (Allen et al., 2008; Graf & Davies, 2014; Matthews et al., 2006). Performing these operations might be more taxing and difficult in a virtual experiment as the interlocutor's perspective is not known (neither is the interlocutor). In addition, the 'pressure' of completing this task in an L2 with different rules regarding subject realisation and referential choice might increase the difficulty of the task further. This explanation would be

able to account for the absence of a proficiency effect: even learners with high levels of proficiency might have struggled to navigate the discourse-related demands of the task.

To counteract the adverse effects of web-based testing, it would be essential to rerun the experiment face-to-face and/or in a task that involves the participants engaging in dialogue with a real conversational partner. Obtaining different results in the replication experiment would confirm the presence of performance limitations in the present experiment. By contrast, obtaining similar results would validate the web-based protocol. A rerun of the experiment could also include an additional comprehension task and/or be coupled with the measurement of reaction times, which can provide an indication of processing speed. The task would shed light on whether the learners' face difficulties with the comprehension of referential expressions and, thus, assess the validity of a processing resources account further. When coupled with accuracy scores on offline tasks, online measures like reaction times can illustrate the difficulties that learners are facing.

5.2. Towards A Usage-Based Account of Transfer

Taken together, the findings of the three experiments suggest that the importance of CLI might not be the same for different populations. In the literature, CLI has been reported to affect simultaneous bilingual children and L2 learners. Its effects on sequential bilingual children, like the ones tested in Experiments 1 and 2, are known to be more limited (e.g., Daskalaki et al., 2019; Hervé et al., 2016; Zdorenko & Paradis, 2011). Indeed, the findings of the first two experiments do not support that traditional determinants of CLI, such as linguistic overlap and/or input quantity, affect the children's use of the target structures. In fact, looking at the child data alone, it is impossible to understand whether the observed effects are due to CLI and are not interpretable just in terms of the amount of English language input that the children have received and, thus, the proficiency levels they have attained. Indeed, the relationships

between the children's task performance and their proficiency scores also seem to point towards this idea.

Nevertheless, looking at the adult data together with the child data allows disambiguating between the phenomena (and underscores the value of comparing the two populations, as was done in the present investigation). Due to their higher cumulative exposure to the L2, as well as their age, the adult learners are further along the L2 developmental continuum relative to their child counterparts. Under this light, the fact that they make the same (omission) mistakes as the child learners seems to suggest that CLI from the learners' dominant language, Greek, to their less dominant language, English. In turn, if CLI is argued to be in operation for the adult participants, it should also be argued to operate in the case of the child participants since a (good) theory of L2 acquisition should account for the observed omissions across the L2 developmental spectrum.

To maintain that CLI did affect the children's results in Experiments 1 and 2, the findings can be evaluated under a usage-based accounts of language acquisition. This account can also capture the patterns observed in the findings of the third experiment – especially if the latter are viewed from the perspective of the learners' associative learning mechanisms (which also fall under the usage-based umbrella). Thus, it could be possible to speculate that the omissions observed in both the children's and the adults' speech do not result from the transfer of specific grammatical properties but, rather, from the transfer of general attentional mechanisms from the L1 that shape the learners' use of the L2 (Shirai, 2018).

To explore this idea further, it is worth observing the asymmetry between the linguistic structures under investigation in Experiments 1 and 2. Sentential subjects proved more problematic for the sequential bilingual participants than subordinate clause markers, and remained problematic even for the EFL participants. The observed asymmetry could indeed

stem from the fact that the production of referential expressions involves the integration of grammatical as well as pragmatic information, while subordinate clause markers do not.

Nevertheless, sentential subjects and subordinate clause markers differ on various other dimensions as well. For instance, the former tend to be salient, frequent and variable in form; rather, the latter are imperceptible, infrequent and stable in form. These characteristics are important in view of usage-based accounts of language acquisition. These do not presuppose that children possess innate knowledge of (grammatical) categories; instead, they suggest that language acquisition involves the storage, analysis and eventual generalisation of (sequences of) words. These processes tend to follow the distributional properties of the input that the children are exposed to (Ambridge et al., 2014; Ambridge & Lieven, 2010; Goldberg, 2006; Wonnacott, et al., 2012; Wonnacott, et al., 2017).

Under this light, the acquisition of salient, frequent and variable structures and/or words should be easier and faster than the acquisition of imperceptible, infrequent and variable ones. Then, due to their characteristics, subordinate clause markers should be more difficult than sentential subjects. This prediction was not borne out in the findings of the present experiments, as the (bilingual) participants appeared to face more problems with the latter than with the former in the English assessment. Yet, it is important to highlight that the children were completing this assessment in their L2; thus, their acquisition might have relied on the distributional properties of these structures in the L1.

Considering the distributional properties of the structures in the L1, it is important to highlight that, in Greek assessment of Experiment 1, both the Greek monolingual and the Greek-English sequential bilingual children did not produce many overt subjects in the comparison condition (where an overt subject was the target structure but, also, the discourse-appropriate option). Daskalaki et al (2019), whose study gave rise to similar results, argued that Greek-English sequential bilingual children are reluctant to produce overt subjects because

the latter are infrequent in children's Greek input. Moreover, the authors highlighted that subjects in pronominal form are ambiguous between a referential and a deictic reading in Greek (Tsimpli & Sorace, 2006). The low frequency and semantic ambiguity of overt (pronominal) subjects in Greek might delay children's acquisition of the structure and influence the Greek-English sequential bilingual children's production of sentential subjects in English.

Turning to the distributional properties of 'to' (and its Greek equivalent, 'na'), the pattern is reversed: although 'to' and 'na' are reasonably frequent and salient in their prepositional and deictic form respectively, they are neither frequent nor salient in their subordinate clause marker form. In their clause marker capacity, these structures appear only after specific verbs (e.g., 'want'). Moreover, in both English and Greek these verbs can also be followed by a noun, offering a direct (and more versatile) competitor structure: 'want+noun' (as opposed to 'want+to+verb'). Yet it could be that similarity between the L1 and L2 provides children with more cumulative exposure to subordinate clause markers and, thus, facilitates their acquisition relative to structures with lower cumulative frequency.

Considering the distributional properties of linguistic structures in children's L1 (or the language in which children have more experience) and the cumulative frequency of the structures in the two languages can, therefore, provide an explanation of the two child experiments' results in tandem. This 'usage-based' account of CLI would accurately predict that Greek-English bilingual children ignore or omit overt subjects in the L2, English, due to the fact that the L1-equivalent structure is infrequent and provides unreliable cues. Moreover, it would accurately predict that the children will perceive and/or produce the subordinate clause markers in English with greater ease than sentential subjects, since the structure (and relevant cues) are available and reliable in the L1.

The usage-based account of CLI and, in particular, associative learning theories would also be able to account for Experiment 3. Associative learning theories suggest that learners'

attention to the input is 'tuned' based on prior experience in the L1, and is likely to be transferred from the L1 to the L2 (Ellis, 2002; 2006; 2019; Ellis, et al., 2014; Ellis & Sagarra, 2010; Wulff & Ellis, 2018).

Ellis and Sagarra's (2010) two experiments illustrate this process among adult L2 learners' acquisition. The authors' first experiment focused on temporal reference (or the means of expressing time) in Latin. It involved 54 L1 English speakers, who had no knowledge of Latin, and were divided in three groups: the first participated in a training session that included a lesson on Latin temporal adverbs and the second participated in a training session that focused on Latin temporal verbs, while the third group did not receive training with either structure. Following training, the participants were tested in a decoding task as well as in a comprehension and a production task. The results suggested that being taught to pay attention to a particular type of structure (i.e., adverb or verb) had a significant effect on the learners' performance in the tasks. During the testing session, the participants in Ellis and Sagarra's (2010) experiment were found to concentrate on the expression of time they had been exposed to during the training session and to ignore the other type. This striking effect was replicated in a second experiment that used the same materials and procedures in a sample of 15 Chinese speakers. The results of Ellis and Sagarra's (2010) experiment reveal that prior use of an L1 shapes the learners' attention to linguistic cues that are found in their input, and has the potential to prevent the acquisition of alternative cues in the L2 input.

The process that the participants of Ellis and Sagarra's (2010) studies (and, also, of the present experiments) seem to undergo is CLI. Under the usage-based account of the phenomenon sketched out in this section, what the L2 learners could be transferring from their L1 to the L2 is language-general attention mechanisms rather than structure-specific mental representations. In other words, a usage-based account of CLI would suggest that both the child and adult L2 learners have 'tuned' their attention based on their Greek input; this leads them

to omit the overt subjects in English due to the lack of convergence between the L1 and L2 cues (and to produce somewhat more subordinate clause markers in English due to the cross-linguistic convergence of cues).

This account seems to have explanatory efficacy in the present context, but would benefit from additional explorations. These should involve additional testing using linguistic structures that are matched in all aspects bar their cross-linguistic convergence, as well as replications with other linguistic combinations (besides English and Greek) that could elucidate the process of CLI further and support or contradict the account.

5.3. A Closer Look into Sequential Bilingualism

Although assessing the validity and reliability of the usage-based account of CLI should involve various learner populations, sequential bilingual children are believed to offer an exciting testing case. These children are first exposed to the L2 after their L1 has already started developing but before it has reached its full potential (Chondrogianni, 2018). Given the special nature of this population, it is not surprising that previous studies have given rise to mixed results about the acquisition process(es) that these children follow. Some report that children who learn their L2 at 3;0 perform on par with simultaneous bilingual children (e.g., Schwartz, Kozminsky & Leikin, 2009; Unsworth, 2013b; Unsworth, Argyri, Cornips, Hulk, Sorace & Tsimpli, 2014). For instance, Schwartz et al (2009) investigated the development of pluralisation in Hebrew-Russian sequential bilingual children (who started acquiring their L2 at 3;0) in a semi-longitudinal design spanning two phases: the first involved 125 bilingual and monolingual children around the age of 7;0, while the second consisted of 119 children around the age of 8;0. Using an oral language elicitation task, the authors found that the bilinguals' acquisition followed a similar trajectory to that of the monolinguals – with irregular plurals being more difficult to acquire than regular plurals over the course of the study. These findings prompted the authors to deem sequential bilinguals close to simultaneous bilinguals.

Nevertheless, other studies find that these children's performance is similar to adult L2 learners (e.g., Granfeldt, Schlyter & Kihlstedt 2007; Meisel, 2008; Paradis & Blom, 2015). This was evidenced in Paradis and Blom's (2015) investigation of the copula 'be' (e.g., 'he is hard-working'). The authors tested 79 children who learned English in a sequential fashion (and had various different L1s) using a grammaticality judgement task as well as an elicitation task. They also divided the participants into two groups based on their Age of Onset (AoO), or the time when their L2 acquisition started: the 'early child L2' group, with a mean AoO of 3;6, and the 'late child L2' group, with a mean AoO of 4;6. The authors reported that all children seemed to use the copula at the same age as monolinguals but, unlike monolinguals, they acquired the structure before other verbal inflections. Furthermore, there was no evidence of development in the children's use – in other words, children's performance was independent of their AoO. This pattern is not attested in simultaneous bilingual children's speech, but it is often found in L2 adults' speech. Given this observation, Paradis and Blom (2015) argued that the children should be characterised as L2 learners.

These findings are in accord with Meisel (2009), who suggests that sequential bilingual children resemble L2 learning adults insofar as they both differ from their simultaneous bilingual counterparts in terms of the acquisition of some linguistic phenomena; these include inflectional morphology (as shown by Paradis & Blom, 2015) but, also, syntactic structures – including referential expressions (as shown by the thesis' experiments). This observation has important implications: as Meisel (2009) argues "L1 learners focus on formal cues, whereas L2 learners tend to rely more strongly on functional cues" (30). In turn, this conclusion resonates with the usage-based account of CLI presented above, according to which child and adult L2 learners are believed to focus on the functional cues offered in their L1 during the course of L2 comprehension or production.

5.4. Limitations

The findings, conclusions and additional observations discussed in the previous sections should be interpreted with the project's limitations in mind. Some of these were discussed earlier in this chapter. One such example is the effect that the web-based administration of Experiment 3 could have had on the participants' use of referential expressions was considered. Indeed, it should be highlighted that referential expressions are context-sensitive structures that require interlocutors to take numerous factors into account.

These include individual differences, like working memory abilities, which were not assessed as part of the relevant experiments. Serratrice and De Cat's (2020) study of 172 children between the ages of 5;0 and 7;0 (87 of whom were learners of English as an Additional Language) highlighted the importance of working memory in informing children's use of referential expressions. The authors aimed to test whether cognitive measures can predict children's referential choice, and to explore whether children's experience with language affects their ability to share their interlocutor's perspective and, subsequently, affect their referential choice. Using numerous measures (including an oral language elicitation task), Serratrice and De Cat (2020) reported working memory, measured through a verbal task, to interact with the learners' proficiency and to result in more target-like performance in the reference task. Moreover, they argued that good working memory abilities might help children make use of corrective feedback they receive from interlocutors in the course of referent use.

These findings highlight the importance of including constructs, like working memory, that can tap into the learners' cognitive abilities in the experimental design. Doing so would have perhaps offered an explanation for the linguistic behaviour of the Greek-English sequential bilingual children in the Greek treatment of Experiments 1 and 2, as well as of the Greek EFL adult participants of Experiment 3. The regression analyses that were conducted to determine which factors are responsible for the use of the target structures in these conditions

did not bear fruit. It is possible that the individual differences could explain (some of) the variance in the obtained results.

Of course, other factors might have impacted the lack of an effect in the correlations and/or regressions; these include issues with the analyses themselves. The power analyses that were conducted to determine the appropriate sample size for each experiment focused on main effects. As correlation and/or regression analyses require more power, it is possible that the experiments were underpowered and, thus, unable to detect interactions and relationships between the focal variables. In addition, there were no item-level analyses that would allow examining whether the learners' performance differed in the items that formed part of each experiment. This was a deliberate decision, as all aspects of items (e.g., individual word frequency and total item length) were carefully controlled, precisely in order to minimise item-level differences. Yet, implementing item-level analyses on the experimental data would be a worthwhile endeavour – not least because it would allow checking whether the design (and controls therein) achieved its purpose.

An alternative reason why the correlation and/or regression analyses did not give rise to significant results could be related to the instruments that were used to derive the measurements that were used in the analyses. For instance, the learners' proficiency was measured using standardised tests of vocabulary and/or grammar. However, the tests' standardisation process involved monolingual, rather than bilingual, populations; this observation renders the use of the tests with a bilingual population problematic. Even though the analyses included the raw (rather than the standardised) scores, it is possible that the tests were not suitable for use in the first place. This observation calls for the creation and/or standardisation of instruments that can be used to accurately assess bilingual learners' lexical and grammatical proficiency in various languages.

Another methodological limitation pertaining to Experiments 1 and 2 is that the English assessment was administered to the sequential bilingual child participants after ten weeks of lockdown due to COVID-19. Following the Greek government's guidelines, eight of these weeks involved some schooling, using a specialised virtual learning platform. However, the schooling lasted fewer hours and using different methods than normal. Children at both the total immersion and the immersion preschool had more or less four hours of school per day. Children in the partial immersion preschool used English for about one of these hours, while children in the total immersion preschool used English throughout the school day. Thus, while the lockdown had the potential to affect all bilingual children's performance, the risk was likely higher for children in partial immersion. Fortunately, it was possible to monitor the potential adverse effect of the lockdown, since pilot data were collected from the same two preschools in 2018, using similar measures. Comparing the children's performance in 2018 and 2020, it can be observed that the scores are lower for vocabulary, but not for grammar or the two linguistic structures under investigation (Appendix V). Nevertheless, these comparisons indicate that the school closures had a negative effect on children's performance and, therefore, the results have to be interpreted with caution.

Another reason why the child participants' results should be interpreted with caution pertains to their socioeconomic status (SES). It is known that SES is an important predictor of linguistic performance and academic achievement, such that children with higher SES tend to outperform those with lower SES in language tasks and school tests (Dalmaijer, Bignardi, Anwyl-Irvine, Smith, Siugzdaite, Uh, Johnson & Astle, 2019; De Cat, 2020; Serratrice & De Cat, 2020). All the child participants were pooled from private schools and had high SES. As such, their results in the first two experiments might not be able to generalise across the population. This issue is not as pressing for the adult participants, as they were recruited using snowball sampling – a method that helps reduce bias in sampling (Rose et al., 2019).

This method also allows recruiting participants with varied characteristics. Based on their answers to the demographics questionnaire, most of the Greek EFL adult participants who completed Experiment 3 had started learning English at primary school (i.e., after the age of 6;0). It was only a handful of the participants who indicated that their acquisition process began at preschool (i.e., after the age of 3;0) and, thus, were identified as sequential bilinguals. If the number of sequential bilingual participants who completed the experiment was higher, it would be possible to assess whether AoO had an impact on the learners' performance in Experiment 3. A (cross-sectional) comparison between adult groups of learners with different AoO would be theoretically valuable (given the collective results of the experiments that form part of the thesis) and interesting to undertake.

5.5. Implications for Practice

Despite their limitations, the experiments shed some light on the nature of omissions in EFL acquisition. Exploring the nature of omissions further would have important implications for research and practice. As research on sequential bilingual children is more limited than research on simultaneous bilingual or L2 learning learners, mapping out the developmental trajectories of this group with respect to specific linguistic structures would be valuable from a theoretical standpoint. In turn, determining what are the processes and strategies that these children use to accomplish the task of language acquisition might help improve the methods and materials that teachers use to introduce them to the L2 (and to develop their L1) and, therefore, would also be valuable from a practical standpoint.

The prominence of English medium programmes is increasing across the world (OECD, 2014). In Greece, where the thesis is set, numerous children are becoming bilingual through attending total immersion or immersion programmes from preschool onwards (Dendrinos et al., 2013). Although these programmes are currently operated in private preschools, English lessons will be introduced in the provision of Greece's state preschools

from September 2021 onwards. Due to this imminent change in national policy, information about the effects of early exposure to a different language (in this case, English) on children's L1 and L2 outcomes, can be beneficial for structuring these lessons so as to maximise children's linguistic potential.

5.5.1. Educational Programmes & Linguistic Outcomes

To explore the effects of early language exposure, it is possible to look at the vocabulary and grammar tests completed as part of the child battery. In general, the results of these tests suggest that the amount of input children receive in each of their languages informs the outcomes they attain. In other words, linguistic exposure is critical in determining bilingual children's lexical and grammatical proficiency (e.g., De Cat, 2020; Hoff, Core, Place, Rumiche, Señor, & Parra, 2012; Thordardottir, 2011). This finding highlights that the design and the implementation of total and partial immersion programmes should consider how much input the children receive in their L1 and their L2 in accord with their (linguistic and/or pedagogical) goals.

As far as children's outcomes in the L2 are concerned, the results of the experiments suggest that children in total immersion might have an (initial) advantage. This finding resonates with Marian, Shook and Schroeder's (2013) study, according to which total immersion-educated children achieved higher gains in L2 language skills than their partial immersion-educated peers in Grade 3, but did not maintain their gains across time: by Grade 5, they were outperformed by the partial immersion-educated participants. Due to the design of the present study, it is not possible to know if the total immersion-educated participants will continue to outperform their partial immersion-educated peers with regards to L2 skills. This observation calls for additional cross-sectional or longitudinal research that assesses the longevity of these effects.

By contrast, as far as L1 outcomes are concerned the data suggest that children in partial immersion have an edge. This finding (also) complements previous research, according to

which children in partial immersion tend to attain higher lexical and grammatical proficiency in their L1 relative to their total immersion-educated peers, as well as to match or surpass age-matched monolinguals in their L2 (for a review, see: Genesee, Lindholm-Leary, Saunders & Christian, 2006). This was evidenced in Hermanto, Moreno and Bialystok's (2012) cross-sectional investigation of 83 L1 English students (50 in Grade 2 and 38 in Grade 5) attending a French monolingual immersion programme in Canada. The authors reported that the children's English vocabulary and grammar scores improved from Grade 2 to Grade 5 as a result of attending the programme. Moreover, comparing their (sequential bilingual) participants to age-matched monolingual participants (tested by Bialystok, Peets & Moreno, 2014), Hermanto et al (2012) suggested that the former surpassed the latter by the end of Grade 5. As the authors' investigation involved older children than those who completed the thesis' experiments, the present findings serve to illustrate that the benefits of immersion programmes can appear from an earlier age than previously reported.

The (differences between the) children's L1 and L2 outcomes are important to consider in the course of designing a curriculum for sequential bilingual children. Since the children's performance in the two languages was found to depend on the quantity of L1 and/or L2 input they received, the most critical step in the process of designing a curriculum is to identify how much exposure to the language(s) is necessary for children to meet their (as well as their parents' and teachers') linguistic goals.

5.5.2. Teaching Approaches

Another important step in the process of designing a suitable curriculum is to decide on the teaching approaches that will allow children to acquire their L2 (whilst also developing their L1, in the case of immersion programmes). The results of the background measures, as well as of the experimental tasks themselves, offer valuable insights to this end.

At present, both the total immersion and partial immersion programmes that are operated in Greece follow a naturalistic approach, such that the children who attend the preschools are expected to pick up the language through mere exposure to input. The importance of input in shaping children's L1 and L2 acquisition has been noted in numerous studies (for L1 acquisition, see: Ambridge et al., 2014; for L2 acquisition, see: Wulff & Ellis, 2018). Given the importance of input, it is not surprising that the children who receive more input in the L2 perform more target-like in the background measures and experimental tasks. If the children's acquisition of sentential subjects and subordinate clause markers proceeded only according to developmental considerations, exposure to input should suffice in order for children to eventually master these structures.

However, as argued earlier on in this chapter, some form of CLI from the learners' L1 to the L2 and in accord with the distributional properties of the L1 might be in place. In that case, as Wulff and Ellis (2018) argue, explicit form-focussed instruction might constitute a useful companion to input. The positive impact of this kind of instruction has been highlighted in multiple studies and aggregated in informative meta-analyses (Goo, Granena, Yilmaz & Novella, 2015; Norris & Ortega, 2000; Spada & Tomita, 2010). This work highlights that 'forcing' L2 learners to focus their attention on specific structures facilitates the acquisition and retention of these structures. Indeed, Schmidt's (1990) Noticing Hypothesis purports that consciously noticing the input allows learners to acquire specific linguistic structures and, thus, advance their acquisition. The Noticing Hypothesis, and the additional attention on specific structures it advocates, might be of particular importance in cases when the L1 input leads learners to develop attention biases that impact their L2 acquisition. Then, explicit form-focussed instruction can work together with implicit learning; following the explicit instruction of a linguistic structure "its use in subsequent implicit processing can update the statistical

tallying of its frequency of usage and probabilities of form-function mapping" (Wulff & Ellis, 2018: 51).

Explicit instruction might have a bigger role to play in cases when the input is minimal. This will be the case in state preschools in Greece after they introduce L2 lessons in their curriculum in September 2021. The L2 lessons are expected to take the form of weekly lessons that will include creative activities led by a qualified English language teacher. Given Greece's established practices surrounding language teaching and learning, it is likely that the creative activities will involve explicit instruction of lexis (and, possibly, grammar). It was through this process that (most of) the Greek EFL learners who completed Experiment 3 learned English at primary school. Since all of them managed to achieve high proficiency in English, it could be argued that targeted practice in the L2 works for primary-level children. More research is required to establish whether it works as well with preschool-level children. This is one of the fascinating opportunities that the educational reform due to take place in Greece offers.

5.6. Conclusions

The thesis added to the literature of CLI as it constituted one of the few investigations of the phenomenon that compared Greek child and adult learners' acquisition of EFL. While CLI has preoccupied researchers for over three decades, there is no definitive answer as to what are the circumstances under which the phenomenon appears. In an effort to pinpoint these circumstances, the thesis comprised three experiments: the first focused on Greek-English sequential bilingual children's use of null and overt subjects in an question-and-answer language elicitation task, the second investigated the same children's use of a less investigated structure, subordinate clause markers, in an explicit correction oral language elicitation task, while the third explored adult Greek EFL learners' production of referential expressions in a narration task. In all three experiments, the Greek child and adult learners were found to differ from their English monolingual peers in that they omitted the target structures more often.

However, the effects were more pronounced in the case of referential expressions (Experiments 1 and 3) than in the case of subordinate clause markers (Experiment 2).

Throughout this chapter, I presented various explanations of these findings. Under more traditional accounts of CLI, the differences between monolinguals and bilinguals in the case of Experiments 1 and 3 could be attributed either to the fact that the latter group's mental representations of subject arguments are not the same as that of the former group (e.g., Prenza & Tsimpli, 2013) or to the fact that the latter group are not as efficient in accessing their mental representations of subject arguments (e.g., Hopp, 2009). Traditional accounts of CLI cannot be used to explain the children's omissions in Experiment 2, as these accounts do not deem the structure CLI-inducing in the first place.

To this end, I proposed an alternative explanation, namely, that the participants' performance is driven by usage-based considerations. These include the specific distributional properties of the structures as well as general cognitive mechanisms that inform learners' attention to linguistic cues and are transferred from the L1 to the L2 (Ellis, 2006). Although this idea has only been investigated in the context of adult L2 acquisition, the similarities between sequential bilingual children and L2 learning adults reported in the present experiments and in previous work (e.g., Meisel, 2008) suggest that it might also account for the children's linguistic behaviour. In a nutshell, this explanation would suggest that CLI plays a role in Greek child and adult learners' acquisition of English insofar as the learners' cognitive mechanisms are tuned to perceive and produce the properties of the structures as acquired in the L1. More research, using different combinations of linguistic structures and learner populations, is required to explore whether this account can explain the criteria that underpin the appearance of CLI.

This is an important endeavour since the backdrop of CLI is language differentiation in the bilingual brain. The question of whether bilingual children's languages are part of one

unified linguistic system or two separate linguistic systems has been the pillar on which bilingualism research was built (Werker & Byers-Heinlein, 2008). The question remains relevant to date. On the one hand, there is substantial psycholinguistic evidence to support that bilingual children are sensitive to the intonational, phonological and grammatical properties of the languages they are exposed to from the first months of life (Byers-Heinlein, 2014; Byers-Heinlein et al., 2010; Genesee & Nicoladis, 2007; Meisel, 2001; Poulin-Dubois & Goodz, 2001; Sebastián-Gallés et al., 2012; Sebastián-Gallés & Bosch, 2002). On the other hand, the notion of translanguaging, according to which the bilingual brain does not contain (let alone differentiate between) named languages (García & Wei, 2014; Wei, 2018; Wei & Ho, 2018), is at the forefront of linguistic (and pedagogical) enquiries.

The experiments presented in the thesis would go against the idea of a unified linguistic system. Whether considered in isolation or together, their results suggest that CLI appears under specific circumstances. The usage-based account of CLI that was proposed in this chapter is tied with associative learning theories, according to which the learners' attention to linguistic cues is shaped by the L1 and transferred to the L2. It would be impossible for the L1 to shape learners' acquisition of the L2 and, thus, for transfer to be substantiated in such a systematic fashion had the learners' two languages existed in a unified linguistic system: if the two languages formed part of a unified linguistic system, the availability (or not) of L1 cues would not necessarily lead to transfer effects and, in turn, the appearance of transfer effects would not be systematic. Instead, the present experiments (as well as the numerous investigations on CLI that have preceded them, irrespective of the theoretical framework they adopted) suggest that child and adult L2 learners alike base their acquisition of the L2 on their existing knowledge of the L1 (regardless of whether this knowledge takes the form of specific linguistic representations or general attention mechanisms). This idea presupposes that the L1

knowledge is separate from (but associated with) the L2 knowledge in the bilingual brain, and gives rise to (testable) a set of systematic criteria that should determine the appearance of CLI.

References

Agouraki, Y., (1991). A Modern Greek complementizer and its significance for UG. *UCL Working Papers in Linguistics, 3*, 1–24.

Ahrenholz, B. (2005). "Reference to Person and Objects in the Function of Subject in Learner Varieties". In H. Hendriks (Ed.). *The Structure of Learner Varieties*. pp. 19-64. Berlin/New York: Mouton de Gruyter.

Alexiadou, A., & Anagnostopoulou, E. (2000). Greek syntax: A principles and parameters perspective. *Journal of Greek Linguistics*, *1*, 171-222.

Allen, S. 2000. A discourse-pragmatic explanation for argument representation in child Inuktitut. *Linguistics*, 38: 483–521

Allen, S., Skarabela, B., & Hughes, M. (2008). "Using corpora to examine discourse effects in syntax". In H. Behrens (Ed.), *Corpora in Language Acquisition Research: Finding Structure in Dat*a. pp. 99-137. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Ambridge, B., Kidd, E., Rowland, C. F., & Theakston, A. L. (2015). The ubiquity of frequency effects in first language acquisition. Journal of Child Language, 42(02), 239-273.

Ambridge, B., & Rowland, C. F. (2013). Experimental methods in studying child language acquisition. *Wiley Interdisciplinary Reviews: Cognitive Science*, *4*(2), 149-168.

Antonijevic, S., Durham, R., & Ní Chonghaile, &. (2017). Language performance of sequential bilinguals on an Irish and English sentence repetition task. *Linguistic Approaches to Bilingualism, 7*(3-4), 359-393.

Ariel, M. (1990). *Accessing Noun Phrase Antecedents*. London: Routledge.

Argyri, E., & Sorace, A. (2007). Crosslinguistic Influence and Language Dominance in Older Bilingual Children. Bilingualism: Language and Cognition, 10(1), 79-99.

Bates, E. & MacWhinney, B. (1989). Functionalism and the Competition Model. In B. MacWhinney and E. Bates (Eds). *The crosslinguistic study of sentence processing.* (pp. 3–73). Cambridge, UK: Cambridge Univ. Press.

Beeckmans, R., Eyckmans, J., Janssens, V., Dufranne, M., & Van de Velde, H. (2001). Examining the Yes/No vocabulary test: Some methodological issues in theory and practice. *Language Testing*, *18*(3), 235-274.

Bialystok, E., Luk, G., Peets, K.F & Yang, S. (2010). Receptive vocabulary differences in monolingual and bilingual children. *Bilingualism: Language and Cognition 13*(4), 525-531.

Bialystok, E., Peets, K. F., & Moreno, S. (2014). Producing bilinguals through immersion education: Development of metalinguistic awareness. *Applied Psycholinguistics*, *35*(1), 177.

Bishop, D. (2003). *Test for Reception of Grammar (TROG-2)*. London: Harcourt Assessment.

Blom, E., Cornips, L., & Schaeffer, J. (2017). Cross-linguistic influence in bilingualism. In E. Blom, L. Cornips and J. Schaeffer (Eds), *Cross-linguistic influence in bilingualism: In honor of Aafke Hulk* (pp. 1-14). Amsterdam/Philadelphia: John Benjamins Publishing Company.

Bloom, P. (1990). Syntactic Distinctions in Child Language. *Journal of Child Language, 17*(2), 343-55.

Bloom, L., Tackeff, J., & Lahey, M. (1984). Learning to in complement constructions. *Journal of Child Language*, *11*(2), 391-406.

Branigan, H., & Pickering, M. (2017). An experimental approach to linguistic representation. *Behavioral and Brain Sciences, 40*, E282.

Brown, R. (1973). *A first language: The early stages.* Cambridge, MA: Harvard University Press.

Bybee, J. (1985). *Morphology: A study on the relation between meaning and form.* Amsterdam/Philadelphia: John Benjamins Publishing Company.

Bybee, J. (1995). Regular morphology and the lexicon. *Language and Cognitive Processes*, *10*(5), 425-455.

Bybee, J. (2010). *Language, cognition, and usage.* Cambridge: Cambridge University Press.

Byers-Heinlein, K. (2014). Languages as categories: Reframing the "one language or two" question in early bilingual development. *Language Learning*, *64*(2), 184-201.

Byers-Heinlein, K., Burns, T. C., & Werker, J. F. (2010). The roots of bilingualism in newborns. *Psychological Science*, *21*(3), 343-348.

Chierchia, G. (1998). Reference to kinds across language. *Natural Language Semantics*, *6*(4), 339-405.

Chini, M. (2005). "Reference to Person in Learner Discourse". In H. Hendriks (Ed.) *The Structure of Learner Varieties*. pp. 65-110. Berlin/New York: Mouton de Gruyter.

Chomsky, N. (1959). On certain formal properties of grammars. *Information and Control*, *2*(2), 137-167.

Chomsky, N. (1981). *Lectures on Government and Binding: The Pisa Lectures*. Berlin/New York: Mouton de Gruyter.

Chomsky, N. (1995). *The Minimalist Program*. Cambridge, MA: The MIT Press.

Christidis, A.F. (1985). Πρόσθετες παρατηρήσεις στις συμπληρωματικές προτάσεις της νέας ελληνικής [Additional observations on the complementary clauses of Modern Greek]. *Studies in Greek Linguistics, 3,* 47-72.

Christidis, A.F. (2002). *Όψεις της Γλώσσας* [*Facets of Language*]. Athens: Nisos Publishing.

Clahsen, H., & Felser, C. (2006). Grammatical processing in language learners. *Applied Psycholinguistics*, *27*(1), 3-42.

Dalmaijer, E., Bignardi, G., Anwyl-Irvine, A., Smith, T. A., Siugzdaite, R., Uh, S., & Astle, D. (2019). Direct and indirect links between children's socio-economic status and education: Pathways via mental health, attitude, and cognition. https://doi.org/10.31234/osf.io/yfn56.

Daskalaki, E., Blom, E., Chondrogianni, V., Paradis, J., & Daskalaki, E. (2020). Effects of Parental Input Quality in Child Heritage Language Acquisition. *Journal of Child Language*, 1-28.

Daskalaki, E., Chondrogianni, V., Blom, E., Argyri, F., & Paradis, J. (2019). Input effects across domains: The case of Greek subjects in child heritage language. *Second Language Research, 35*(3), 421-445.

De Cat, C. (2020). Predicting Language Proficiency in Bilingual Children. *Studies in Second Language Acquisition, 42*(2), 279-325.

Dendrinos, B., Zouganeli, A. & Karavas, E. (2013), *Foreign Language Learning in Greek Schools: European Survey on Language Competences.* Athens: RCeL.

Diessel, H. (2004). *The Acquisition of Complex Sentences*. Cambridge: Cambridge University Press.

Diessel, H. (2019). *The grammar network: How linguistic structure is shaped by language use.* Cambridge, UK: Cambridge University Press.

Dunn, L. M., Dunn, D., Styles, B., & Sewell, J. (2009). *The British Picture Vocabulary Scale - 3rd edition*. *BPVS III*. London: GL Assessment.

Eilers, R. E., & Oller, D. K. (1976). The role of speech discrimination in developmental sound substitutions. *Journal of Child Language*, *3*(3), 319-329.

Eilers, S., Tiffin-Richards, S. P., & Schroeder, S. (2018). Individual differences in children's pronoun processing during reading: Detection of incongruence is associated with

higher reading fluency and more regressions. *Journal of Experimental Child Psychology*, *173*, 250-267.

Ellis, N. C. (2002). Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition, 24*(2), 143–188.

Ellis, N. C. (2006). Selective attention and transfer phenomena in SLA: Contingency, cue competition, salience, interference, overshadowing, blocking, and perceptual learning. *Applied Linguistics, 27*(2), 1–31.

Ellis, N. C. (2019). Essentials of a theory of language cognition. *The Modern Language Journal*, *103*, 39-60.

Ellis, N. C., Hafeez, K., Martin, K. I., Chen, L., Boland, J., & Sagarra, N. (2014). An eye-tracking study of learned attention in Second Language Acquisition. *Applied Psycholinguistics, 35*, 547–579.

Ellis, N. C., & Sagarra, N. (2010). The bounds of adult language acquisition: Blocking and learned attention. *Studies in Second Language Acquisition, 32*(4), 553–580.

Enever, J. (2011). *ELLiE: Early Language Learning in Europe: Evidence from the ELLiE study*. British Council.

Eurydice, (2017). *Key Data on Teaching Languages at School in Europe – 2017 Edition. Eurydice Report*. Luxembourg: Publications Office of the European Union.

Faitaki, F., Hessel, A., & Murphy, V. A. (2020). "Vocabulary and Grammar Development in Young Learners of English as an Additional Language". In M. Schwartz (Ed). *International Handbook of Early Language Education*, (pp. 1-21). Dodrecht: Springer.

Faitaki, F. & Murphy, V.A. (2020). "Oral elicitation tasks in language acquisition research", in H. Rose and J. McKinley (Eds.). *The Routledge Handbook of Research Methods in Applied Linguistics* (pp. 360-369). London: Routledge.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175-191.

Fiotaki, A. (2014). *Study of –na and –oti complementizers in Modern Greek: an LFG/XLE treatment*. Master Thesis. University of Athens.

García, O., & Wei, L. (2014). *Translanguaging: Language, bilingualism and education*. Palgrave Pivot, London.

Genesee, F. (1989). Early bilingual development: One language or two. *Journal of Child Language, 16*, 161-179.

Genesee, F. (2009). Early childhood bilingualism: Perils and possibilities. *Journal of Applied Research on Learning, 2*(2), 1-21.

Genesee, F. (2016). "North America: Rethinking early childhood education for English language learners: the role of language". In V. A. Murphy & M. Evangelou (Eds). *Early Childhood Education in English for Speakers of Other Languages*. pp. 21-42. London: The British Council

Genesee, F., Lindholm-Leary, K., Christian, D., Saunders, W., & Saunders, B. (2006). *Educating English language learners: A synthesis of research evidence*. Cambridge University Press.

Genesee, F., & Nicoladis, E. (2007). "Bilingual First Language Acquisition". In E. Hoff, & M., Shatz, (Eds). *Blackwell Handbook of Language Development*. pp. 324-342. Malden, MA; Oxford: Blackwell Pub.

Genesee, F., Nicoladis, E., & Paradis, J. (1995). Language differentiation in early bilingual development. *Journal of Child Language, 22*(3), 611-631.

Goldberg, Adele E. 2006. *Constructions at work: The nature of generalization in language.* Oxford: Oxford University Press.

Goo, J., Granena, G., Yilmaz, Y., & Novella, M. (2015). "Implicit and explicit instruction in L2 learning". In P. Rebuschat (Ed.), *Implicit and explicit learning of languages*. pp. 443–482. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Graf, E., & Davies, C. (2014). "The production and comprehension of referring expressions". In D. Matthews (Ed.). *Pragmatic Development in First Language Acquisition*. pp. 161-182. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Granfeldt, J., Schlyter, S. & Kihlstedt, M. (2007). French as cL2, 2L1 and L1 in pre-school children. *PERLES: Petites Études Romanes de Lund, 21*, 6–43.

Hacohen, A., & Schaeffer, J. (2007). Subject realization in early Hebrew/English bilingual acquisition: The role of crosslinguistic influence. *Bilingualism: Language and Cognition, 10*(3), 333.

Hendriks, H. (2003). "Using Nouns for Reference Maintenance: A Seeming Contradiction in L2 Discourse". In A. Giacalone Ramat (Eds.). *Typology and Second Language Acquisition*. pp. 291-326. Berlin/New York: Mouton de Gruyter.

Hermanto, N., Moreno, S., & Bialystok, E. (2012). Linguistic and metalinguistic outcomes of intense immersion education: How bilingual?. *International Journal of Bilingual Education and Bilingualism*, *15*(2), 131-145.

Hervé, C., Serratrice, L., & Corley, M. (2016). Dislocations in French–English bilingual children: An elicitation study. *Bilingualism: Language and Cognition, 19*(5), 987-1000.

Hoff, E., Core, C., Place, S., Rumiche, R., Señor, M., & Parra, M. (2012). Dual language exposure and early bilingual development. *Journal of Child Language, 39*(1), 1-27.

Hopp, H. (2009). The syntax–discourse interface in near-native L2 acquisition: Off-line and on-line performance. *Bilingualism: Language and Cognition*, *12*(4), 463-483.

Huibregtse, I., Admiraal, W., & Meara, P. (2002). Scores on a yes-no vocabulary test: Correction for guessing and response style. *Language Testing*, *19*(3), 227-245.

Hulk, A., & Müller, N. (2000). Bilingual first language acquisition at the interface between syntax and pragmatics. *Bilingualism: Language and Cognition, 3*(3), 227-244.

Jackendoff, R. (2002). *Foundations of language: Brain, meaning, grammar, evolution*. Oxford: Oxford University Press.

Joseph, B.D. (1981). On the synchrony and diachrony of Modern Greek na. *Byzantine and Modern Greek Studies, 7.* 139-154.

Kang, J. Y. (2004). Telling a coherent story in a foreign language: Analysis of Korean EFL learners' referential strategies in oral narrative discourse. *Journal of Pragmatics, 36*(11), 1975-1990.

Kirjavainen, M., Lieven, E. V. M., & Theakston, A. L. (2017). Can Infinitival to Omissions and Provisions Be Primed? An Experimental Investigation Into the Role of Constructional Competition in Infinitival to Omission Errors. *Cognitive Science, 41*(5), 1242-1273.

Kirjavainen, M., Theakston, A., Lieven, E., & Tomasello, M. (2009). 'I want hold Postman Pat': An investigation into the acquisition of infinitival marker 'to'. *First Language*, *29*(3), 313-339.

Kirjavainen, M., & Theakston, A. (2011). Are infinitival to omission errors primed by prior discourse? The case of WANT constructions. *Cognitive Linguistics,22*(4), 629-657.

Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. *Behavior Research Methods*, *44*(4), 978-990.

Kupisch, T. (2012). Specific and generic subjects in the Italian of German-Italian simultaneous bilinguals and L2 learners. *Bilingualism: Language and Cognition, 15*(4), 736-756.

Langacker, R. W. (1987.) *Foundations of cognitive grammar.* Stanford, CA: Stanford University Press.

Leclercq, P., & Lenart, E. (2013). Discourse cohesion and accessibility of referents in oral narratives: A comparison of L1 and L2 acquisition of French and English. *Discours: Revue de linguistique, psycholinguistique et informatique*, 12, 1-33.

Lenneberg, E. H. (1967). The biological foundations of language. *Hospital Practice*, *2*(12), 59-67.

MacWhinney, B. (2000). *The CHILDES Project: Tools for Analyzing Talk* (3rd Edition). Mahwah, NJ: Lawrence Erlbaum Associates.

Marian, V., Shook, A., & Schroeder, S. R. (2013). Bilingual two-way immersion programs benefit academic achievement. *Bilingual Research Journal*, *36*(2), 167-186.

Marinis, T. (2003). *The acquisition of the DP in Modern Greek*. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Marinis, T., & Armon-Lotem, S. (2015). "Sentence repetition". In S. Armon-Lotem, J. de Jong, & N. Meir (Eds.). *Assessing multilingual children: Disentangling bilingualism from language impairment*. pp. 95-24. Bristol: Multilingual Matters.

Marsden, E., Mackey, A., & Plonsky, L. (2016). *Advancing methodology and practice: The IRIS repository of instruments for research into second languages.* New York, NY: Routledge.

Mastropavlou, M., & Tsimpli, I. M. (2010). Complementizers and subordination in typical language acquisition and SLI. *Lingua*, *121*(3), 442-462.

Matthews, D., Lieven, E., & Tomasello, M. (2006). How toddlers and preschoolers learn to uniquely identify referents for others: A training study. *Child Development*, *78*(6), 1744-1759.

Meara, P. M. (1992). *EFL vocabulary tests*. New York: ERIC Clearinghouse.

Meara, P., & Buxton, B. (1987). An alternative to multiple choice vocabulary tests. *Language Testing*, *4*(2), 142-154.

Meisel, I. M. (2001). "The simultaneous acquisition of two first languages". In J. Cenoz, & F. Genesee (Eds). *Trends in bilingual acquisition*. pp. 11-42. Amsterdam; Philadelphia: John Benjamins Publishing Company.

Meisel, J. M. (2008). "Child second language acquisition or successive first language acquisition". In B. Haznedar and E. Gavruseva (Eds.) *Current Trends in Child Second Language Acquisition: A Generative Perspective* (pp. 55-80), Amsterdam/Philadelphia: John Benjamins Publishing Company.

Menard S. Applied Logistic Regression Analysis. 2nd edition. SAGE Publications, Inc; 2001.

Miralpeix, I., & Muñoz, C. (2018). Receptive vocabulary size and its relationship to EFL language skills. *International Review of Applied Linguistics in Language Teaching*, *56*(1), 1-24.

Mochida, K., & Harrington, M. (2006). The Yes/No test as a measure of receptive vocabulary knowledge. *Language Testing*, *23*(1), 73-98.

Muñoz, C., & Singleton, D. (2011). A critical review of age-related research on L2 ultimate attainment. *Language Teaching*, *44*(1), 1.

Murphy, V.A. (2014). *Second Language Learning in the Early School Years Trends and Contexts*. USA: Oxford University Press.

Murphy, V.A., Evangelou, M., (2016). *Early Childhood Education in English for Speakers of Other Languages*. London: British Council.

Nation, I. S. P., & Beglar, D. (2007). A Vocabulary Size Test. *The Language Teacher 31*(7), 9-13.

Nicoladis, E. (2012). Cross-linguistic influence in French–English bilingual children's possessive constructions. *Bilingualism: Language and Cognition*, *15*(2), 320-328.

Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning, 50*, 417–528.

OECD. (2014). *Education at a Glance 2014: OECD Indicators*. OECD Publishing.

Papadopoulou, D., Peristeri, E., Plemenou, E., Marinis, T., & Tsimpli, I.M. (2014). Pronoun ambiguity resolution in Greek: Evidence from monolingual adults and children. *Lingua,* 155, 98-120.

Paradis, J. (2011). Individual differences in child English second language acquisition Comparing child-internal and child-external factors. *Linguistic Approaches to Bilingualism*, *1*(3), 213–237.

Paradis, J., & Blom, E. (2015). Do early successive bilinguals show the English L2 pattern of precocious BE acquisition? *Bilingualism*, *19*(03), 630–635.

Paradis, J., & Genesee, F. (1996). Syntactic Acquisition in Bilingual Children: Autonomous or Interdependent? *Studies in Second Language Acquisition, 18*(1), 1-25.

Paradis, J., & Navarro, S. (2003). Subject realization and crosslinguistic interference in the bilingual acquisition of Spanish and English: What is the role of the input? *Journal of Child Language, 30*(2), 371-393.

Pellicer-Sánchez, A., & Schmitt, N. (2012). Scoring Yes–No vocabulary tests: Reaction time vs. nonword approaches. *Language Testing*, *29*(4), 489-509.

Philippaki-Warburton, I., & Veloudis, I. (1985). Η υποτακτική στις συμπληρωματικές προτάσεις [The subjunctive in complement clauses]. *Studies in Greek Linguistics*, *5*, 87-104.

Pierce, L. J., Genesee, F., & Paradis, J. (2012). Acquisition of English grammatical morphology by internationally adopted children from China. *Journal of child language*, *40*(5), 1076-1090.

Pinto, M. (2007). “Subject pronouns in bilinguals: interference or maturation”. In L. Escobar & V. Torrens (Eds.), *The acquisition of syntax in Romance languages*. pp. 331-352. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Postal, P. M. and Pullum, G. K. (1978) Traces and the description of English complementizer contraction. *Linguistic Inquiry, 9*, 1–29.

Poulin-Dubois, D., & Goodz, N. (2001). Language differentiation in bilingual infants: Evidence from babbling. *Trends in Bilingual Acquisition*, *1*, 95-106.

Prenza, A.I. (2014). Can Greek Learners Acquire the Overt Subject Property of English? A Pilot Study. *Theory and Practice in Language Studies, 4*(9), 1770-1777.

Prentza, A., & Tsimpli, I. M. (2013). The interpretability of features in second language acquisition: Evidence from null and postverbal subjects in L2 English. *Journal of Greek Linguistics, 13*(2), 323-365.

Pullum, G. K. (1982) Syncategorematicity and English infinitival *to*. *Glossa, 16*, 181–215.

Redlinger, W. E., & Park, T. Z. (1980). Language mixing in young bilinguals. *Journal of Child Language, 7*(2), 337-352.

Rixon, S (2013) *British Council survey of policy and practice in primary English language teaching worldwide.* London: The British Council.

Rose, H., McKinley, J., & Biggs Baffoe-Djan, J. (2020). *Data collection research methods in applied linguistics*. London: Bloomsbury Academic.

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11,*129-158.

Schmitz, K., & Müller, N. (2008). Strong and clitic pronouns in monolingual and bilingual acquisition of French and Italian. *Bilingualism: Language and Cognition*, *11*(1), 19-41.

Schmitz, K., Patuto, M., & Müller, N. (2012). The null-subject parameter at the interface between syntax and pragmatics: Evidence from bilingual German–Italian, German–French and Italian–French children. *First Language*, *32*(1), 205-238.

Schwartz, B. D. (2004). On child L2 development of syntax and morphology. *Lingue e Linguaggio*, *3*(1), 97-132.

Schwartz, M., Kozminsky, E. & Leikin, M. (2009). Delayed acquisition of irregular inflectional morphology in Hebrew in early sequential bilingualism. *International Journal of Bilingualism, 13*(4), 501–522.

Sebastián-Gallés, N., Albareda-Castellot, B., Weikum, W. M., & Werker, J. F. (2012). A bilingual advantage in visual language discrimination in infancy. *Psychological Science*, *23*(9), 994-999.

Sebastián-Gallés, N., & Bosch, L. (2002). Building phonotactic knowledge in bilinguals: Role of early exposure. *Journal of Experimental Psychology*. *28*(4), 974.

Serratrice, L. (2007). Cross-linguistic influence in the interpretation of anaphoric and cataphoric pronouns in English–Italian bilingual children. *Bilingualism: Language and Cognition*, *10*(3), 225-238.

Serratrice, L. (2013). Cross-linguistic influence in bilingual development: Determinants and mechanisms. *Linguistic Approaches to Bilingualism. 3*(1), 3-25.

Serratrice, L., & De Cat, C. (2020). Individual differences in the production of referential expressions: The effect of language proficiency, language exposure and executive function in bilingual and monolingual children. *Bilingualism: Language and Cognition, 23*(2), 371-386.

Serratrice, L., Sorace, A., Filiaci, F., & Baldo, M. (2009). Bilingual children's sensitivity to specificity and genericity: Evidence from metalinguistic awareness. *Bilingualism: Language and Cognition*, *12*(2), 239-257.

Serratrice, L., Sorace, A., Filiaci, F., & Baldo, M. (2012). Pronominal objects in English–Italian and Spanish–Italian bilingual children. *Applied Psycholinguistics, 33*(4), 725-751.

Serratrice, L., Sorace, A., & Paoli, S. (2004). Cross-linguistic influence at the syntax–pragmatics interface: Subjects and objects in Italian–English bilingual and monolingual acquisition. *Bilingualism: Language and Cognition, 7*(3), 183–205.

Shin, N. L., & Cairns, H. S. (2012). The development of NP selection in school-age children: Reference and Spanish subject pronouns. *Language Acquisition*, *19*(1), 3-38.

Shirai, Y. (2018). *Connectionism and second language acquisition*. NY/London: Routledge.

Simos, P. G., Sideridis, G. D., Protopapas, A., & Mouzaki, A. (2011). Psychometric evaluation of a receptive vocabulary test for Greek elementary students. *Assessment for Effective Intervention*, *37*(1), 34-49.

Slabakova, R. (2008). *Meaning in the Second Language*. Berlin/New York: Mouton De Gruyter.

Sorace, A. (2006). Gradience and optionality in mature and developing grammars. In G. Fanselow, C. Fery, M. Schlesewsky, & R. Vogel (Eds.), *Gradience in Grammar: Generative Perspective* (pp. 106-123). Oxford: Oxford University Press.

Sorace, A. (2011). Pinning down the concept of "interface" in bilingualism. *Linguistic Approaches to Bilingualism*, *1*(1), 1-33.

Sorace, A., & Filiaci, F. (2006). Anaphora Resolution in Near-Native Speakers of Italian. *Second Language Research, 22*(3), 339-368.

Sorace, A., & Serratrice, L. (2009). Internal and external interfaces in bilingual language development: Beyond structural overlap. *International Journal of Bilingualism*, *13*(2), 195-210.

Sorace, A., Serratrice, L., Filiaci, F., & Baldo, M. (2009). Discourse conditions on subject pronoun realization: Testing the linguistic intuitions of older bilingual children. *Lingua*, *119*(3), 460-477.

Spada, N. & Tomita, Y. (2010). Interactions Between Type of Instruction and Type of Language Feature: A Meta-Analysis. *Language Learning 60*(2): 263–308.

Staehr, L. S. (2008). Vocabulary size and the skills of listening, reading and writing. *Language Learning Journal*, *36*(2), 139-152.

Stavrakaki, S., & Tsimpli, I. M. (2000). Diagnostic verbal IQ test for Greek preschool and school age children: Standardization, statistical analysis, psychometric properties. In *Proceedings of the 8th Symposium of the Panhellenic Association of Logopedists* (pp. 95-106). Athens: Ellinika Grammata.

Strand, S. & Hessel, A. (2018). *English as an Additional Language, proficiency in English and pupils' educational achievement: An analysis of Local Authority data*. Bell Foundation.

Swain, M. & Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied Linguistics, 16*, 371-391.

Theakston, A. L., Lieven, E. V. M., Pine, J. M., & Rowland, C. F. (2001). The role of performance limitations in the acquisition of verb-argument structure: an alternative account. *Journal of Child Language, 28*, 127-152.

Thordardottir, E. (2011). The relationship between bilingual exposure and vocabulary development. *International Journal of Bilingualism, 15*(4), 426-445.

Tomasello, M. (2003). *Constructing a Language: a Usage-Based Theory of Language Acquisition*. Cambridge MA: Harvard University Press.

Tsimpli, l.-M. (1990). The clause structure and word order of Modern Greek. *UCL Working Papers in Linguistics, 2*, 220-258.

Tsimpli, I.M. (2005). Peripheral Positions in early Greek. In M. Stavrou and A. Terzi (eds.) *Advances in Greek Generative Syntax,* Linguistics Today (pp. 179-216). Amsterdam: John Benjamins.

Tsimpli, I. M. (2014). Early, late or very late?: Timing acquisition and bilingualism. *Linguistic Approaches to Bilingualism*, *4*(3), 283-313.

Tsimpli, I.M. & A. Roussou. (1991). Parameter-Resetting in L2? *UCL Working Papers in Linguistics, 3*(1), 149-169.

Tsimpli, I.M., & Sorace, A. (2006). Differentiating interfaces: L2 performance in syntax-semantics and syntax-discourse phenomena. In D. Bamman, T. Magnitskaia, and C. Zaller (eds.), *Proceedings of the 30th annual Boston University Conference on Language Development* (pp. 653-664). Somerville, MA: Cascadilla Press.

Unsworth, S. (2005). Child L2, Adult L2, Child L1: Differences and Similarities. A study on the acquisition of direct object scrambling in Dutch. *Language Acquisition, 14(2)*, 215-217.

Unsworth, S. (2013a). Assessing the role of current and cumulative exposure in simultaneous bilingual acquisition: The case of Dutch gender. *Bilingualism: Language and Cognition, 16*(1), 86-110.

Unsworth, S. (2013b). Assessing age of onset effects in (early) child L2 acquisition. *Language Acquisition*, *20*, 74–92.

Unsworth. S. (2015). "Amount of exposure as a proxy for dominance in bilingual language acquisition". In C. Silva-Corvalan & J. Treffers-Daller (Eds.), *Language dominance in bilinguals: Issues of measurement and operationalization.* pp. 156-173. Cambridge: Cambridge University Press.

Unsworth, S., Argyri, F., Cornips, L., Hulk, A., Sorace, A., & Tsimpli, I. (2014). The role of age of onset and input in early child bilingualism in Greek and Dutch. *Applied Psycholinguistics, 35*(4), 765-805.

Valian, V. (1991). Syntactic subjects in the early speech of American and Italian children. *Cognition*, *40*(1-2), 21-81.

Valian, V., & Eisenberg, Z. (1996). The development of syntactic subjects in Portuguese-speaking children. *Journal of Child Language*, *23*(1), 103-128.

Valian, V., Hoeffner, J., & Aubry, S. (1996). Young children's imitation of sentence subjects: Evidence of processing limitations. *Developmental Psychology*, *32*(1), 153.

Vihman, M. M. (1985). Language differentiation by the bilingual infant. *Journal of Child Language, 12*(2), 297-324.

Volterra, V., & Taeschner, T. (1978). The acquisition and development of language by bilingual children. *Journal of Child Language*, *5*(2), 311-326.

Wagner, S. (2018). Never saw one–first-person null subjects in spoken English. *English Language & Linguistics*, *22*(1), 1-34.

Wei, L. (2018). Translanguaging as a practical theory of language. *Applied Linguistics*, *39*(1), 9-30.

Wei, L., & Ho, W. Y. J. (2018). Language learning sans frontiers: A translanguaging view. *Annual Review of Applied Linguistics*, *38*, 33-59.

Werker, J. F., & Byers-Heinlein, K. (2008). Bilingualism in infancy: First steps in perception and comprehension. *Trends in Cognitive Sciences, 12*(4), 144-151.

White, L. (1985). The "pro-drop" parameter in adult second language acquisition. *Language Learning, 35*(1), 47-62.

White, L. (1986). “Markedness and parameter setting: some implications for a theory of adult second language acquisition”. In F. Eckman, E. Moravcsik, & J. Wirth (Eds). *Markedness*. pp. 309-327. New York: Plenum.

Wonnacott, E., Boyd, J.K, Thomson. J.J., & Goldberg, A.E. (2012) Input effects on the acquisition of a novel phrasal construction in 5 year olds. *Journal of Memory and Language, 66*, 458-478.

Wonnacott, E., Brown, H., & Nation, K. (2017) Skewing the evidence: The effect of input structure on child and adult learning of lexically-based patterns in an artificial language. *Journal of Memory and Language, 95*, 46-48.

Wulff, S. & Ellis, N.C. (2018). “Usage-based approaches to second language acquisition”. In D. Miller, F. Bayram, J. Rothman & L. Serratrice (Eds). *Bilingual Cognition and Language: The state of the science across its subfields.* pp. 37-56. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Yip, V., & Matthews, S. (2000). Syntactic Transfer in a Cantonese-English Bilingual Child. *Bilingualism: Language and Cognition, 3*(3), 193-208.

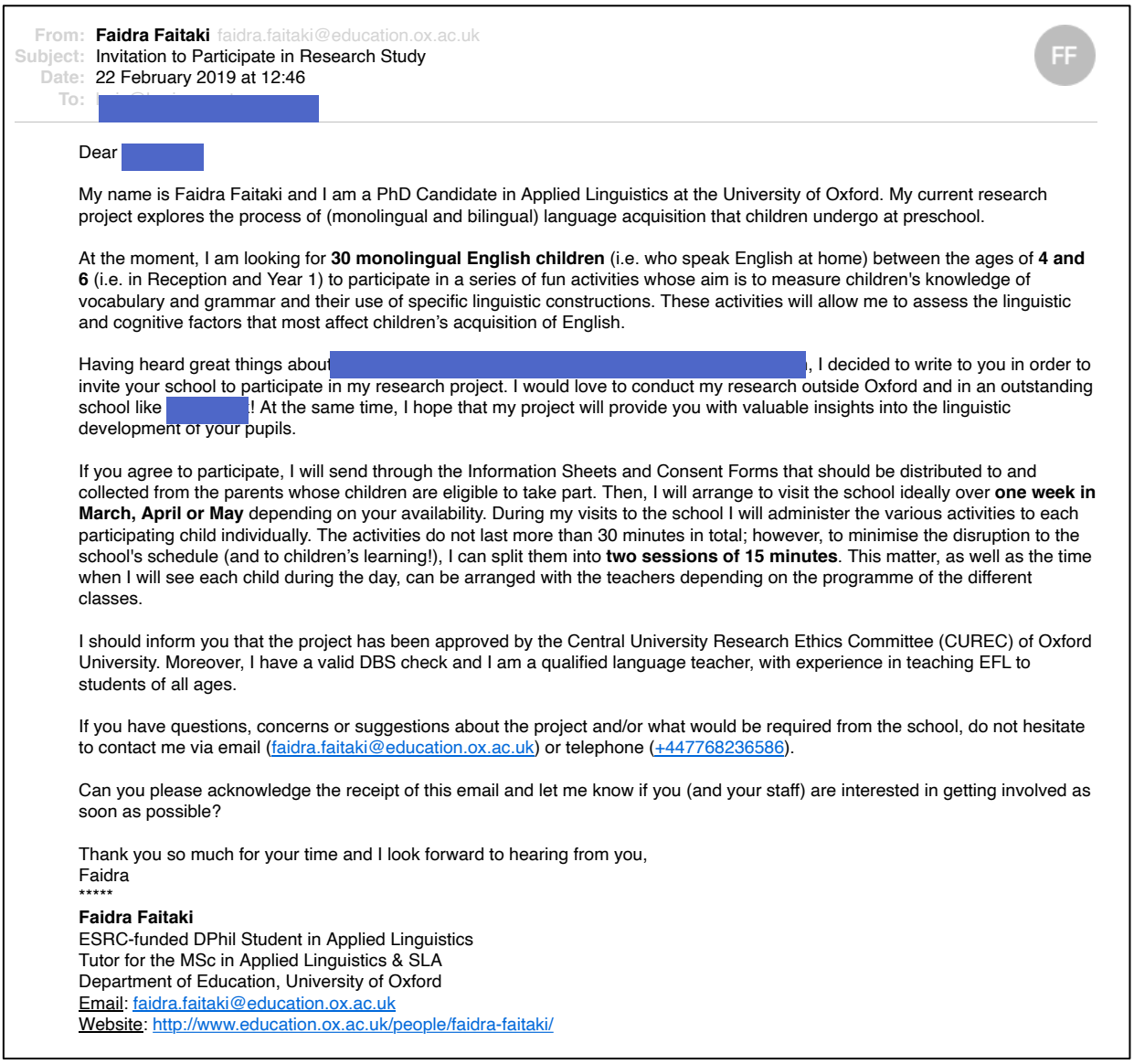
Zdorenko, T., & Paradis, J. (2008). The acquisition of articles in child second language English: fluctuation, transfer or both?. *Second language research*, *24*(2), 227-250.

Zdorenko, T., & Paradis, J. (2011). Articles in child L2 English: When L1 and L2 acquisition meet at the interface. *First Language*, *32*(1-2), 38-62.

Appendices

Appendix A: Email Invitation to Eligible Schools

The email invitation is presented in its English format (sent out to schools in the UK to recruit English monolingual children). Greek schools were invited using the same email invitation, translated into Greek and altered to match the population's characteristics.



From: **Faidra Faitaki** faidra.faitaki@education.ox.ac.uk
Subject: Invitation to Participate in Research Study
Date: 22 February 2019 at 12:46
To: [redacted]

FF

Dear [redacted]

My name is Faidra Faitaki and I am a PhD Candidate in Applied Linguistics at the University of Oxford. My current research project explores the process of (monolingual and bilingual) language acquisition that children undergo at preschool.

At the moment, I am looking for **30 monolingual English children** (i.e. who speak English at home) between the ages of **4 and 6** (i.e. in Reception and Year 1) to participate in a series of fun activities whose aim is to measure children's knowledge of vocabulary and grammar and their use of specific linguistic constructions. These activities will allow me to assess the linguistic and cognitive factors that most affect children's acquisition of English.

Having heard great things about [redacted], I decided to write to you in order to invite your school to participate in my research project. I would love to conduct my research outside Oxford and in an outstanding school like [redacted]! At the same time, I hope that my project will provide you with valuable insights into the linguistic development of your pupils.

If you agree to participate, I will send through the Information Sheets and Consent Forms that should be distributed to and collected from the parents whose children are eligible to take part. Then, I will arrange to visit the school ideally over **one week in March, April or May** depending on your availability. During my visits to the school I will administer the various activities to each participating child individually. The activities do not last more than 30 minutes in total; however, to minimise the disruption to the school's schedule (and to children's learning!), I can split them into **two sessions of 15 minutes**. This matter, as well as the time when I will see each child during the day, can be arranged with the teachers depending on the programme of the different classes.

I should inform you that the project has been approved by the Central University Research Ethics Committee (CUREC) of Oxford University. Moreover, I have a valid DBS check and I am a qualified language teacher, with experience in teaching EFL to students of all ages.

If you have questions, concerns or suggestions about the project and/or what would be required from the school, do not hesitate to contact me via email (faidra.faitaki@education.ox.ac.uk) or telephone (+447768236586).

Can you please acknowledge the receipt of this email and let me know if you (and your staff) are interested in getting involved as soon as possible?

Thank you so much for your time and I look forward to hearing from you,
Faidra
*****
Faidra Faitaki
ESRC-funded DPhil Student in Applied Linguistics
Tutor for the MSc in Applied Linguistics & SLA
Department of Education, University of Oxford
Email: faidra.faitaki@education.ox.ac.uk
Website: http://www.education.ox.ac.uk/people/faidra-faitaki/

Appendix B: Experiment 1 Stimuli

Sentence Type	Sentence	Word Count	Elicitation Target	Target Verb
Practice | Experimental	Every day after lunch, Nick sleeps for an hour. Today…	10	He sleeps on the sofa	Sleep
Practice | Experimental	Anne is driving something at the fun park every weekend. Today…	11	She is driving a car	Drive
Practice | Control	Anne's grandparens sit somewhere to rest in the evenings. Today…	10	Grandma sits on a chair	Sit
Experimental [HE]	Nick drinks something every morning, before going to school. Today…	10	He drinks some juice	Drink
Experimental [HE]	Nick buys treats from the shop on his walk home. Today…	11	He buys a chocolate	Buy
Experimental [HE]	Nick goes for a walk in the neighborhood every evening. Today…	11	He goes to the park	Go
Experimental [HE]	Nick sees the animals at the zoo on Saturdays. Today…	10	He sees the zebras	See
Experimental [HE]	Nick tells his friend different kinds of story at school. Today...	11	He tells a scary story	Tell
Control	Nick and his teacher get each other a gift sometimes. Today…	11	The teacher gets Nick a book	Get
Control	One of Nick's parents gives him a present sometimes. Today…	10	Nick's Mum gives Nick a toy	Give
Control	Nick's siblings open things for him because he is little. Today…	11	Nick's brother opens a bottle	Open
Control	Nick's parents run somewhere near the house in the weekend. Today…	11	Nick's Mum runs at the beach	Run
Control	Nick's parents stand somewhere waiting for him in the morning. Today…	11	Nick's Mum stands at the door	Stand
Experimental [SHE]	Anne looks at something in the house every evening. Today…	10	She looks at a painting	Look
Experimental [SHE]	Anne has a new toy to play with every day. Today…	11	She has a ball	Have

Experimental [SHE]	Anne eats one of her favourite fruits in the evenings. Today...	11	She is eating an apple	Eat
Experimental [SHE]	Anne works in a shop from time to time. Today…	10	She works in the bakery	Work
Experimental [SHE]	Anne reads something every night before going to bed. Today…	10	She reads a book	Read
Control	One of Anne's parents makes her something special sometimes. Today…	10	Anne's Dad makes a card	Make
Control	Anne's parents find her toys in their bag sometimes. Today...	10	Anne's Dad finds her doll	Find
Control	Anne's parents play various musical instruments in the weekend. Today…	10	Anne's Dad plays the drums	Play
Control	Anne's grandparents show her something special when she visits. Today…	10	Granma shows Anne her books	Show
Control	Anne and her friend climb something in the weekends. Today...	10	Anne's friend climbs the trees	Climb

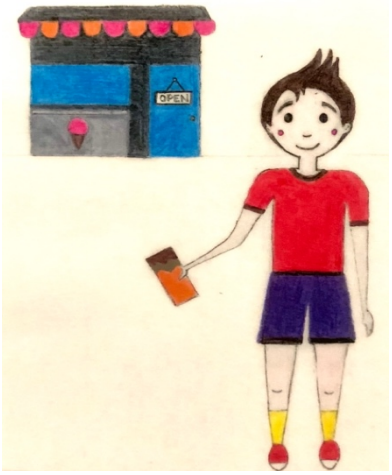
Appendix C: Experiment 1 Pictures


Sleep


Drive


Sit


Drink



Buy


Go


See


Tell

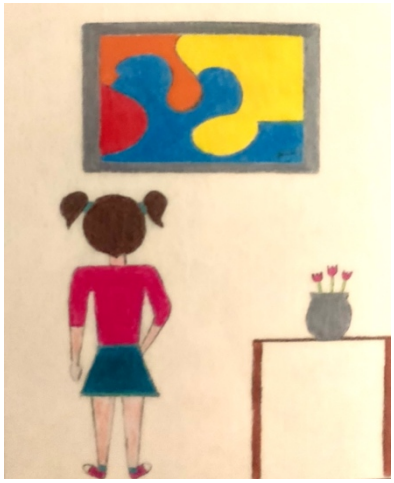
Get

Give

Open

Run

Stand



Look

Have

Eat

Work

Read

Make

Find

Play

Show

Climb

Appendix D: CUREC Approval for the DPhil Project

From: Liam Gearon <liam.gearon@education.ox.ac.uk>
Subject: CUREC
Date: 6 March 2018 at 19:10:31 GMT
To: Faidra Faitaki <faidra.faitaki@education.ox.ac.uk>
Cc: Victoria Murphy <victoria.murphy@education.ox.ac.uk>, Education Research Office <research.office@education.ox.ac.uk>, Kate Nation <kate.nation@psy.ox.ac.uk>

Dear Faidra,

Under Crosslinguistic Influence: Investigating the Linguistic Development of Greek-English Successive Bilingual Preschool Children

The above application has been considered on behalf of the Departmental Research Ethics Committee (DREC) in accordance with the procedures laid down by the University for ethical approval of all research involving human participants.

I am pleased to inform you that, on the basis of the information provided to DREC, the proposed research has been judged as meeting appropriate ethical standards, and accordingly, approval has been granted.

If your research involves participants whose ability to give free and informed consent is in question (this includes those under 18 and vulnerable adults), then it is advisable to read the following NSPCC professional reporting requirements for cases of suspected abuse

http://www.nspcc.org.uk/globalassets/documents/information-service/factsheet-child-abuse-reporting-requirements-professionals.pdf

Should there be any subsequent changes to the project which raise ethical issues not covered in the original application you should submit details to research.office@education.ox.ac.uk for consideration.

Good luck with your research study.

Sincerely,

With kind regards,

Liam

Dr LF Gearon
Chair DREC, Department of Education
University of Oxford

Appendix E: Participant Information Sheet (Experiments 1 & 2)

UNIVERSITY OF OXFORD
DEPARTMENT OF EDUCATION

15 Norham Gardens, Oxford, OX2 6PY
Tel: (+44)(0)1865274024 Fax: (+44)(0)1865274027
general.enquiries@education.ox.ac.uk
www.education.ox.ac.uk

Faidra Faitaki, DPhil Student
St. Edmund Hall, Queen's Lane, Oxford, OX1 4AR
(+44)(0)7768236586 (UK), (+30)6947760923 (Greece)
faidra.faitaki@education.ox.ac.uk
www.education.ox.ac.uk/people/faidra-faitaki

Under Crosslinguistic Influence:
Investigating the Linguistic Development of Successive Bilingual Preschool Children

INFORMATION FOR PARENTS/GUARDIANS

Your children's school has agreed to take part in a University of Oxford study, which investigates how bilingual and monolingual children develop their English. We would like to invite your child to be part of this study! Before you let your child take part, it is important that you understand what will the study involve.

What are we trying to find out?

In this study, we investigate how children develop their English at preschool. A substantial body of research reports that children's language skills at the start of school forecast their educational success in the future. With this information in mind, the study will delve into language learning at preschool. It will explore various factors that might affect children's language development at this stage, such as their knowledge of English vocabulary and grammar, their cognitive skills and their use of specific linguistic structures. We hope that this exploration will help support children -and those around them- in their linguistic journey(s)!

Why has my child been invited to take part?

We invited your child to take part because he/she is aged between 4 and 6 years old, and speaks English at home.

Does my child have to take part?

No. You can ask questions about the study before deciding whether or not to allow your child to participate. In fact, even if you agree to participation, you may withdraw your child from the study at any time, without giving a reason and without penalty, by advising us of this decision.

Who is conducting this research?

The research project is organised by Faidra Faitaki, a DPhil student at the University of Oxford, supervised by Professor Victoria Murphy and Professor Kate Nation. Faidra is also a qualified language teacher, with experience in teaching English and Greek to pupils of all age groups. Her research is funded by the Economic and Social Research Council (ESRC), and her study has been reviewed and received ethics clearance by the University of Oxford's Central University Research Ethics Committee.

Why should I allow my child to take part?

Do you want to find out about your child's linguistic development, and how to best support them? Do you simply want your child to have a unique, educational and recreational experience? By allowing your child to participate, you can meet these goals!

What will happen if my child takes part?

The main researcher, Faidra, will either arrange a visit to your children's school. During her visit, Faidra will play various games with your child, lasting around 30 minutes in total. These games explore how children use specific linguistic structures but also, they are great fun! The visits will take place at a time and place that the teacher considers suitable and not disruptive to your child's schedule.

What will happen to any samples taken from my child?

We will ensure that the samples remain safe throughout and after the study. All personal information you provide will remain confidential. Children's performance in the tasks will be audio-recorded, and the recordings will be transcribed and analysed by Faidra. The recordings, transcriptions and analyses will be anonymised (each child's samples will be identified by a number, rather than by the child's name). All data will be stored in password-protected laptop folder and deposited in a secure database (maintained by the University of Oxford), where it will be retained after the study's completion in accord with the guidelines of the University of Oxford and the ESRC. It should be noted that data might be used for different projects, publications and presentations after the completion of the DPhil.

What happens to the results of the study?

A of the study's findings will be given to the school and become available to all interested families, so you can ask your child's school for a copy. Overall, the research will be written up as a thesis. On successful submission of the thesis, it will be deposited in print and online in the University of Oxford archives. The thesis will be published with open-access and will be available to all internet users. We also aim to publish our findings in scientific journals, but this may be after end of the study.

What if I have questions or concerns?

If you have a concern about any aspect of this project, please speak to Faidra Faitaki, who will do her best to answer your query. She will acknowledge your concern within 10 working days and suggest how she intends to deal with it. If you wish to make a formal complaint, please contact the chair of Sciences & Humanities Inter-Divisional Research Ethics Committee (ethics@socsci.ox.ac.uk; Research Services, University of Oxford, Wellington Square, Oxford OX1 2JD) who will seek to resolve the matter in a reasonably expeditious manner.

What should I do next?

Please fill in the enclosed form and return it to your child's class teacher if you would like your child to take part in this study. If you would like to discuss the research with someone beforehand (or if you have questions afterwards), please contact Faidra and/or her main supervisor, Professor Victoria Murphy:

Faidra Faitaki, DPhil Student
Address: St Edmund Hall, Queen's Ln, Oxford, OX1 4AR
Mobile: +44(0)7768236586
Email Address: faidra.faitaki@education.ox.ac.uk

Dr Victoria Murphy, Professor in Applied Linguistics
Address: Department of Education, 15 Norham Gardens, Oxford, OX2 6PY
Telephone: +44(0)1865274042
Email Address: victoria.murphy@education.ox.ac.uk

Note that the above form was used to recruit the English monolingual participants. The forms used to recruit the Greek monolingual and the Greek-English bilingual participants were similar but were translated into Greek and contained a modified version of the text in the second

question (i.e. “Why has my child been invited to take part”), in order to reflect the children’s Greek monolingual or Greek-English bilingual status.

Appendix F: Consent Form (Experiments 1 & 2)

UNIVERSITY OF OXFORD
DEPARTMENT OF EDUCATION

15 Norham Gardens, Oxford, OX2 6PY
Tel: (+44)(0)1865274024 Fax: (+44)(0)1865274027
general.enquiries@education.ox.ac.uk
www.education.ox.ac.uk

Faidra Faitaki, DPhil Student
St. Edmund Hall, Queen's Lane, Oxford, OX1 4AR
(+44)(0)7768236586 (UK), (+30)6947760923 (Greece)
faidra.faitaki@education.ox.ac.uk
www.education.ox.ac.uk/people/faidra-faitaki

Investigating the Linguistic Development of Successive Bilingual Preschool Children

CONSENT FORM FOR PARENTS/GUARDIANS

Your child's school has agreed to take part in a study run by Faidra Faitaki, DPhil Student at the University of Oxford, looking at how children develop their English at preschool.

If your child takes part, Faidra will play various fun games with them in English during school hours. Your child will be audio-recorded for the duration of the games, and the recordings will then be transcribed and analysed anonymously.

If you are happy for your child to take part, please fill in the form below and return it to your child's teacher as soon as possible.

To find out more about the study, please read the attached information sheet. If you have questions/concerns, you can contact Faidra by email at faidra.faitaki@education.ox.ac.uk, or on (+44)(0)7768236586.

Child's Name: _____

Child's Date of Birth: _____

I have read and understood the details of the study. I understand that the project has received ethics clearance through the University of Oxford's ethical approval process for research involving human participants as well as who will have access to the data, how it will be stored and what will happen to it at the end of the study. I understand that participation is voluntary and that my child and I are free to withdraw at any time, without giving any reason and without my child's education being affected in any way. I understand how to raise a concern or make a complaint. I have had the opportunity to ask questions and I have received satisfactory answers to my questions (if any).

I give permission for my child to take part in the above study.

Parent/Guardian Name: _____

Signature:_____**Date:** _____

Appendix G: Pilot Results

The first of the pilot's questions asked whether the task used affects children's use of the two target structures. Children's mean score for subject pronouns was 6.41 (*SE*=.45, *SD*=2.79) in the Repetition Task (RT) and 5.64 (*SE*=.54, *SD*=3.38) in the Elicitation Task (ET). Children's mean score for infinitival TO was 6.77 (*SE*=.47, *SD*=2.94) in the RT, and 5.08 (*SE*=.59, *SD*=3.71) in the ET. Taken together, these findings suggest that the task does affect children's production, as mean scores were higher in the RT for both linguistic structures.

To look deeper into the differences between the tasks and answer the pilot's second question (if children who attend different educational programmes perform differently in the two tasks), a series of independent samples t-tests (with Bonferroni correction applied) was conducted. In these tests, the educational programme functioned as the grouping variable. The t-tests revealed that children attending submersion and immersion performed alike in the RT. Specifically, for subject pronouns the submersion group's mean score was 6.40 (*SD*=3.40) and the immersion group's was 6.42 (*SD*=2.41). The difference, -.02, was not significant: $t(37)=-.18$, $p=.99$, $d=.003$. For infinitival TO, the former group's score was 6.53 (*SD*=2.99) and the latter group's was 6.92 (*SD*=2.96). Once again, the difference between the groups, -.38, was not significant: $t(37)=-.39$, $p=.69$, $d=.15$.

The two groups did not perform similarly in the ET. For subject pronouns, the submersion group scored 7.40 (*SD*=3.31) and the immersion group 4.54 (*SD*=2.99). This difference, 2.86, was marginally significant, $t(37)=2.79$, $p=.008$, and represented a large effect $d=.91$. Likewise, the submersion group scored 7.80 (*SD*=2.31) for infinitival TO, and the immersion group scored 3.38 (*SD*=3.42) for the same structure. The difference, 4.42, was highly significant $t(37)=4.82$, $p=.000$ and represented a very large effect $d=1.51$.

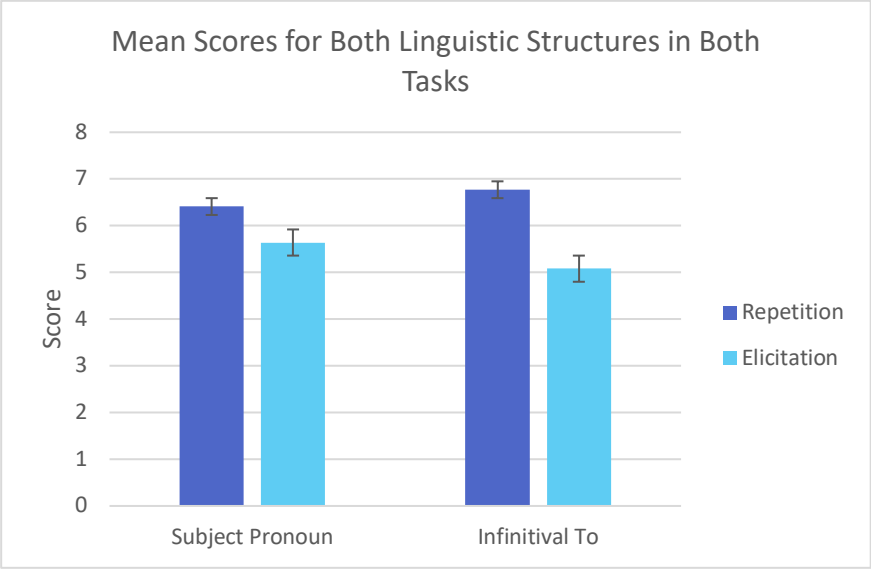


Figure 1: Children's mean scores for each task per linguistic structure, with standard error

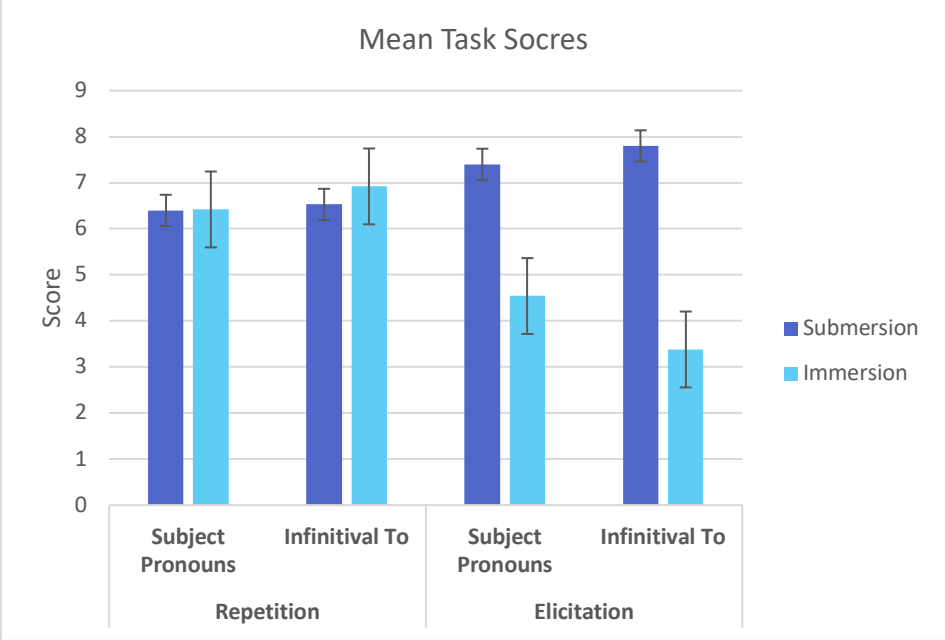


Figure 2: Mean scores with standard error for both linguistic structures and in both tasks per education group

Turning to the pilot's third question, if children's production of the two structures in the two tasks is linked to their knowledge of vocabulary and grammar, Table 1 presents the mean scores for the PPVT and TROG. Overall, children in the submersion programme performed better in both language tests.

Table 1: Mean scores, standard error and standard deviation for the PPVT and TROG

	Total	Submersion	Immersion
PPVT	49.08 (*SE*=2.34, *SD*=16.4)	52.85 (*SE*=4.46, *SD*=19.95)	46.48 (*SE*=2.45, *SD*=13.19)
TROG	20.25 (*SE*=1.15, *SD*=7.94)	24.50 (*SE*=1.94, *SD*=8.24)	17.70 (*SE*=1.27, *SD*=6.66)

Two independent-samples t-tests (one for each standardised test) were conducted to compare the performance of the educational groups in the PPVT and the TROG. According to the t-tests, and as Figure 3 illustrates, the difference between the groups on the PPVT, 6.37 was not significant $t(47)=1.35$, $p=.18$ and represented a small effect size $d=.38$. Rather, the groups' mean difference on the TROG, 6.8, was significant $t(46)=3.13$, $p=.003$, and represented a large effect size $d=.91$ as shown in Figure 4.

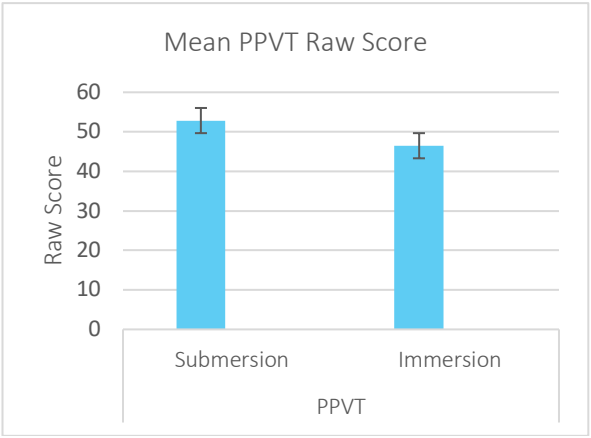


Fig. 3: Mean PPVT raw scores for both groups

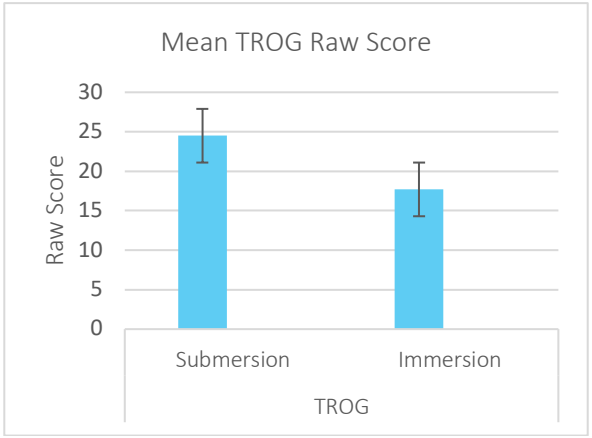


Fig. 4: Mean TROG raw scores for both groups

To explore these differences further, a bivariate Spearman's correlation was performed. There was a positive, significant, medium correlation between PPVT and the TROG raw scores, suggesting that children who score high on the PPVT also score high on the TROG. Moreover, the PPVT raw score was significantly correlated with the subject pronoun score in the RT, while the TROG raw score was significantly correlated with both subject pronoun and infinitival TO scores in the ET.

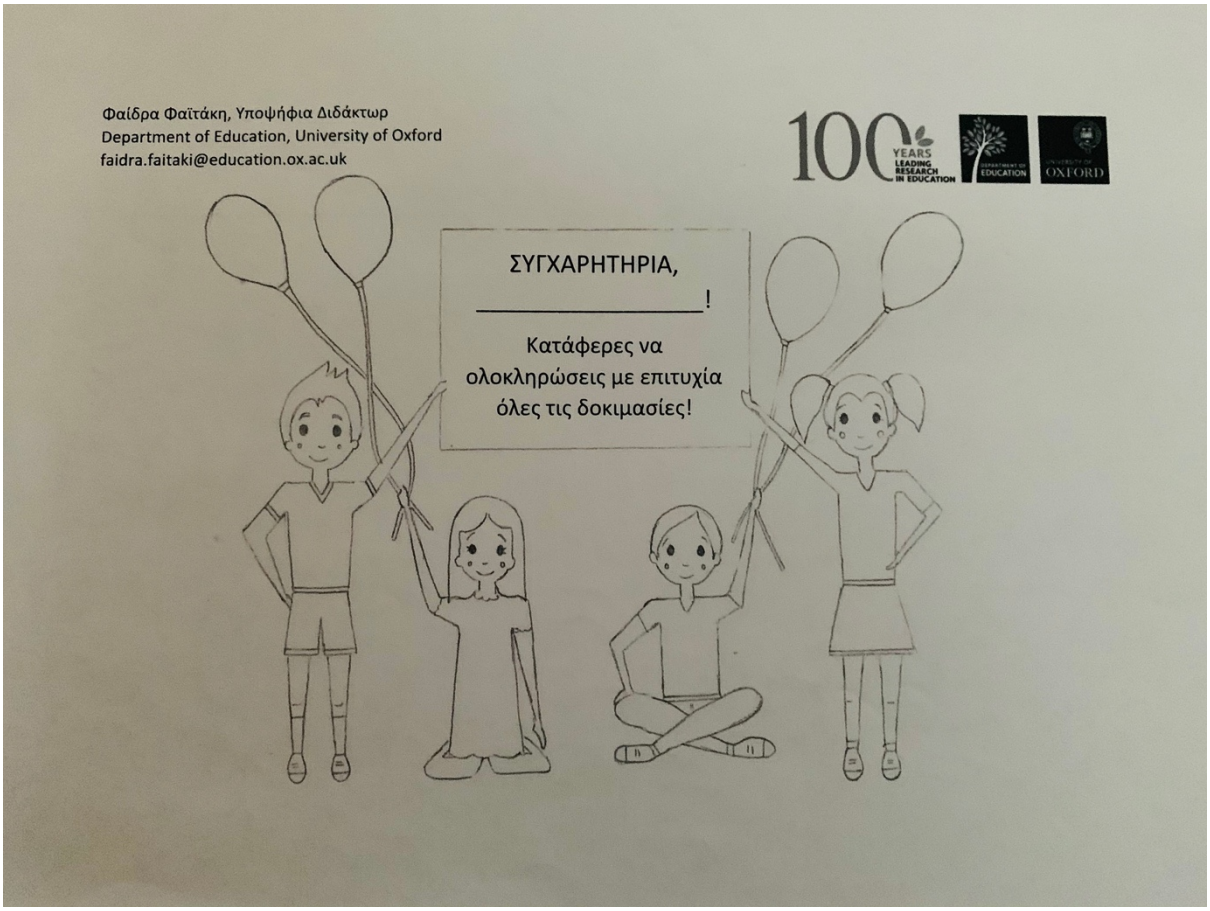
Table 2: Correlations for the PPVT scores, TROG scores, and both linguistic structures per task
Note that *: p<.05, **: p<.01, ***: p<.001

				RT		ET	
		PPVT	TROG	Pronouns	Infinitival To	Pronouns	Infinitival To
	PPVT	-					
	TROG	.529***	-				
RT	Pronouns	.365	.173	-			
	Infinitival To	-.076	.306	.211	-		
ET	Pronouns	.295	.370*	.463**	.074	-	
	Infinitival To	.304	.466**	.279	.025	.424**	-

The reported positive, medium correlations between the children's raw scores for TROG and the two linguistic structures in the ET suggest that children with high scores on the test perform better in the task. For the RT, the correlation coefficients between the scores for this task and for the PPVT and TROG are descriptively lower (or negative) and non-significant. Thus, the role of children's lexical and grammatical knowledge on their performance in the RT cannot be fully determined.

Appendix H: Certificate of Participation

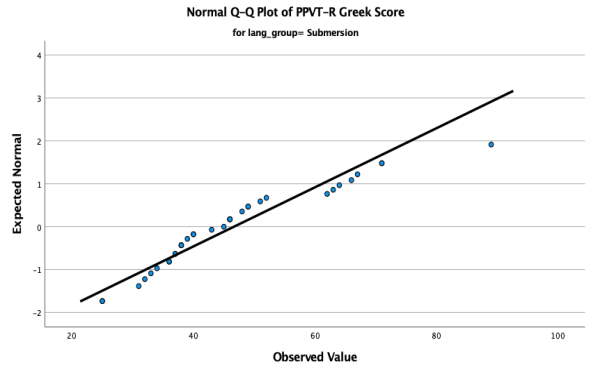
Note that the certificate below was used in the Greek treatment. The certificate that children received after the English treatment was the same, but translated into English. The text in the box reads: “Congratulations! You successfully completed all the activities!”.

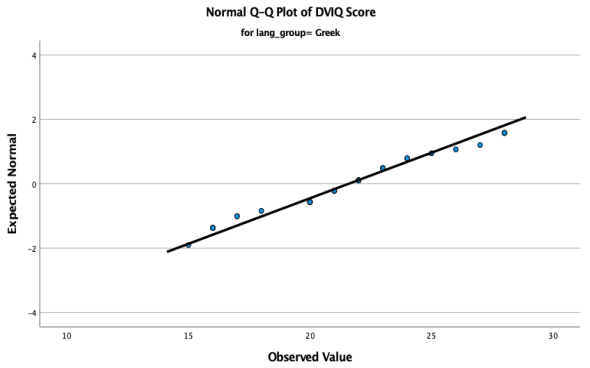


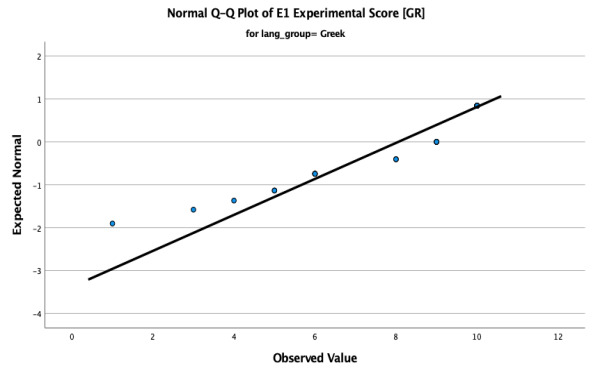
Appendix I: Normality & Homogeneity of Variances for Greek (Experiment 1)

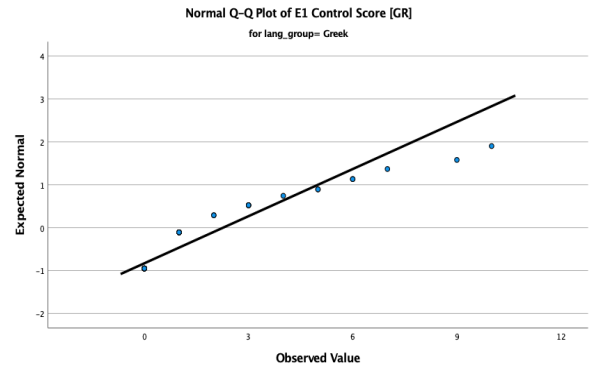
Normal QQ Plots (assessing normality) of: vocabulary test scores (PPVT-R Greek); grammar test scores (DVIQ) & Experiment 1 scores.

Monolingual Greek Children

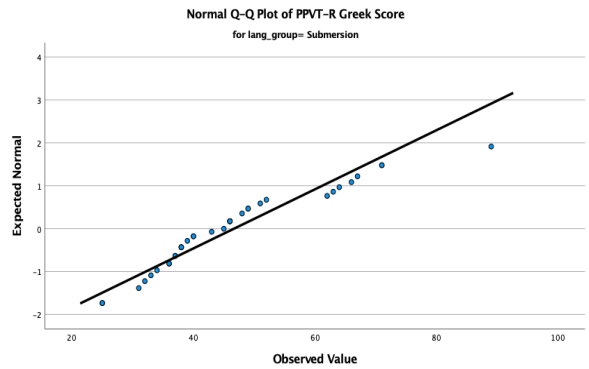


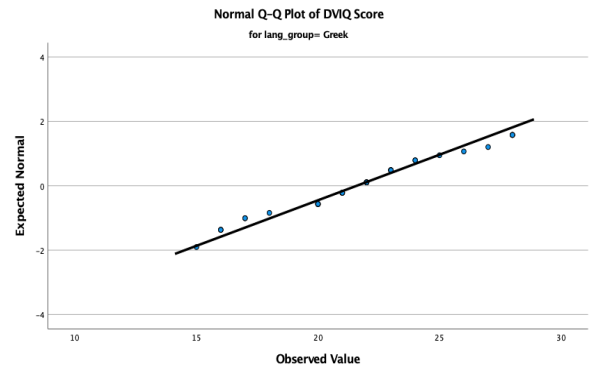


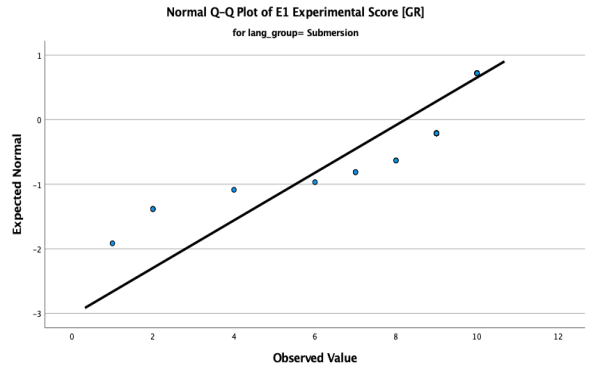


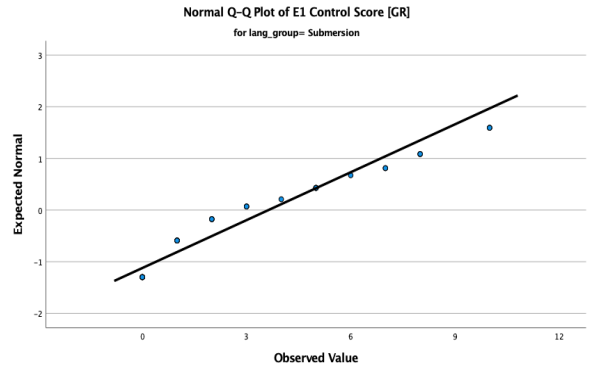


Bilingual Children (Total Immersion)

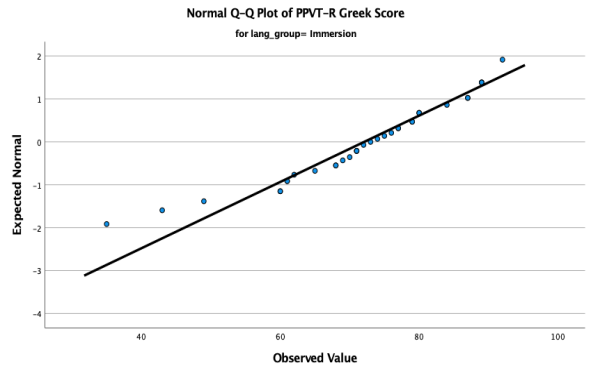


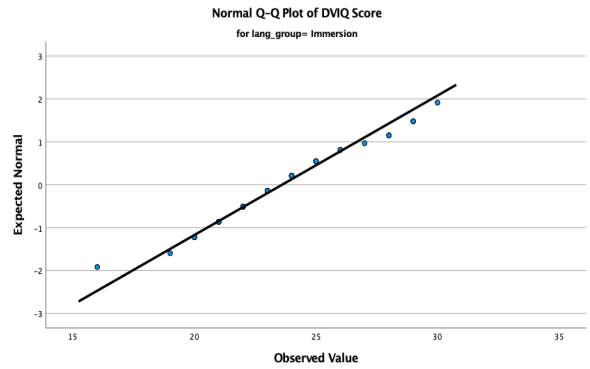


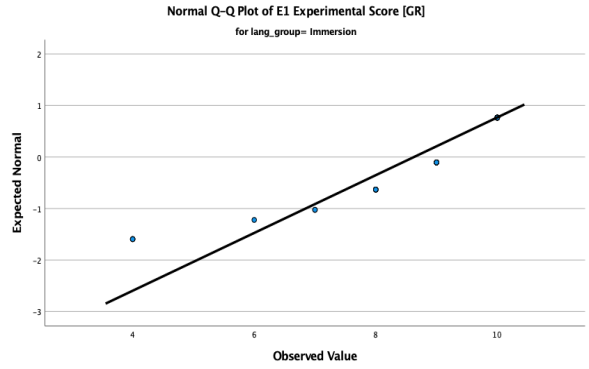


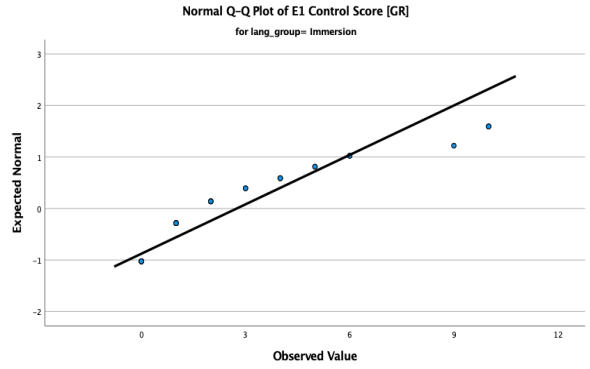


Bilingual Children (Partial Immersion)









Levene's Test (assessing homogeneity of variances) for the three groups': vocabulary test scores (PPVT-R Greek); grammar test scores (DVIQ) & Experiment 1 scores.

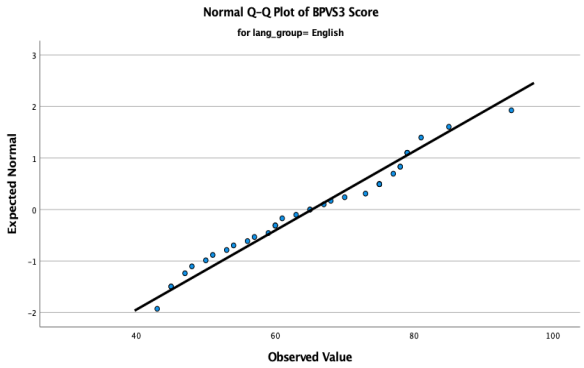
Tests of Homogeneity of Variances

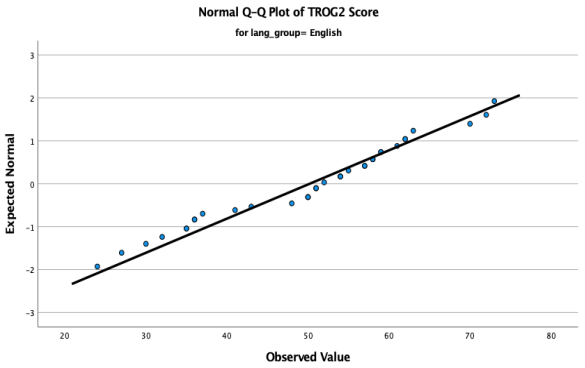
		Levene Statistic	df1	df2	Sig.
PPVT-R Greek Score	Based on Mean	1.401	2	101	.251
	Based on Median	1.386	2	101	.255
	Based on Median and with adjusted df	1.386	2	100.208	.255
	Based on trimmed mean	1.409	2	101	.249
DVIQ Score	Based on Mean	.734	2	101	.482
	Based on Median	.646	2	101	.526
	Based on Median and with adjusted df	.646	2	98.945	.526
	Based on trimmed mean	.741	2	101	.479
E1 Experimental Score [GR]	Based on Mean	2.185	2	101	.118
	Based on Median	.730	2	101	.484
	Based on Median and with adjusted df	.730	2	87.791	.485
	Based on trimmed mean	1.673	2	101	.193
E1 Control Score [GR]	Based on Mean	1.109	2	101	.334
	Based on Median	1.127	2	101	.328
	Based on Median and with adjusted df	1.127	2	97.870	.328
	Based on trimmed mean	1.201	2	101	.305

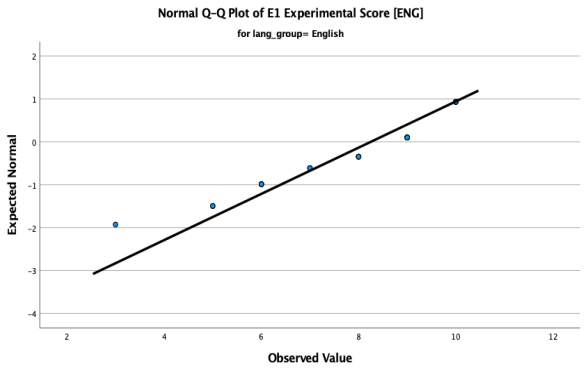
Appendix J: Normality & Homogeneity of Variances for English (Experiment 1)

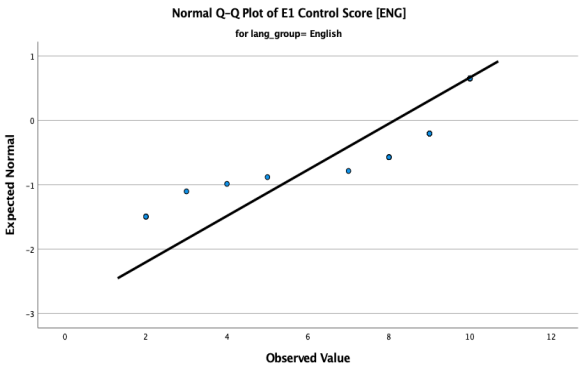
Normal QQ Plots (assessing normality) of: vocabulary test scores (BPVS3); grammar test scores (TROG2) & Experiment 1 scores.

Monolingual English Children

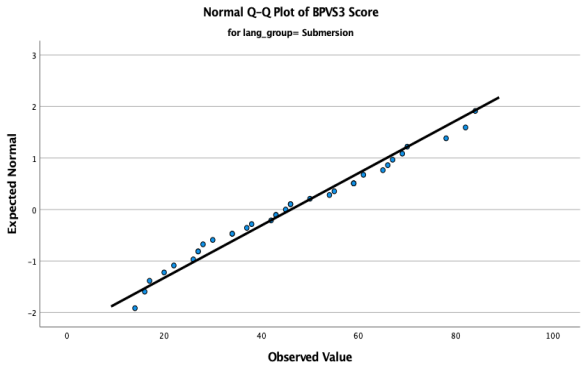


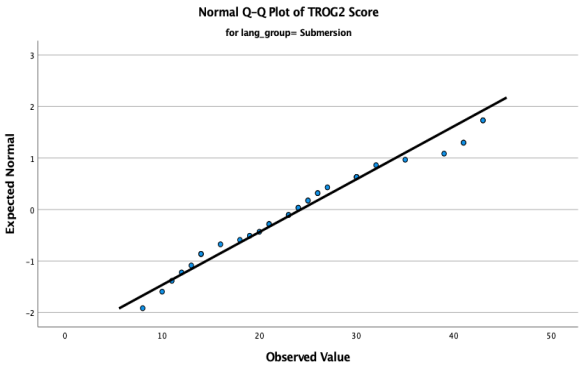


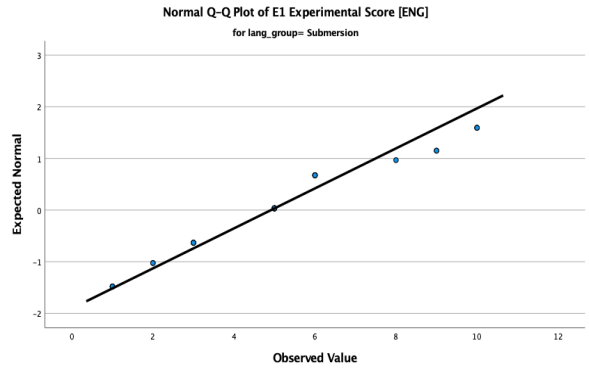


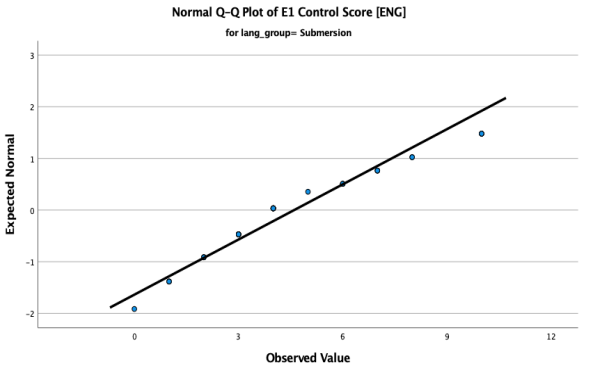


Bilingual Children (Total Immersion)

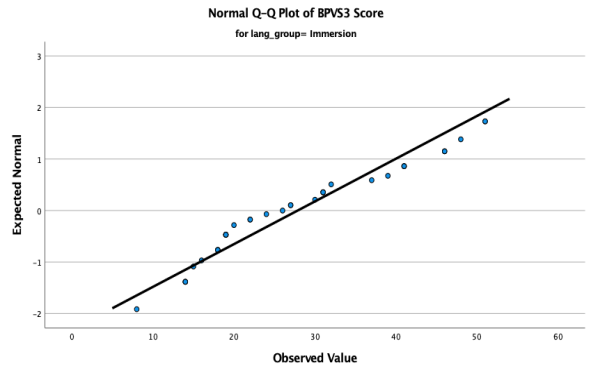


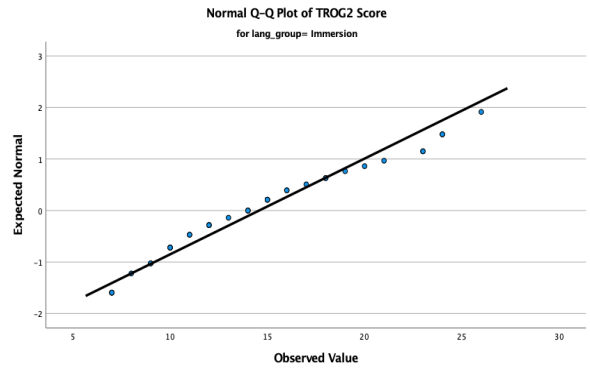


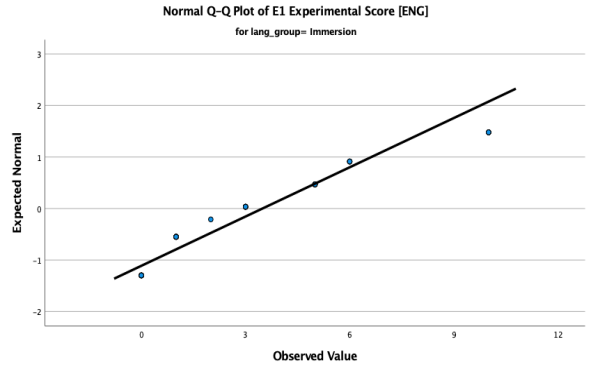


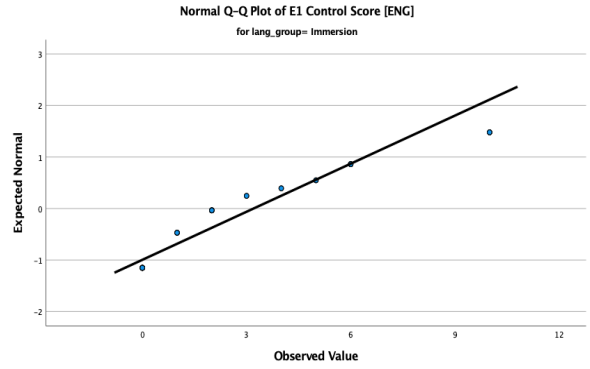


Bilingual Children (Partial Immersion)









Levene's Test (assessing homogeneity of variances) for the three groups': vocabulary test scores (BPVS3); grammar test scores (TROG2) & Experiment 1 scores.

Tests of Homogeneity of Variances

		Levene Statistic	df1	df2	Sig.
BPVS3 Score	Based on Mean	6.180	2	103	.003
	Based on Median	6.002	2	103	.003
	Based on Median and with adjusted df	6.002	2	84.211	.004
	Based on trimmed mean	6.142	2	103	.003
TROG2 Score	Based on Mean	8.286	2	103	.000
	Based on Median	7.624	2	103	.001
	Based on Median and with adjusted df	7.624	2	76.205	.001
	Based on trimmed mean	8.184	2	103	.001
E1 Experimental Score [ENG]	Based on Mean	4.294	2	103	.016
	Based on Median	3.491	2	103	.034
	Based on Median and with adjusted df	3.491	2	96.908	.034
	Based on trimmed mean	4.014	2	103	.021
E1 Control Score [ENG]	Based on Mean	.824	2	103	.442
	Based on Median	.618	2	103	.541
	Based on Median and with adjusted df	.618	2	98.289	.541
	Based on trimmed mean	.807	2	103	.449

Appendix K: Experiment 2 Stimuli

Sentence Type	Sentence	Word Count	Elicitation Target	Target Verb
Practice | Experimental	Mary has a nap after lunch. She wants to sleep on the sofa.	13	She wants to sleep on the bed	Sleep
Practice | Control	John drives something every time he goes at the funpark. He drives a car.	14	He drives an airplane	Drive
Practice | Filler	Mary stands on something for fun sometimes. Today, she stands on the table.	13		Stand
Experimental	Every day, John walks in the evening. He wants to go to the beach.	14	He wants to go to the forest	Go
Experimental	Every day, John gets dressed in the morning. He wants to wear trousers.	13	He wants to wear socks	Wear
Experimental	Every week, John gets a new toy. Today, he wants to have a ball.	14	He wants to have a car	Have
Experimental	Every week, John gives his teacher something. Today, he wants to get her balloons.	14	He wants to get flowers	Get
Experimental	Every day, John buys a treat. Today, he wants to take a sweet.	13	He wants to take popcorn	Take
Experimental	Every week, Mary goes to the zoo. Today, she wants to see the giraffes.	14	She wants to see the gorillas	See
Experimental	Every week, Mary makes something for Mum. Today, she wants to make a drawing.	14	She wants to make a card	Make
Experimental	Every day, Mary has to open something. Today, she wants to open a bottle.	14	She wants to open a box	Open
Experimental	Every week, Mary looks in Mum's wardrobe. Today, she wants to find a skirt.	14	She wants to find a t-shirt	Find
Experimental	Sometimes, Mary stays in the living-room. Today, she wants to sit on the sofa.	14	She wants to sit on the chair	Sit
Control	Mary talks about her favourite animal to her friend. She thinks of dolphins.	13	She thinks of lions	Think

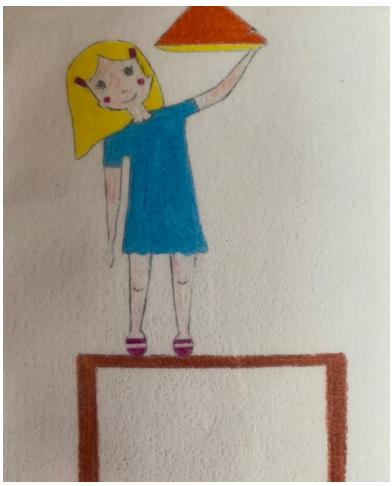
Control	Mary goes to her grandad's house on Saturdays. She comes in through the door.	14	She comes in through the window	Come
Control	Mary goes for a run in the weekends. She runs with her cat.	13	She runs with her dog	Run
Control	Mary must find her pink bag before school. She looks on the bed.	13	She looks on the table	Look
Control	Mary eats one of her favourite fruits every evening. She eats an orange.	13	She eats a banana	Eat
Control	John learns to play a musical instrument at his school. He plays the drums.	14	He plays the piano	Play
Control	John is making pictures for his art class. He is drawing a tree.	13	He draws a flower	Draw
Control	John goes to the museum with his sister. He is showing her a statue.	14	He shows her a painting	Show
Control	John tells his Grandma what to make for dinner. He says spaggetti with meatballs	14	She says burger with chips	Say
Control	John earns some pocket money in the weekends. He works in a restaurant.	13	He works in a book shop	Work
Filler	Mary goes to the beach in the summer. She wants to build a castle.	14		Build
Filler	Mary is at the supermarket for the week's groceries. She wants to buy apples.	14		Buy
Filler	Mary gets something to drink in the morning. Tonight, she wants to drink water.	14		Drink
Filler	Mary goes to the forest on Sunday. He wants to climb a tree.	13		Climb
Filler	John reads something different every evening. Today, he wants to read a book.	13		Read
Filler	John gives Mum something for her birthday every year. He gives her a present.	14		Give
Filler	Mary makes breakfast for her Mum and Dad sometimes. She cooks a cake.	13		Cook

Filler	John tells his friend what his favourite colour is. He likes the colour blue.	14	Like
Filler	John tidies her toys in the evening. He puts them in a box.	13	Put
Filler	John is talking to his friend at school. He tells them a funny story.	14	Tell

Appendix L: Experiment 2 Pictures


Sleep


Drive


Stand

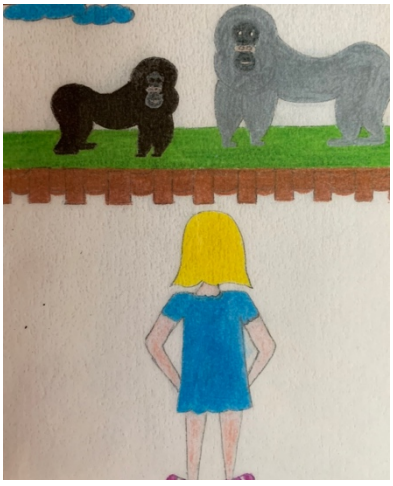

Go

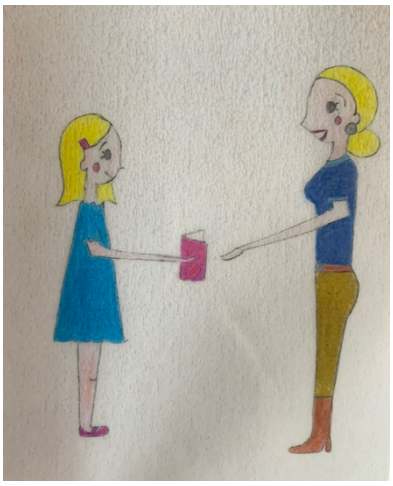

Wear


Have


Get


Take

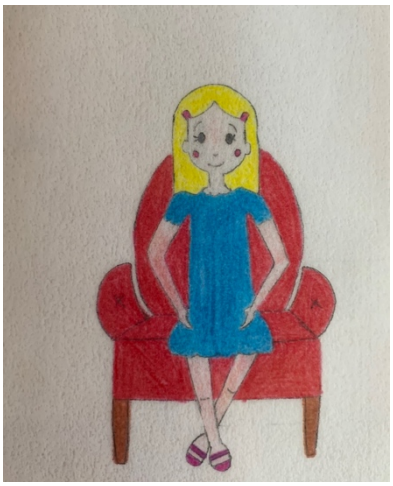

See



Make

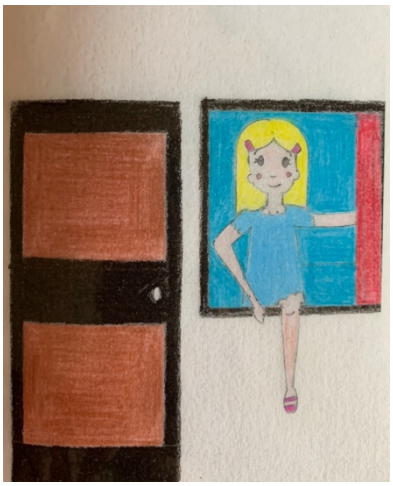
Open

Find



Sit

Think



Come

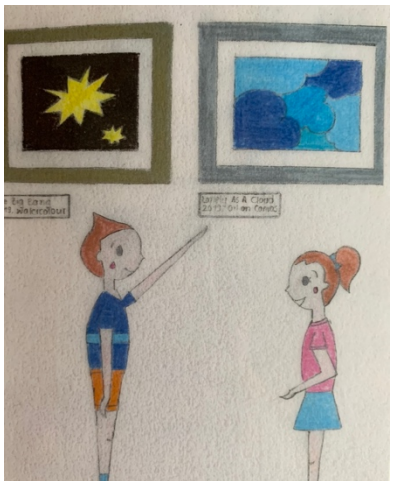
Run

Looks

Eat

Play

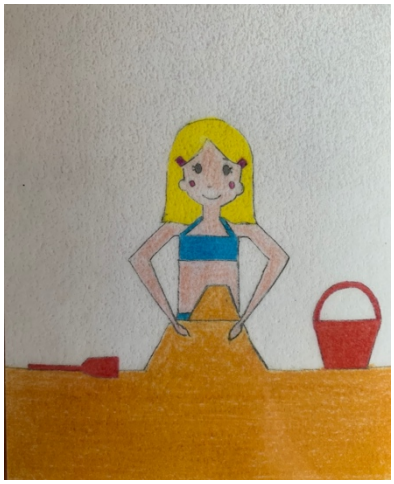
Draw



Show

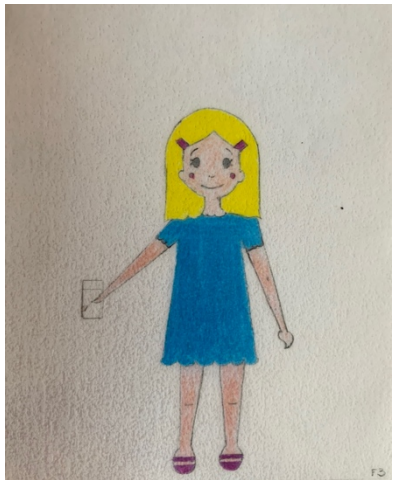
Say

Work



Build

Buy



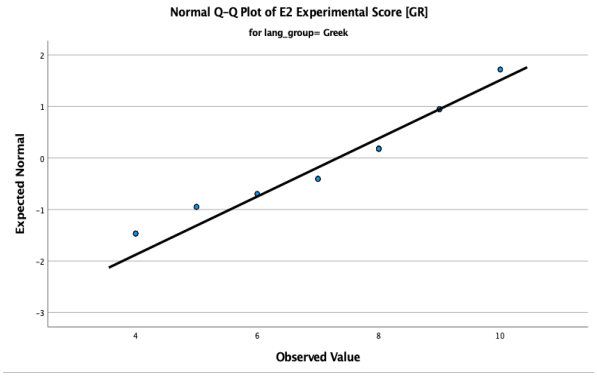
Drink

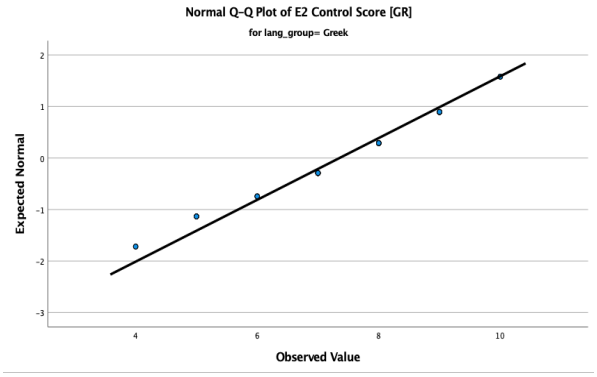
Climb

Appendix M: Normality & Homogeneity of Variances for Greek (Experiment 2)

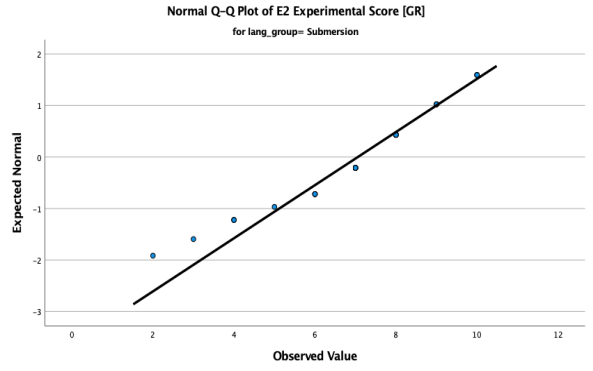
Normal QQ Plots (assessing normality) of Experiment 2 scores.

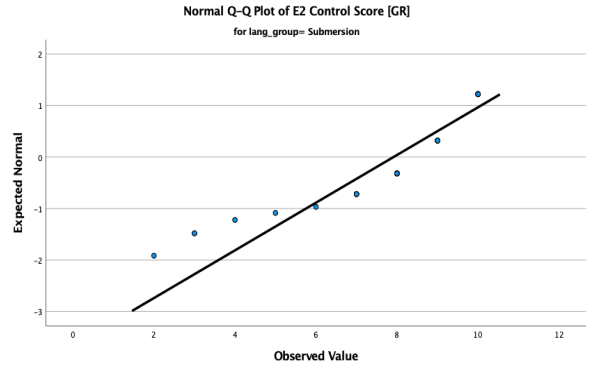
Monolingual Greek Children



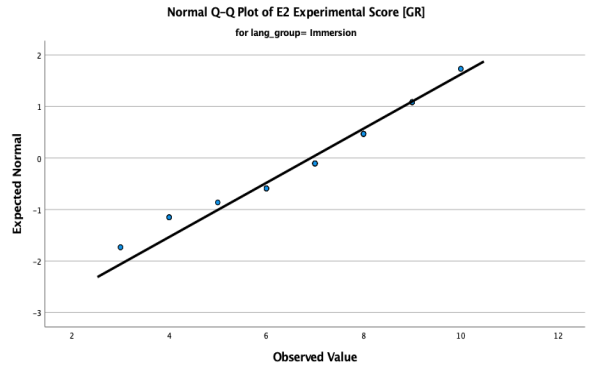


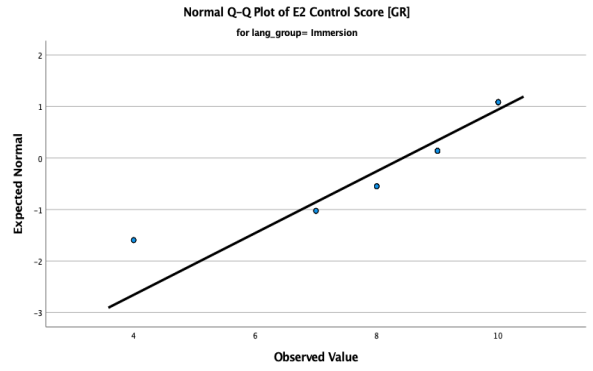
Bilingual Children (Total Immersion)





Bilingual Children (Partial Immersion)





Levene's Test (assessing homogeneity of variances) for the three groups Experiment 2 scores.

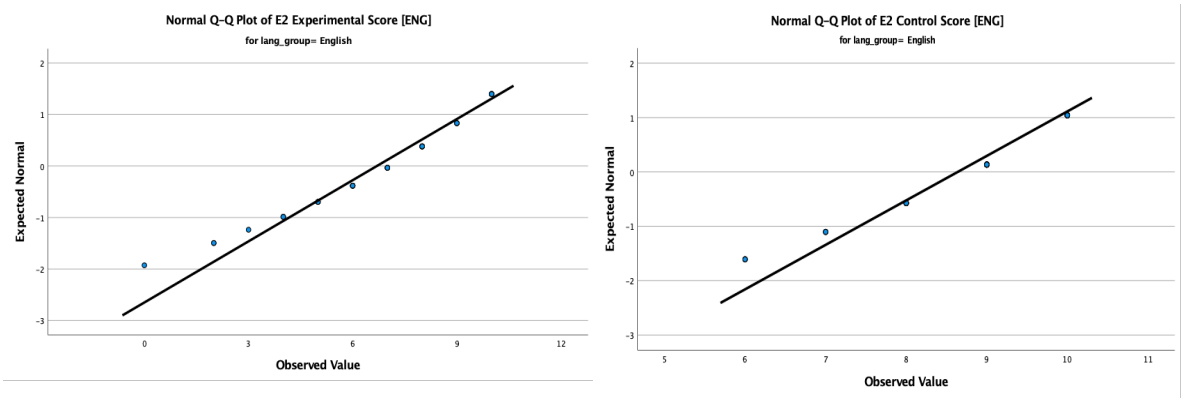
Tests of Homogeneity of Variances

		Levene Statistic	df1	df2	Sig.
E2 Experimental Score [GR]	Based on Mean	.019	2	101	.981
	Based on Median	.104	2	101	.901
	Based on Median and with adjusted df	.104	2	100.127	.901
	Based on trimmed mean	.008	2	101	.992
E2 Control Score [GR]	Based on Mean	.933	2	101	.397
	Based on Median	.673	2	101	.512
	Based on Median and with adjusted df	.673	2	79.803	.513
	Based on trimmed mean	.815	2	101	.445

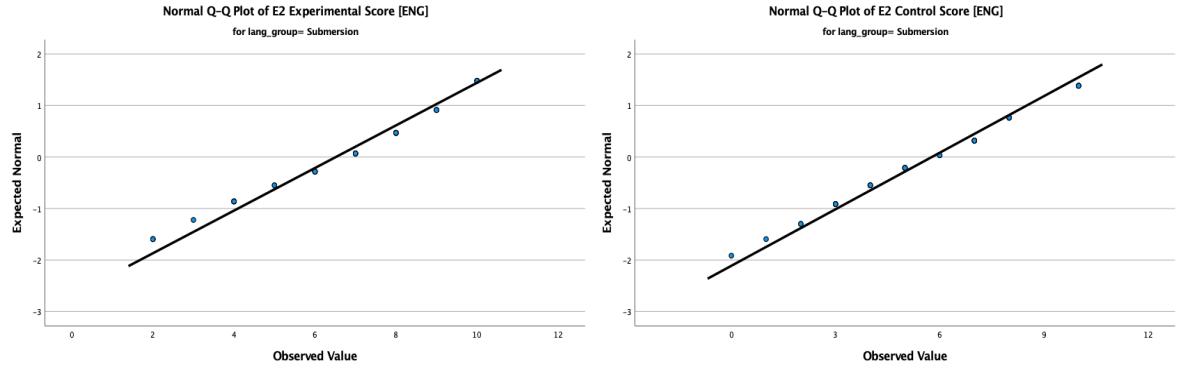
Appendix N: Normality & Homogeneity of Variances for English (Experiment 2)

Normal QQ Plots (assessing normality) of Experiment 2 scores.

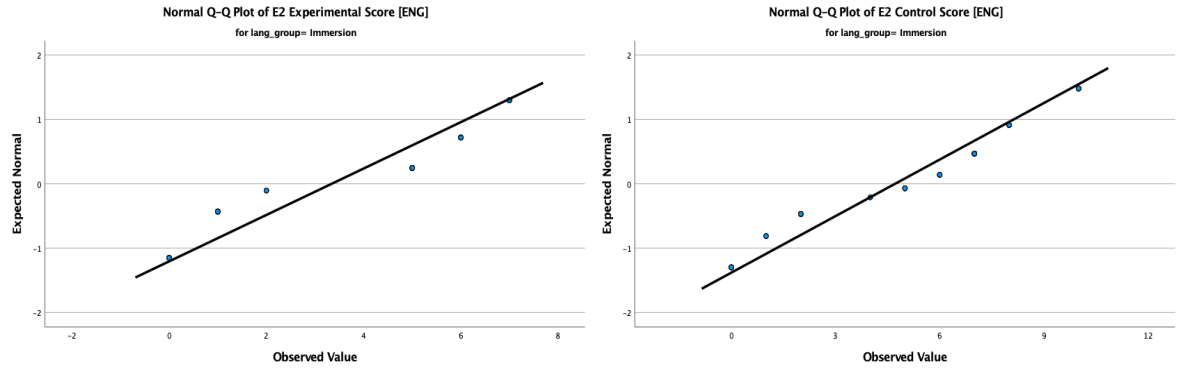
Monolingual English Children



Bilingual Children (Total Immersion)



Bilingual Children (Partial Immersion)

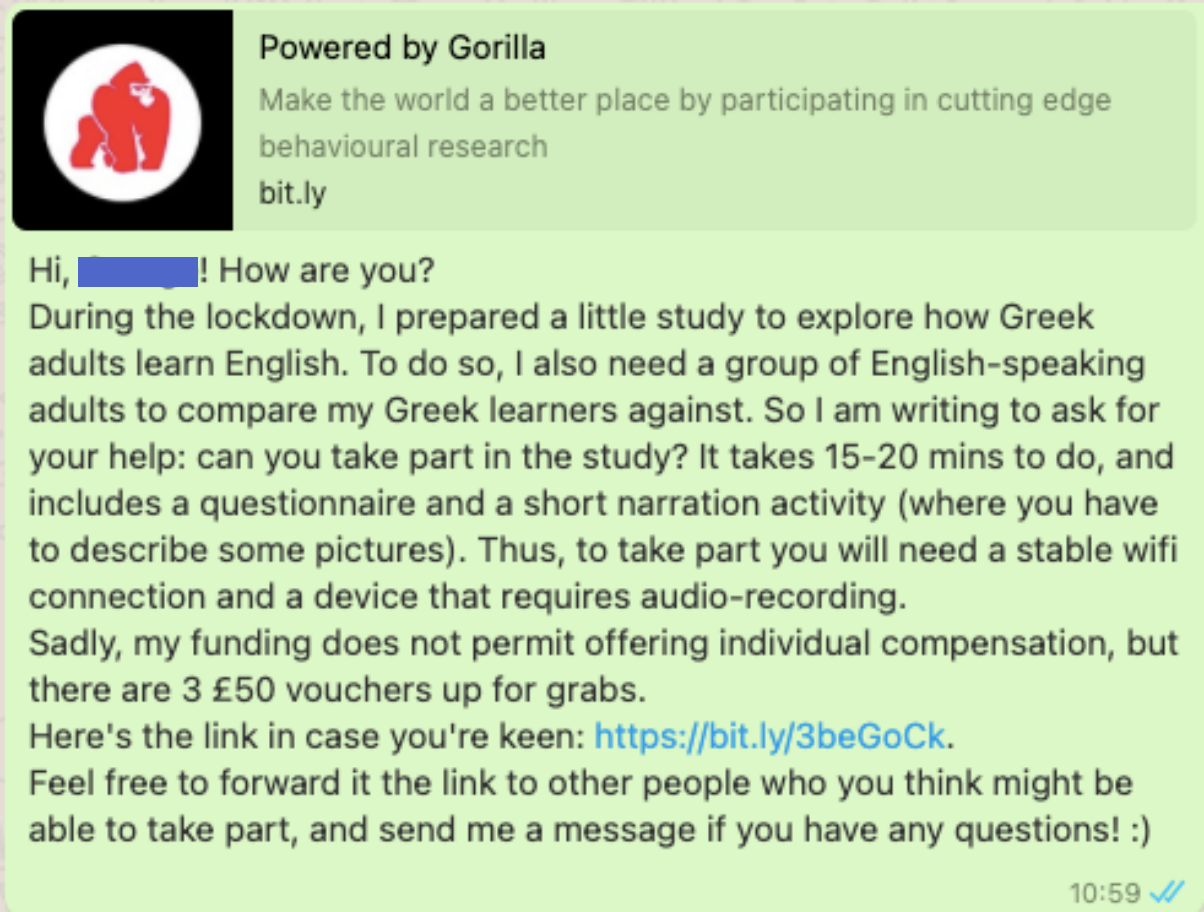


Levene's Test (assessing homogeneity of variances) for the three groups Experiment 2 scores.

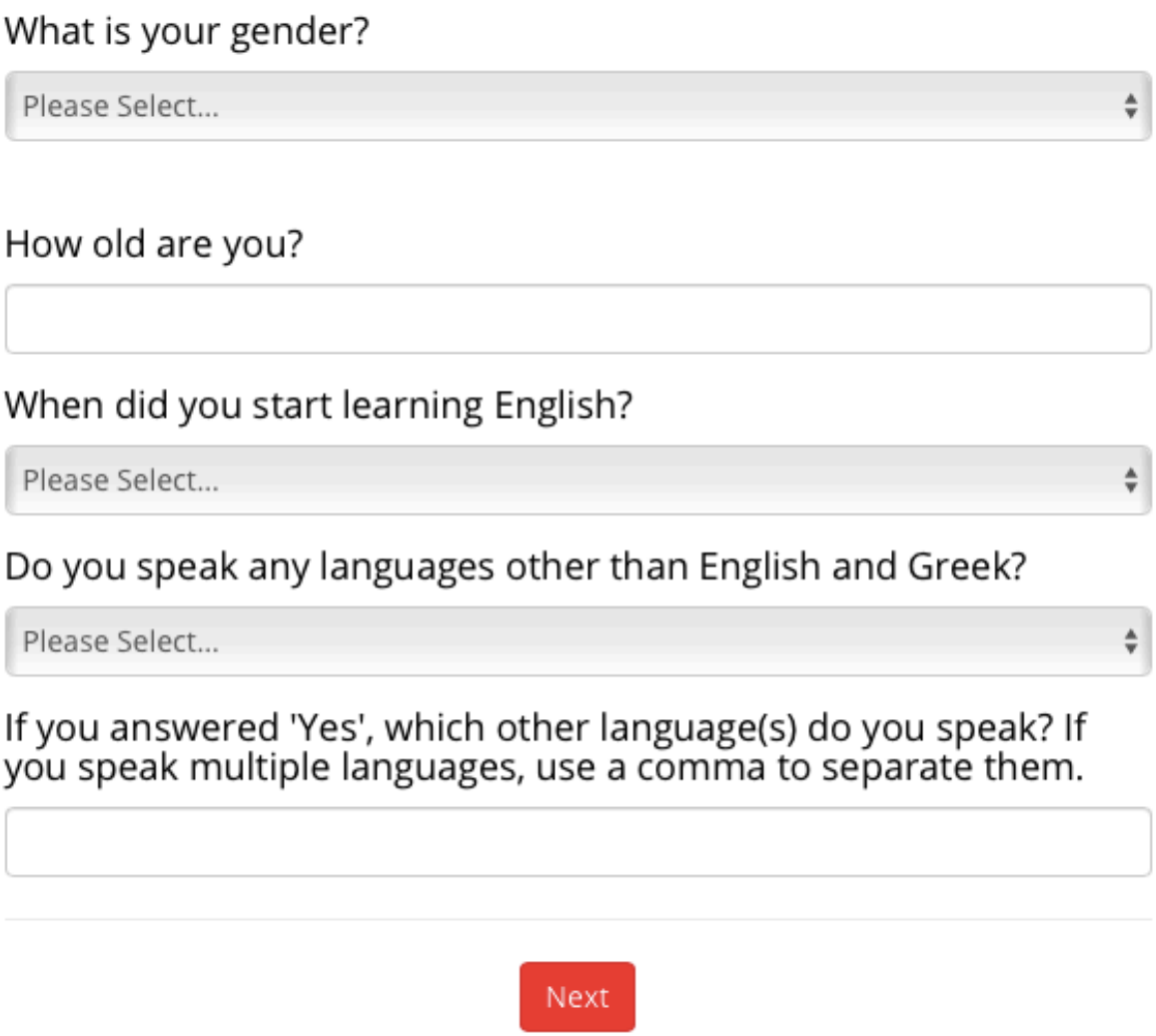
Tests of Homogeneity of Variances

		Levene Statistic	df1	df2	Sig.
E2 Experimental Score [ENG]	Based on Mean	2.809	2	103	.065
	Based on Median	1.526	2	103	.222
	Based on Median and with adjusted df	1.526	2	96.892	.223
	Based on trimmed mean	2.860	2	103	.062
E2 Control Score [ENG]	Based on Mean	22.974	2	103	.000
	Based on Median	21.019	2	103	.000
	Based on Median and with adjusted df	21.019	2	86.489	.000
	Based on trimmed mean	23.332	2	103	.000

Appendix O: Invitation to Adult Participants



Appendix P: Demographics Questionnaire



Note that this is the version of the questionnaire presented to Greek EFL learners. The version of the questionnaire presented to English L1 speakers did not contain the third question ("When did you start learning English?") and contained a modified version of the fourth question (i.e., "Do you speak any languages other than English?").

Appendix Q: Language Use Questionnaire

Note that this is the version of the questionnaire presented to Greek EFL learners. The version of the questionnaire presented to English L1 speakers was identical for page 1, but somewhat different for page 2. In particular, instead of "relative to Greek" the questions used in the English version stated: "relative to other languages".

Page 1:

Which country do you live in?

What is the main language used in your house?

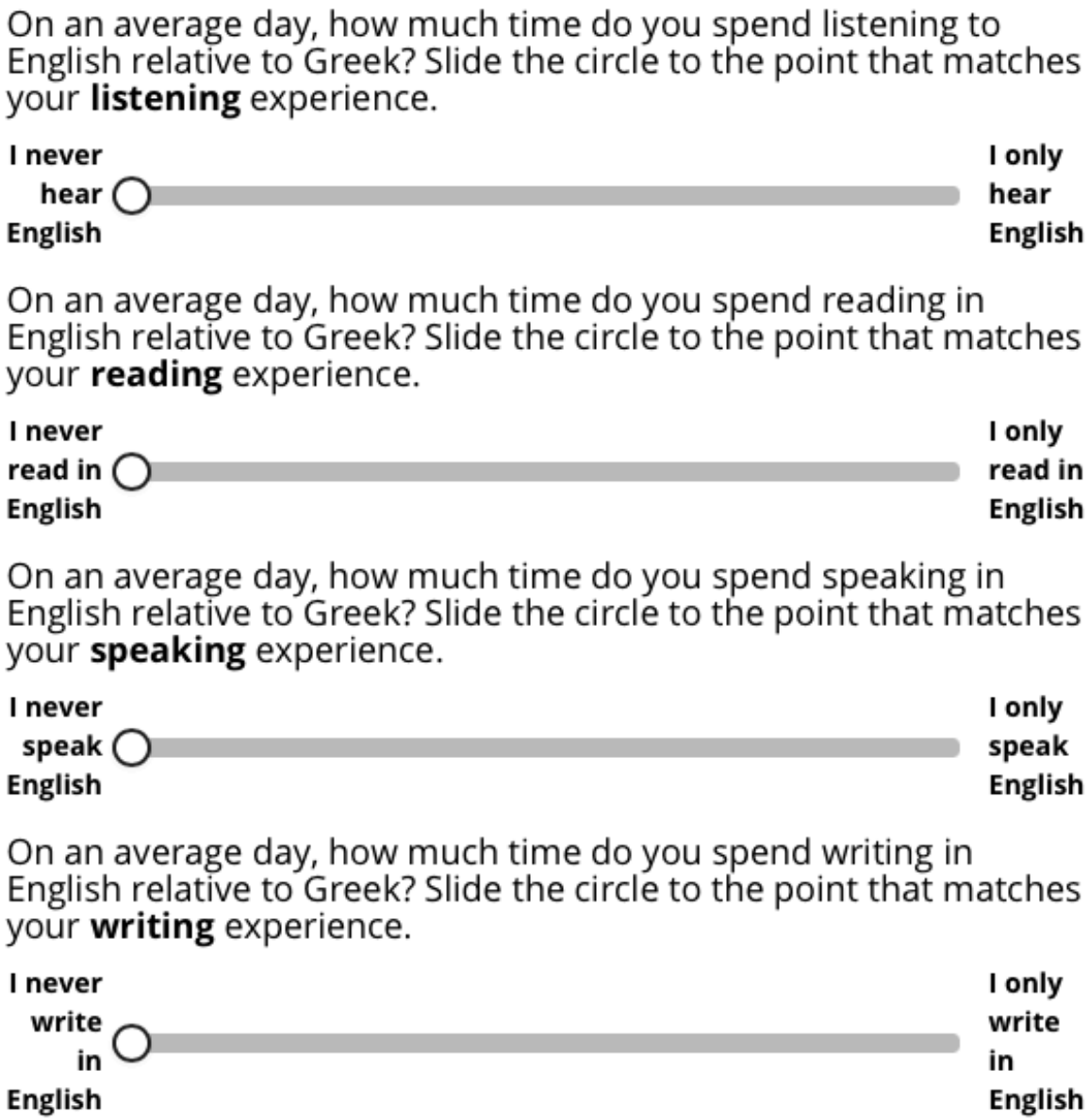
What is the main language used in your workplace?

What language do you prefer/tend to use with family?

What language do you prefer/tend to use with friends?

Next

Page 2


On an average day, how much time do you spend listening to English relative to Greek? Slide the circle to the point that matches your **listening** experience.
I never hear English
I only hear English
On an average day, how much time do you spend reading in English relative to Greek? Slide the circle to the point that matches your **reading** experience.
I never read in English
I only read in English
On an average day, how much time do you spend speaking in English relative to Greek? Slide the circle to the point that matches your **speaking** experience.
I never speak English
I only speak English
On an average day, how much time do you spend writing in English relative to Greek? Slide the circle to the point that matches your **writing** experience.
I never write in English
I only write in English
Next

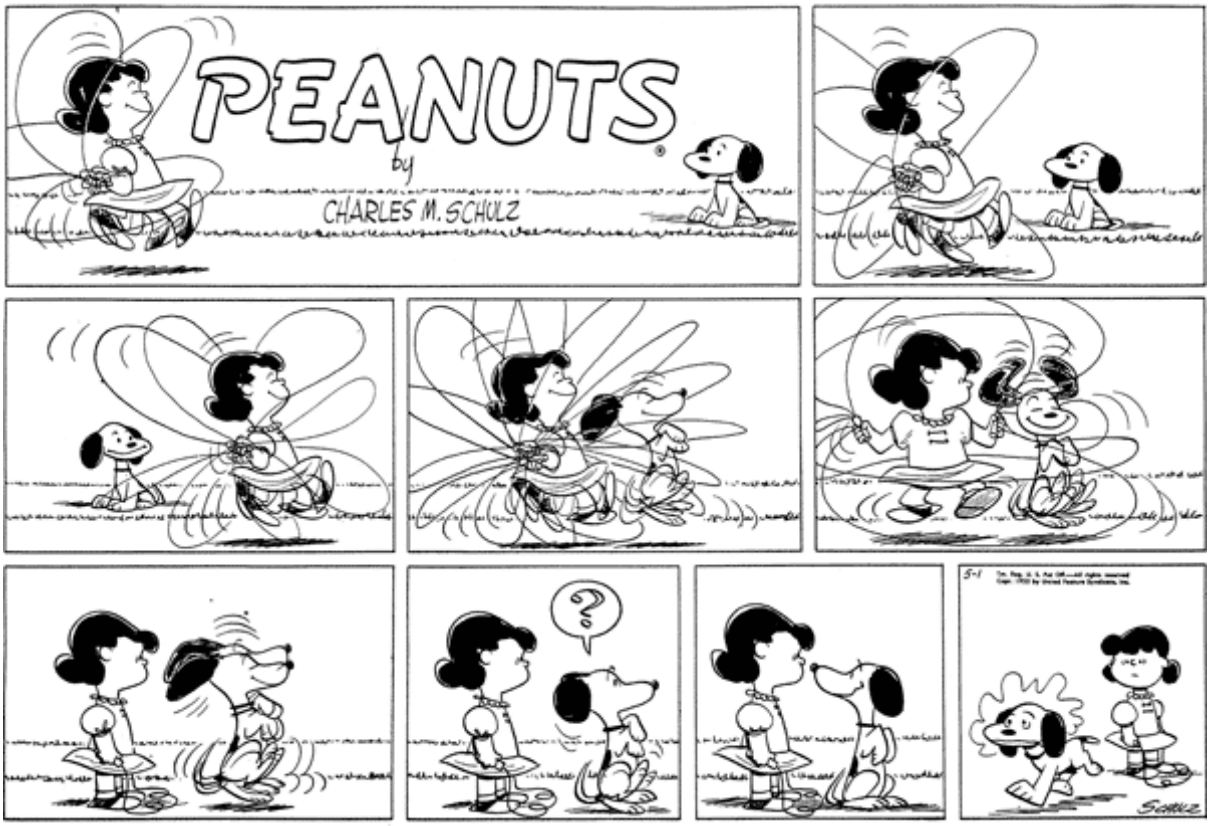
Appendix R: Generated Pseudowords

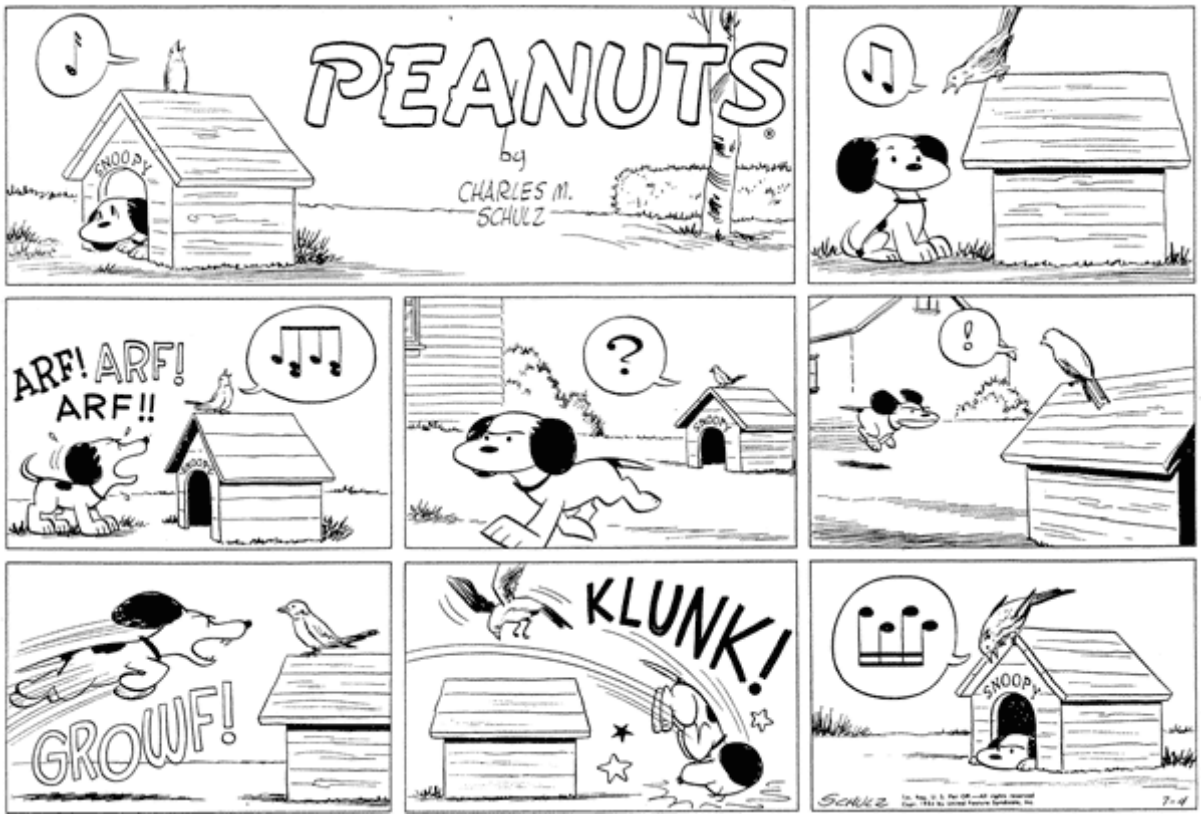
2K Heads	**3K Heads**	**4K Heads**
entrave	detailoring	regrain
wilding	homoglyph	ralling
baldock	oxylate	draconite
hapgood	custony	horozone
overend	horobin	imigate
suddery	gumm	dogmatile
haque	venn	lauder
balfour	scother	stace
duffin	abstair	quorant
twose	tooly	wray
forfert	dowrick	cambule
kiley	eldred	eckett

Appendix S: Yes/No Vocabulary Test Lists

2K Heads	3K Heads	4K Heads
baldock	abstair	aggravate
balfour	alley	bay
bracket	carol	bitch
bugger	certificate	blob
campaign	chapel	cambule
chase	crawl	corrupt
chuck	custony	depart
demand	detailoring	dispute
detect	dowrick	dogmatile
dictionary	egypt	doom
duffin	eldred	draconite eckett
entrave	enemy	fixture
escape	gumm	glen
examine	hammer	horozone
forfert	height	imigate
generate	homoglyph horobin	lauder
hapgood	invent	manifesto
haque	jog	mechanism
hunt	menu	perspective
kiley	nod	pork
opposite	oxylate	quorant
overend	scother	ralling
recall	sew	regrain
relevant	silver	rice
song	squash	socket
suddery	storm	spider
tradition	surgeon	stace
twose	tooly	superior
typical	tragic	territory
upset	venn	woo
wilding	wallpaper	wray
wing		

Appendix T: Experiment 3 Pictures


PEANUTS
by
CHARLES M. SCHULZ
?


PEANUTS
by
CHARLES M. SCHULZ
SNOOPY
ARF! ARF! ARF!!
?
!
GROWF!
KLUNK!
SNOOPY

Appendix U: Consent Form (Experiment 3)

The linguistic development of Greek learners of English

You are invited to take part in a study run by Faidra Faitaki, PhD Candidate at the University of Oxford.

The aim of the study is to investigate Greek speakers' acquisition of English as a Foreign Language relative to English speakers.

The study is funded by the United Kingdom's Economic and Social Research Council (ESRC), and has received ethics clearance by the University of Oxford's Central University Research Ethics Committee (CUREC).

To take part, you must:

1. be over 18 years old;
2. have Greek as your only first language;
3. know English well enough to understand and give answers to the questions of this form;
4. possess a device that has a stable wifi connection and a microphone (i.e. computer, tablet, phone).

If you take part in the study, you will have to:

1. answer a few questions about your experience of English;
2. complete a short task about your knowledge of English;
3. look at some comic strips;
4. describe what happens in the comic strips in an audio-recording.

At the end of the study, we will draw three £50 voucher prizes! We will randomly choose the three winners among the respondents who leave their email address at the end of the study.

All the personal information, answers to questions and audio-recordings that you provide will be anonymised (saved using a number rather than your name), and remain safe and confidential throughout the course of this study. The data will form part of Faidra's doctoral thesis, and might also be used for different projects, publications and presentations.

If you have concerns about any aspect of this study, please contact Faidra (faidra.faitaki@education.ox.ac.uk) and/or her supervisor, Professor Victoria Murphy (victoria.murphy@education.ox.ac.uk), who will do their best to answer.

If you meet the inclusion criteria, and give your consent for the data to be used as described in the previous paragraphs, please tick the box below.

Thank you for your interest in the study!

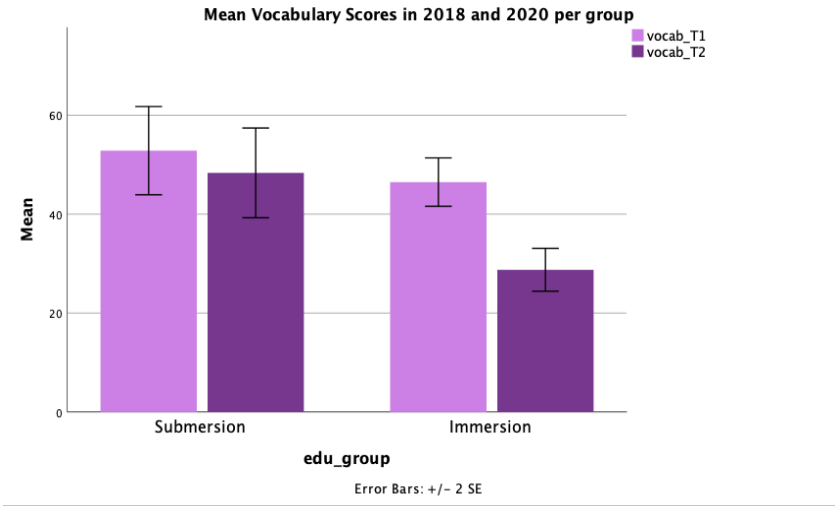
☐ **I understand the above information and I want to take part.**

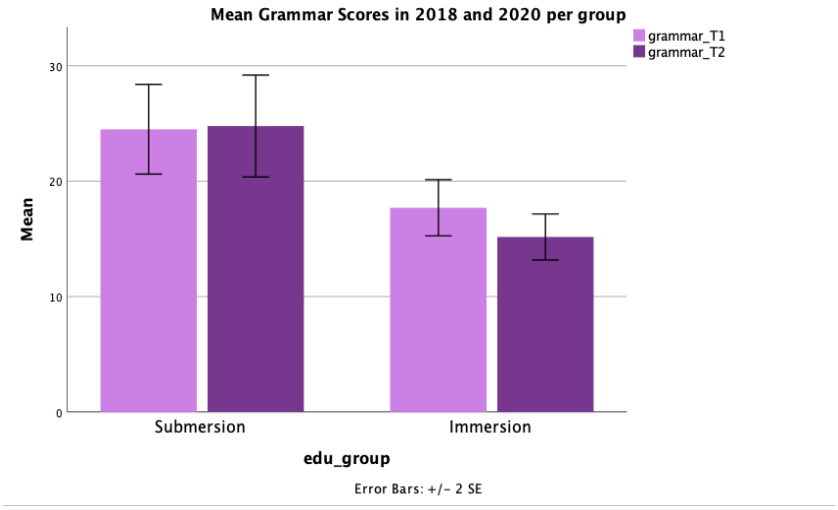
Appendix V: Comparison of 2018 and 2020 Results

Due to COVID-19, the administration of the English treatment took place online, rather than face-to-face as planned, for the bilingual groups that were included in the child study. This methodological change is substantial, and could have impacted the obtained results. The substantive pilot data presented in the above section can be used to determine whether and to what extent children's performance was altered as a result of the administration method.

Although the children tested in 2018 (when the pilot took place) and 2020 (when data collection proper was conducted) were not the same, there were enough contextual similarities to safeguard the accuracy of the comparison: namely, the schools that participated in both studies were the same, and the measures used were similar – an exception being the PPVT, which was substituted by the BPVS, as the latter was deemed more suitable given the characteristics of the investigated population.

Thus, two paired-samples t-test were conducted, one for vocabulary and one for grammar. As the two figures below illustrate, for vocabulary, the difference between children's 2018 and 2020 performance, 12.33 (SD=14.58), was significant and corresponded to a large effect: $p<.001$, $d=.85$. Further paired-samples t-tests conducted to explore this effect per education group suggested that both the submersion and the immersion children had lower vocabulary scores in 2020. In particular, the submersion group had a mean difference of 4.50 (SD=7.96) between the two timepoints, which was significant and represented a medium effect: $p=.02$, $d=.57$. A larger and significant mean difference, 17.72 (SD=15.73), was observed for the immersion group: $p<.001$, $d=1.13$. Turning to children's performance on the TROG, the mean difference between 2018 and 2020 scores, 1.48 (SD=7.23) was not significant: $p=.16$, $d=.20$.





In sum, it appears that children's English vocabulary suffered a greater loss than their grammar in the aftermath of COVID-19 (and as a result of the school closures). The reason behind this pattern might be that the acquisition (and retention) of vocabulary requires that children experience words in numerous and varied contexts, while the acquisition of grammar does not. As vocabulary is more likely to be impacted by frequency, the reduction in input for a prolonged period of time that the participants who completed the study in 2020 had to deal with could have resulted in lower scores. Frequency does not have the same impact on grammar – once the child has been exposed to, abstracted and represented a rule or pattern, he/she can retain it; this observation can account for the fact that the loss of L2 grammatical knowledge during lockdown was not as grave.

In addition to the reduction in input, another explanation for these findings could be that the vocabulary test is less stable in online settings – after all, it is designed for in-person administration and, as such, children's lower vocabulary scores in 2020 could be the result of the reduced validity of the procedure. However, if this was the case, children's grammar scores in 2020 should also have been lower. Indeed, the BPVS/PPVT and the TROG have similar formats: they are multiple choice tests that assess children's receptive knowledge of vocabulary or grammar. Thus, the lack of a drop in children's grammar scores speaks against there being general issues with the online format (and its validity).