



Università
Ca' Foscari
Venezia

Master's Degree
in Language Sciences

Final Thesis

**The Acquisition of *Wh*-Questions
in Italian Learners of English:
Evidence from Elicited Written
Production**

Jessie Quinn

Supervisor

Prof. Giulia Bencini

Co-supervisors

Prof. Virginia V. Valian

Michaela M. Vann,

PhD Student in Language Sciences

Student number

874173

Academic Year

2019 / 2020

ABSTRACT

The particular syntactic structure of English questions has made them a topic of interest in both first language (L1) and second language (L2) acquisition research. *Wh*-questions pose difficulties in acquisition for both monolingual English-speaking children and second language learners of English. English main questions require subject-auxiliary inversion while embedded questions do not allow for it. Studies have shown that main questions are acquired before embedded questions and, due to the structural asymmetry between the two, learners often incorrectly apply subject-auxiliary inversion in embedded contexts. The goal of the present study is to investigate how intermediate/advanced L2 English learners acquire both main and embedded *wh*-questions, focusing on subject-auxiliary inversion errors. In addition to analyzing sentence structure type, the present study examines the effect of *wh*-word, focusing on the arguments *what*, *who*, and *whose* and the adjuncts *where* and *why*. Through an elicited written production task, L1 Italian participants were asked to produce both main and embedded English questions. The results show low inversion rates in *why*-questions, a phenomenon that has been noted in both L1 and L2 acquisition research, higher rates of inversion in arguments in embedded questions, and particular difficulty acquiring and producing *whose*-questions. Within errors of non-inversion in embedded questions, learners overgeneralize obligatory subject-auxiliary inversion from English main questions to embedded questions, rather than transferring L1 properties, indicating an interference of the L2. Furthermore, pilot data was gathered from highly advanced/bilingual L2 learners living in an English-speaking country. Overall, participants produced more correct responses in both experiments compared to the learners in a non-immersive context. However, the results indicate that errors in *whose*-questions and errors of non-inversion in embedded questions persist in highly advanced, bilingual speakers.

Italian summary

La struttura sintattica delle domande interrogative in inglese presenta delle peculiarità e, per questo motivo, un'ampia letteratura nel campo dell'acquisizione del linguaggio si concentra sulle strutture interrogative; sia per quanto riguarda l'acquisizione dell'inglese come lingua nativa (L1), sia per quanto riguarda l'acquisizione dell'inglese come seconda lingua (L2). In letteratura, diversi studi si sono concentrati sull'acquisizione delle interrogative inglesi da parte di bambini L1, rivelando una difficoltà nell'acquisizione di queste strutture sintattiche simile a quella dimostrata da parlanti bilingui apprendenti inglese L2.

Principalmente, la difficoltà nel padroneggiare le interrogative dirette *wh-* è dovuta dalla presenza del verbo ausiliare che antecede il soggetto, seguito dal verbo principale. In altre parole, la formulazione di una struttura interrogativa diretta inglese richiede un'inversione del soggetto e del verbo ausiliare. Diversamente, le domande indirette non ammettono la stessa inversione; pertanto, è possibile affermare che esista un'asimmetria fra le strutture sintattiche delle interrogative dirette e indirette.

Al contrario, la lingua italiana non prevede lo stesso tipo di asimmetria tra una struttura interrogativa diretta e indiretta. Infatti, sia le domande dirette che quelle indirette richiedono l'inversione del verbo e del soggetto. Mentre in inglese la formulazione di una domanda prevede l'inversione del soggetto e del verbo ausiliare, in italiano è necessaria l'inversione del soggetto (se fosse espresso) e del verbo principale o dell'intero sintagma verbale. In altre parole, tutte e due lingue prevedono l'inversione, però l'elemento che si inverte è diverso.

Nello studio del bilinguismo esistono diverse teorie che discutono l'influenza della lingua nativa sulla L2. Alcuni studi affermano che i parametri della L1 influenzano la L2 e che i parlanti trasferiscono le strutture dalla prima lingua nelle lingue successivamente acquisite. In letteratura, questo fenomeno è tipicamente definito *L1 transfer*.

Nello specifico, diversi studi che si sono concentrati sull'acquisizione delle domande inglesi da parte di parlanti L2 hanno osservato che le domande dirette vengono tipicamente acquisite prima delle domande indirette (Pienemann, 1998; Pienemann & Keßler, 2011; Pozzan, 2011; Pozzan & Quirk, 2014) e che vi è una comune tendenza nell'inversione del verbo ausiliare e del soggetto nelle strutture interrogative indirette, nonché la conseguente produzione di frasi agrammaticali. È dunque possibile affermare che la difficoltà nell'acquisizione delle strutture interrogative inglesi sta nel sapere *quando* e *cosa* invertire.

Inoltre, diversi studi hanno osservato un'asimmetria nella produzione delle domande introdotte dall'elemento interrogativo *why*, in quanto si osservano percentuali di inversione più

basse, se paragonate alle strutture interrogative introdotte da altri elementi *wh-* (*what, when, where, whose e who*).

Rizzi (2001) ha dimostrato che la struttura sintattica introdotta da *perché* e *come mai* presenta delle asimmetrie rispetto alle domande introdotte da altri elementi interrogativi, in quanto è rappresentata come un *higher adverbial* (e cioè, l'elemento *perché* è rappresentato come un avverbio che occupa una posizione più alta nella struttura sintattica della frase) e per questo motivo non richiede l'inversione. Inoltre, la mancata inversione soggetto-ausiliare per quanto riguarda le interrogative introdotte da *why* si osserva omogeneamente in diverse popolazioni di parlanti con diverse L1: questo dimostra che la ragione per la quale gli apprendenti italiani tendono a non invertire questi due elementi in fase di produzione non può essere attribuita solamente al fenomeno di *L1 transfer*.

Lo studio presentato in questa tesi indaga come parlanti nativi italiani e apprendenti inglese L2 acquisiscono, e dunque producono, le strutture interrogative in inglese. Attraverso due esperimenti, sul modello dello studio di Pozzan (2011) e Pozzan e Quirk (2014), questo studio mira ad analizzare gli errori nella produzione scritta delle domande interrogative, focalizzandosi nello specifico sugli errori di inversione soggetto-ausiliare. Per indagare se questi errori siano possibilmente dovuti ad un effetto dell'elemento *wh-* che introduce la domanda, nel materiale sono state incluse cinque *wh-words*: gli argomenti *what, who* e *whose* e gli aggiunti *where* e *why*. Inoltre, al fine di osservare se gli errori in produzione siano dovuti al verbo contenuto nella frase, il materiale sperimentale è stato costruito includendo sia i verbi ausiliari (*to be; to have*) che i verbi lessicali.

Lo studio si rivolge a parlanti nativi italiani con intermedi e/o avanzati livelli di competenza in inglese ed è stato diviso in due esperimenti, entrambi svolti sulla piattaforma online Pavlovia. Il primo esperimento è stato progettato per elicitarne la produzione di domande dirette in inglese ed è stato strutturato come segue: ai partecipanti è stato presentato uno studente dal nome Phil, descritto come un personaggio timido e che aveva dunque bisogno di aiuto per poter porre delle domande al suo insegnante, Ms. Smith.

Il compito dell'esperimento prevedeva dunque l'esposizione a frasi come: "*Jane is reading something to the students. Ask Ms. Smith what.*" e successivamente, ai partecipanti veniva chiesto di produrre la domanda interrogativa diretta in inglese. Pertanto, la domanda a seguito dell'esempio sopra presentato sarebbe: "*What is Jane reading to the students?*".

Nel secondo esperimento, al fine di elicitarne la produzione di domande interrogative indirette, ai partecipanti è chiesto di aiutare la curiosa Bridget, che era interessata a sapere cosa Phil stesse chiedendo a Ms. Smith. In linea con la routine precedentemente presentata, ai

partecipanti è stata mostrata una domanda diretta come: “*Why is Mark complaining to the server?*” e nella schermata successiva, Bridget chiedeva: “*What does Phil want to know?*”. Le risposte dei partecipanti venivano elicitate con la presentazione dell’*incipit* della struttura interrogativa indiretta sulla terza schermata: “*Phil wants to know...*”. Di conseguenza, la risposta target prevista sarebbe: “*Phil wants to know why Mark is complaining to the server.*”

I risultati di entrambi gli esperimenti sono stati analizzati sulla base delle percentuali di interrogative dirette e indirette corrette prodotte dai partecipanti, nonché sulle percentuali di inversione soggetto-ausiliare. Pertanto, il numero totale di interrogative corrette osservate è stato diviso per il numero totale di risposte prodotte dai partecipanti. Al fine di calcolare la percentuale di domande interrogative che rispettassero l’inversione soggetto-ausiliare, il numero totale di interrogative presentanti tale inversione, fossero esse grammaticali e non, è stato diviso per il numero totale di interrogative che permettevano la valutazione invertita. In altre parole, le risposte che non permettevano la valutazione invertita, per esempio una risposta che manca il verbo ausiliare, non sono state considerate nell’analisi.

I risultati hanno mostrato una maggior percentuale di risposte corrette prodotte nel primo esperimento, e quindi, la percentuale di domande dirette corrette è maggiore rispetto alla percentuale di domande indirette corrette prodotte nel secondo esperimento. È dunque possibile affermare che i risultati dei due esperimenti confermano l’ipotesi secondo la quale le strutture delle domande dirette sono acquisite prima delle strutture delle domande indirette.

Inoltre, l’analisi delle risposte ha evidenziato una minor percentuale di inversione nella produzione di domande dirette e indirette introdotte da *why*. Nella produzione di domande dirette non è stata osservata alcuna asimmetria argomento-aggiunto. Le percentuali di risposte corrette e invertite degli argomenti e gli aggiunti erano infatti comparabili.

Nel secondo esperimento, sono state osservate più risposte corrette (e dunque non invertite) nella produzione di interrogative aggiuntive, se paragonate alla produzione di interrogative argomentative: è stata dunque osservata un’asimmetria argomento-aggiunto nella produzione di domande indirette.

Inoltre, i partecipanti hanno mostrato difficoltà nella produzione delle domande interrogative introdotte da *whose*. In questo caso, le interrogative inglesi presentano una struttura particolare che si discosta dalla struttura interrogativa italiana introdotta da *di chi*. Per questo motivo, è stata ipotizzata una più alta percentuale errori nella produzione di questa struttura. I risultati ottenuti hanno confermato l’ipotesi, infatti è stata osservata una bassa percentuale di risposte corrette e invertite introdotte da *whose* nell’esperimento 1; nel secondo

esperimento, la percentuale di risposte corrette e non invertite era più bassa di quella ottenuta nel primo esperimento.

Nell'analisi dei dati, abbiamo focalizzato la nostra attenzione sulla produzione di strutture marcate come: “*Whose is the book that Joe is carrying to class?*?”. Sebbene queste strutture marcate siano state osservate in produzione da parte di parlanti nativi inglesi, abbiamo deciso di condurre un ulteriore studio per investigare l'accettabilità delle interrogative introdotte da *whose* presentanti la struttura marcata precedentemente esemplificata. I risultati hanno rivelato che i parlanti inglese L1 preferiscono la struttura non marcata come: “*Whose book is Joe carrying to class?*”

Per esplorare ulteriormente *quando* e *come* parlanti nativi italiani apprendenti inglese acquisiscano le domande *wh-* in inglese, gli esperimenti sono stati distribuiti a parlanti italiani esposti all'input linguistico in contesto L2 da almeno due anni con un alto livello di competenza in inglese. I dati iniziali ottenuti dal *pilot* dell'esperimento hanno mostrato che gli errori di non inversione prodotti nelle domande indirette, così come gli errori prodotti nelle domande introdotte da *whose*, persistono anche da parte parlanti con alti livelli di competenza ed esposti all'input frequentemente.

Per concludere, è possibile affermare che i risultati dello studio presentato in questa tesi supportano l'ipotesi secondo la quale le strutture delle domande dirette sono acquisite prima delle interrogative indirette e inoltre, confermano l'esistenza di un'asimmetria nella percentuale di errore osservata nelle domande introdotte da *why*: se paragonate alle domande introdotte dagli altri elementi *wh-*, si osserva una percentuale di inversione soggetto-ausiliare più bassa sia nelle domande dirette che indirette.

Inoltre, i partecipanti hanno rivelato una difficoltà nella produzione delle domande introdotte da *whose*, producendo sia frasi agrammaticali, che strutture marcate rispettanti l'ordine degli elementi delle strutture italiane introdotte da *di chi*. Infatti, nella produzione di domande indirette, i partecipanti hanno applicato la regola dell'inversione soggetto-ausiliare senza applicare l'inversione dell'intero sintagma verbale, come previsto dalla lingua italiana. Pertanto, è possibile assumere che gli errori di non inversione osservati nella produzione delle domande indirette siano principalmente dovuti ad una interferenza della L2, e non al fenomeno di *L1 transfer*.

Infine, gli errori non inversione prodotti nelle domande indirette, nonché errori prodotti nelle interrogative che esprimono il possesso (*whose*) sono stati osservati in produzione da parte di parlanti con alti livelli di competenza in inglese e frequentemente esposti all'input linguistico in contesto L2. È perciò plausibile concludere che queste strutture siano

difficilmente padroneggiate da nativi italiani e che ai fini di una corretta acquisizione di queste peculiari interrogative, sarebbe necessario un insegnamento esplicito della struttura sintattica introdotta da *whose* e della struttura sintattica delle interrogative indirette.

ACKNOWLEDGMENTS

Writing this thesis has been a very long and stressful process but I have learned so much and it's been so worth it. Much of the time was spent in lockdown and the support of loved ones near and far has made completing this thesis in such an unprecedented time possible.

Firstly, I'd like to thank my parents for always encouraging me and being the best examples of how to selflessly support your child when she decides to move thousands of miles away. I wish you could be here.

I'd like to thank my supervisor, Professor Giulia Bencini, for motivating me and opening my eyes to the world of psycholinguistics. Thank you to my co-supervisor, Professor Virginia Valian, for the invaluable advice and help throughout the whole process. To my co-supervisor, Michaela Mae Vann, thank you for the constant support and guidance. I look up to each one of you and I feel very privileged to have learned from you all.

I'd like to also thank my friends. To Emma, thank you for always entertaining me with a post library spritz when I was in need. To Bree, thank you for the frequent videocalls and support through everything. Thank you to Bridget, Teal, Connor, Sophia, Katie and everyone else who piloted my experiments, helped me recruit participants and listened to me talk about this project for months.

Finally, thank you to Martina Pucci for the technical and coding support and to my fellow MA students, Giulia Sfriso, Ilaria Venagli, and Micol Zanaga, for providing a collaborative working environment and helping me with Italian when I needed it. It's been a pleasure working with you all.

TABLE OF CONTENTS

Chapter 1. Introduction	12
1.1 L2 acquisition	12
1.1.1 Theories of L2 acquisition	13
1.2 English questions	15
1.2.1 English main questions	15
1.2.2 English embedded questions	16
1.3 Italian questions	17
1.3.1 Italian main questions	17
1.3.2 Italian embedded questions	19
1.4 English and Italian comparison	20
1.4.1 Comparing English <i>whose</i> to Italian <i>di chi</i>	21
1.5 L2 acquisition of English questions	23
1.5.1 Effect of L1 transfer on L2 acquisition	25
1.5.2 Effect of <i>wh</i> -word	28
Chapter 2. Current study	30
2.1 Experiment 1: Main questions	30
2.1.1 Method	31
2.1.1.1 Participants	31
2.1.1.2 Materials	32
2.1.1.3 Procedure	33
2.1.1.4 Transcription and coding	34
2.1.2 Results and discussion	35
2.1.2.1 <i>Wh</i> -word	37
2.1.2.2 Verb type	39
2.1.2.3 <i>Whose</i>	39
2.1.2.4 Proficiency level	40
2.2 Experiment 2: Embedded questions	41
2.2.1 Method	42
2.2.1.1 Participants	42
2.2.1.2 Materials	42
2.2.1.3 Procedure	44
2.2.1.4 Transcription and coding	44
2.2.2 Results and discussion	45

2.2.2.1 <i>Wh</i> -word.....	48
2.2.2.2 Verb type.....	49
2.2.2.3 <i>Whose</i>	49
2.2.2.4 Proficiency level	50
2.3 L2 speakers in an immersive context	51
2.3.1 Method.....	51
2.3.1.1 Participants.....	51
2.3.1.2 Materials	52
2.3.1.3 Procedure	52
2.3.1.4 Transcription and coding	53
2.3.2 Results and discussion	53
2.3.2.1 Experiment 1: main questions.....	53
2.3.2.2 Experiment 2: embedded questions	54
2.4 Acceptability judgment task on English questions	56
2.4.1 Method.....	57
2.4.1.1 Participants.....	57
2.4.1.2 Materials	57
2.4.1.3 Procedure	59
2.4.2 Results and discussion	59
Chapter 3. Discussion	62
3.1 Question type.....	62
3.2 L1 transfer	64
3.3 <i>Wh</i> -word.....	65
3.4 Proficiency	66
Chapter 4. Conclusions.....	68
4.1 Answered questions.....	68
4.2 Unanswered questions and opportunities for follow-up studies	69
4.3 Implications for teaching English as a second language.....	70
Appendix A	71
Appendix B	74
Appendix C	75
Appendix D.....	77
Appendix E	78
Appendix F	79

Appendix G.....85
References.....86

List of Tables

Table 1: Word order in English and Italian *wh*-questions 21
Table 2: English *whose* and Italian *di chi* comparison 23
Table 3: Order of acquisition of questions in L2 English..... 24
Table 4: L1 Italian participant demographic data 32
Table 5: L2 learners' written production of main *wh*-questions by coding category (strict) .. 36
Table 6: L2 learners' written production of main *wh*-questions by coding category (lax)..... 37
Table 7: L2 learners' written production of main *wh*-questions by *wh*-word..... 37
Table 8: L2 learners' written production of main *wh*-questions by *wh*-type..... 38
Table 9: L2 learners' written production of main *wh*-questions by *wh*-type (excluding *whose*)
..... 38
Table 10: L2 learners' written production of main *wh*-questions by verb type..... 39
Table 11: L2 learners' written production of *whose*-questions by CEFR self-rate level..... 40
Table 12: L2 learners' written production of main *wh*-questions by self-rated proficiency ... 40
Table 13: L2 learners' written production of embedded *wh*-questions by coding category
(strict)..... 46
Table 14: L2 learners' written production of embedded *wh*-questions by coding category (lax)
..... 47
Table 15: L2 learners' written production of embedded *wh*-questions by *wh*-word 48
Table 16: L2 learners' written production of embedded *wh*-questions by *wh*-type 49
Table 17: L2 learners' written production of embedded *wh*-questions by verb type 49
Table 18: L2 learners' written production of embedded *wh*-questions by self-rated
proficiency 50
Table 19: L1 Italian participants living in an immersive environment demographic data 52
Table 20: L2 immersive learners' written production of main *wh*-questions by coding
category (strict) 53
Table 21: L2 immersive learners' written production of main *wh*-questions by coding
category (lax) 54
Table 22: L2 immersive learners' written production of main *wh*-questions by *wh*-word..... 54
Table 23: L2 immersive learners' written production of embedded *wh*-questions by coding
category (strict) 55

Table 24: L2 immersive learners' written production of embedded <i>wh</i> -questions by coding category (lax)	55
Table 25: L2 immersive learners' written production of embedded <i>wh</i> -questions by <i>wh</i> -word	56
Table 26: L2 immersive learners' written production of embedded <i>wh</i> -questions by <i>wh</i> -type	56
Table 27: Native English speakers' mean age by English variety	57
Table 28: Native English speakers' average rating of <i>whose</i> marked and unmarked structures by English variety and sentence type	59
Table 29: Average ratings of <i>whose</i> marked structures and ungrammatical structures by English variety and sentence type	60
Table 30: L2 learners' inversion rates in main <i>wh</i> - and embedded <i>wh</i> -questions by participant	63
Table 31: L2 immersive learners' inversion rates in main <i>wh</i> - and embedded <i>wh</i> -questions by participant	64

List of Figures

Figure 1: L2 learners' written production of main <i>wh</i> -questions by coding category and <i>wh</i> -word (strict).....	38
Figure 2: Correct responses of main <i>wh</i> -questions by CEFR self-rated level	41
Figure 3: L2 learners' written production of embedded <i>wh</i> -questions by coding category and <i>wh</i> -word (strict).....	48
Figure 4: Correct responses of embedded <i>wh</i> -questions by CEFR self-rated level.....	51

1. INTRODUCTION

English questions have long been popular topics of study within language acquisition research due to their syntactic complexity. The syntactic structure of English questions poses difficulties in acquisition for both monolingual English-speaking children acquiring English, and those who learn English as a non-native language. The study presented in this thesis examines the acquisition of English questions by non-native English speakers. The experiments included in the study were designed to elicit written production of both English main clause and embedded clause questions. Through analyzing the production errors made by the non-native English learners, it's possible to gain insight as to how second language learners acquire these structures and what factors have an effect on their acquisition and production. Before presenting the present study and results, it is necessary to understand how learners acquire a second language and, specifically, how learners acquire the structure of English questions. The present study deals with native Italian speakers and therefore, the following sections will present an overview of both English and Italian interrogative structures as both are relevant to interpreting the errors made in the production of English questions by native Italian speakers. After discussing the language acquisition literature, the structure of questions in English and Italian and reviewing previous findings regarding the acquisition of English questions, Chapter 2 discusses the current study and results. Chapters 3 and 4 discuss the present study's findings, connecting the results to previous research, attempting to understand how Italian learners of English acquire English questions, what are the major factors that affect their acquisition and what does this mean for second language learners as a whole.

1.1. L2 Acquisition

The main focus of the current study is to investigate how second language (L2) learners of English acquire the structure of English questions. To begin, it is necessary to understand the basics of first language (L1) acquisition in order to understand how one acquires an L2. The two main theories concerning the acquisition of an L1 are known as empiricism and nativism. The former claims that language is acquired through exposure to the input. Children learn their mother tongue and acquire adultlike levels of grammar through exposure to syntactic features through the input alone. On the other hand, the nativist view claims that children are born with an innate, abstract grammar in which they structure and form into the L1 grammar based on the input. Support for nativism comes from studies showing that children quickly acquire intricate concepts and structures that cannot be accounted for simply due to the input (see Bencini & Valian, 2008; Fisher et al., 2010; Stromswold, 1990; Valian & Casey, 2003).

Because of this, there must be an underlying innate grammar that allows the child to make abstract generalizations and assumptions when acquiring her L1 (White, 2003). For example, Stromswold (1990) found that from an early age, English-speaking children are able to distinguish the auxiliaries *be*, *do* and *have* from their main verb counterparts. The English auxiliary system is elaborate and difficult to acquire. Stromswold attributes children's quick acquisition of the auxiliary system as a reflection of Universal Grammar.

Universal Grammar (UG) was first introduced by Chomsky in the 1960s (see Chomsky, 1965, 1986). UG is defined as the built-in linguistic principles and parameters that constrain and structure a speaker's grammar. In terms of first language acquisition, UG is the child's initial state. It is the underlying abstract system a child unconsciously utilizes in order to constrain and form her grammar based on the received input. These principles and parameters account for the unlearnable aspects of language that are too complex to be acquired without some sort of innate faculty (White, 2003).

In L1 acquisition there is a problem in the mismatch between the input and the output, the output being the child's grammar. UG is what accounts for this mismatch. White (2003) explores if this mismatch is also a problem in L2 acquisition. That is, if UG principles hold when a learner is acquiring her second language. Interlanguage grammar is a term used to refer to a learner's non-native, L2, grammar. The same principles and parameters of UG that constrain an L1 grammar are said to constrain the interlanguage grammar (White, 2003). The question of L1 influence on learning an L2 deals with to what extent a learner can restructure these set parameters and principles, if at all, in order to acquire the grammar of the L2. If the phenomenon is not available in the L1 and cannot be acquired from input alone, it can be said that full acquisition of this phenomenon is due to UG. The level of availability of UG and the extent of L1 influence on L2 acquisition is still debated, with different theories claiming different degrees of involvement.

1.1.1 Theories of L2 acquisition

Current theories of L2 acquisition, with the exception of Full Transfer/No Access (Bley-Vroman, 1989; Schachter, 1990) generally agree that learners of a second language are able to access UG, but the level of availability and whether or not there is influence from the L1 on the L2 processes are still debated.

One theory that claims a large role of the L1 is Full Transfer/Full Access (Schwartz & Sprouse, 1994, 1996). This hypothesis claims that the L1 grammar is the initial state of L2 acquisition (Full Transfer). The influence of the L1 is active throughout L2 development.

According to Full Transfer/Full Access, learners draw from UG when their L1 grammar is insufficient (Full Access). L2 acquisition is determined by the L1 grammar, the L2 input and UG.

In opposition to Full Transfer/Full Access, the Minimal Trees hypothesis (Vainikka & Young-Scholten, 1994, 1996) denies any L1 influence on the acquisition of the L2. This hypothesis claims that L2 learners start with a lexical grammar without functional categories. These functional projections develop later during acquisition and are dependent on the L2 input. The authors deny any influence of the L1, claiming that the development of functional categories in L2 learners will be the same regardless of the learner's L1.

Also contradicting the Full Transfer/Full Access theory, Pienemann's (Pienemann 1998, 2015; Pienemann & Keßler, 2011) Processability Theory claims that L2 learners can only acquire linguistic forms which they can process. Each learner, regardless of her L1, follows the same acquisition stages and these stages act as building blocks in that they depend on each other. The top, highest stage cannot be reached without each level below it. Because this hierarchy is applicable to all languages, Processability Theory claims a Lexical-Functional Grammar as the starting point of L2 acquisition. Learners begin with the lexical representation before its grammatical category can be assigned. Processability Theory does not identify a role of Universal Grammar in that it claims that learners are constrained not by UG but by the hierarchy and their processing capacity. The only innate linguistic knowledge a learner relies on is the basic notion of constituency and the one-to-one mapping of semantic roles (Pienemann, 2015). The rest is developed from the input.

Although, according to Processability Theory, all learners move through the same predictable stages for L2 acquisition, learner variation is still apparent. Processability Theory accounts for learner variation claiming that it arises from the choices an individual learner makes when solving developmental problems and the way the learner shapes her interlanguage. The Developmentally Moderated Transfer Hypothesis (DMTH) argues that L1 transfer is restricted by processability. Learners can only transfer from their L1 what can be processed. Therefore, L1 transfer only occurs when the learner is developmentally ready to acquire such features in the L2. Håkansson and Pienemann (2002) found that L1 Swedish learners of German did not transfer word order to the L2 despite the structure being identical in both languages. Instead, the participants produced a word order that is ungrammatical in the L1 and the L2. They claimed the data supports the DMTH because the forms higher up in the processability hierarchy were not transferred initially even if they are identical in both

languages. This completely contradicts Full Transfer/Full Access that states that the L1 grammar is the initial state of the L2.

Finally, Full Access theory (Epstein et al., 1996) claims that “UG in its entirety constrains L2 acquisition” (p. 677). The L2 learner has full access to the UG principles and parameters, and there is no L1 transfer on L2 acquisition.

Having discussed some of the current theories regarding L2 acquisition and the extent of UG availability and L1 transfer, the following sections explore the structures of interrogatives in both English and Italian, both of which are important to understand for the present study.

1.2 English Questions

English is an SVO language with a strict word order due to a weak inflectional system. The position of the subject is not flexible. This strict word order is apparent in the structures of English interrogatives. This thesis will focus mainly on the role of subject-auxiliary inversion in English main and embedded object questions with a particular focus on *wh*-questions.

1.2.1 English main questions

There are two types of main clause, or direct, object questions in English: *yes/no* and *wh*-. Both types require subject-auxiliary inversion (although non-inverted *yes/no* questions can occur in the input). The subject must follow the auxiliary and precede main verb. The *yes/no* question in (1a) shows this subject-auxiliary inversion. This phenomenon is also known as T-to-C movement because the tensed auxiliary moves into the head position of the Complementizer Phrase (CP). In English, only auxiliary verbs are able to move and therefore, in declarative sentences with a lexical verb (1b), *do*-support is needed in order to turn the declarative sentence into an interrogative (1c). The *do* auxiliary must precede the subject and it receives inflection for tense and number while the main verb *work* is untensed.

- (1) a. Have you eaten?
- b. She works today.
- c. Does she work today?

The present study focuses on inversion in *wh*-questions. In English *wh*-object questions, the *wh*-element is fronted to the beginning of the interrogative and appears before the auxiliary. The obligatory word order is shown as WH + AUX + SUBJ + V. Examples (2a*) and (2b*)

demonstrate the ungrammatical outcome when subject-auxiliary inversion is not applied. Also, it's worth noting that *how come* is the only English *wh*-phrase that does not permit inversion (3).

- (2) a. What is she doing?
*What she is doing?
- b. Why have you done that?
*Why you have done that?
- (3) How come he left the party?
*How come did he leave the party?

1.2.2 English embedded questions

English embedded clause questions, or indirect questions, show an asymmetry with main questions in that they do not allow for subject-auxiliary inversion. Take the examples (4) below in which the structure of a main question applied in an embedded context becomes ungrammatical (4b). The embedded clause in (4c) shows no inversion.

- (4) a. What is she doing?
b. *Her mom wants to know what is she doing.
c. Her mom wants to know what she is doing.

Embedded questions in English only apply *wh*- movement and do not trigger T-to-C movement. Therefore, the insertion of *do*-support in embedded contexts with lexical verbs is not necessary (5b).

- (5) a. *I wonder where does he live.
b. I wonder where he lives.

The purpose of this study is to investigate how Italian learners of English acquire the complex structures of English questions. The study examines errors made by Italian learners in the production of English main and embedded questions. In order to analyze these errors, it is necessary to understand the structure of interrogatives in Italian.

1.3 Italian Questions

Italian is an SVO language that allows for a flexible word order due to a strong inflectional system. Examples (6a) and (6b) show that the subject position is flexible. Additionally, like other Romance languages, Italian is a pro-drop language and therefore allows for the subject to be omitted completely (6c).

- (6) a. Alessia lavora qui.
b. Lavora qui Alessia.
c. Lavora qui.

1.3.1 Italian main questions

In Italian main questions, subject-verb inversion is mandatory and triggers T-to-C movement. Unlike in English main questions, when there is an auxiliary and subject present, the subject cannot occur between the auxiliary and the main verb (7*) and (8*).

- (7) *Che cosa ha Gianni fatto?
Che cosa ha fatto Gianni?
What has done Gianni?
- (8) *Dove sono Martina e Sofia andate?
Dove sono andate Martina e Sofia?
Where have gone Martina and Sofia?

Additionally, the adjacency requirement of Italian questions with a bare *wh*-element requires the verb to follow the *wh*-phrase, not permitting the separation of the two (Calabrese, 1982; Rizzi, 1996, 2001, as cited in Bianchi, Bocci & Cruschina, 2017, p. 234). Examples (9–11) show the arguments *what* (*che cosa*) and *who* (*chi*) and the adjunct *where* (*dove*) trigger obligatory T-to-C movement. The ungrammatical questions are a result of the subject occurring between the *wh*-word and the inflected verb.

- (9) a. *Che cosa Gianni ha fatto?
What Gianni has done?

b. Che cosa ha fatto Gianni?

What has done Gianni

(10) a. *Chi Gianni ha incontrato?

Who Gianni has met?

b. Chi ha incontrato Gianni?

Who has met Gianni?

(10b) can be interpreted as the subject question 'Who met Gianni?' or the object 'Who did Gianni meet?'

(11) a. *Dove Gianni è andato?

Where Gianni has gone?

b. Dove è andato Gianni?

Where has gone Gianni?

The case of 'perché'

Rizzi (2001) shows that subject-verb inversion is not mandatory in *why* (*perché*)-questions in Italian and that *perché* does not trigger obligatory T-to-C movement. The subject in *perché*-questions is able to appear in a preverbal (12, 14) or a postverbal (13, 15) position. While *wh*-elements that correspond to arguments or lower adverbials (e.g., *what*, *where*, *how*) require inversion, *perché* and *come mai* (*how come*) behave like higher adverbials and do not require inversion (Rizzi, 2001). These special adverbials "fill a position distinct from and higher than the position of ordinary *wh*-elements in main clauses" (p. 7).

(12) Perché Gianni è venuto?

Why Gianni has come?

(13) Perché è venuto Gianni?

Why has come Gianni?

(14) Come mai Gianni è partito?

How come Gianni has left?

- (15) Come mai è partito Gianni?
How come has left Gianni?

A study done by Bianchi, Bocci and Cruschina (2017) investigated the flexible subject position in *perché*-questions and found that 62% of the L1 Italian participants preferred a preverbal subject position in *perché*-questions in neutral contexts. However, when the context was not neutral but focused on emphasizing the subject, the participants favored postverbal subjects (66.5%) over preverbal. In the above examples (12) and (13), the latter structure seems to be preferred by speakers when emphasizing that it was Gianni who came and not someone else. Similar results were found in a study by Bocci and Pozzan (2014). They found that when the subject is able to appear in either the preverbal or postverbal position, the case with *perché*-questions, it was shown that speakers tend to choose preverbal subjects in neutral contexts. In *perché* contexts that license a narrow focus interpretation, postverbal subjects are often preferred over preverbal. Taken together, these results support the hypothesis that in neutral contexts, Italian speakers do not apply subject-verb inversion in *perché*-main questions.

1.3.2 Italian embedded questions

Unlike English, there is no asymmetry between Italian main and embedded questions. The structure of embedded questions is the same as main questions, allowing for subject-verb inversion (Guasti, 1996). Compare (16a) to (16b) below. Both show the same structure in the main question in (16a) and the embedded clause in (16b).

- (16) a. Che cosa ha fatto Gianni?
What has done Gianni?
- b. Elena vuole sapere che cosa ha fatto Gianni.
Elena wants to know what has done Gianni.

The special case of *perché*-main questions holds in embedded contexts in that the subject position is flexible and possible either preverbally or post verbally. The previously mentioned study by Bocci and Pozzan (2014) found that, in addition to main questions, monolingual Italian speakers preferred a preverbal subject in neutral embedded contexts when this preverbal position was available. The authors included both questions containing *dove* (*where*) and *perché* (*why*) and found significant contrast between the two, with postverbal

subjects strongly preferred in *dove* questions and preverbal subjects strongly preferred in *perché*-embedded questions. In the examples below, (17a) was shown to be preferred over (17b) when there was no narrow focus interpretation.

- (17) a. Non so perché Marta balla.
b. Non so perché balla Marta.
I don't know why Marta dances.

The previous sections have summarized the interrogative structures in both English and Italian. In order to understand potential difficulties in the acquisition of English questions by Italian learners, it is necessary to directly compare the two languages.

1.4 English and Italian comparison

The previous sections have shown that in English there is an asymmetry between main and embedded questions, and this asymmetry does not occur in Italian. Subject-auxiliary inversion is required in English main questions but not allowed in English embedded questions. In Italian, subject-verb inversion is required in both main and embedded *wh*-questions, except in those that include *perché* and *come mai*.

Subject inversion, or T-to-C movement, is triggered in both English and Italian main questions. However, the element that raises is different in each language. English allows for only the auxiliary verb to raise to T while Italian does not allow for the auxiliary verb to move on its own. Therefore, while both Italian and English require T-to-C movement, the element moving from T-to-C is different, resulting in different word orders (Pozzan, 2011). Compare the word order of both English and Italian main *wh*-questions in Table 1. The main distinction is that in no case in Italian can the subject intervene between the auxiliary and the verb, while this movement is required in English. Table 1 shows the auxiliary verb *has eaten* and as mentioned before, when the main verb in an English declarative sentence is a lexical verb, it is obligatory to insert the auxiliary *do* into the question structure. Additionally, Table 1 shows the asymmetry between English main and embedded *wh*-questions that does not appear in Italian.

Table 1. *Word order in English and Italian wh-questions*

	Question type	Word order	Example
English	Main	WH+AUX+SUBJ+V	What has John eaten?
	Embedded	(main clause)+WH+SUBJ+AUX+V	I don't know what John has eaten.
Italian	Main	WH+AUX+V+SUBJ	Che cosa ha mangiato Gianni?
	Embedded	(main clause)+WH+AUX+V+SUBJ	Non so che cosa ha mangiato Gianni.

1.4.1 Comparing English 'whose' to Italian 'di chi'

The examples of English questions shown in previous sections include a bare *wh*-element. The English genitive *whose* has lexical restrictions in that the *wh*-word cannot move on its own. The noun phrase in which *whose* appears must be pied-piped to the specifier of the complementizer phrase (Gavruseva & Thornton, 2001). *Whose* cannot be subextracted from the NP. Take example (18b) in which *whose* has been moved out of the entire NP *whose car* resulting in an ungrammatical question. (18a) shows that *whose* main questions in English still undergo subject-auxiliary inversion and *wh*-fronting, but the entire NP must be fronted with the *wh*-element *whose*.

- (18) a. Whose car did Joe drive?
 b. *Whose did Joe drive car?

L1 English-speaking children have shown difficulties in the acquisition of *whose*-questions. Gavruseva and Thornton (2001) found that L1 English-speaking children split *whose*-questions and strand the NP in the embedded clause. This is shown in example (19), taken from Gavruseva and Thornton (2001), in which the child has stranded the possessive 's. Additionally, the study found that the children produced structures with a double genitive as in (20) and subextracted the possessive *whose* as in (21).

- (19) *Who do you think's flower fell off?
 (20) *Whose do you think's hat is on the skis?
 (21) *Whose do you think ball went in the cage?

Gavruseva (1997) analyzed data from two L1 Russian child learners of English. The children in the study were aged 6;5 and 5;11. Compared to the data with monolingual English-speaking children, the L1 Russian children also produced split *whose*-questions but omitted the possessive 's morpheme as in (22) below.

- (22) *Who do you think Pocahontas likes the chair?
(Target: Whose chair do you think Pocahontas likes?)

The acquisition of the possessive *whose* poses challenges in L1 acquisition as well as in L2 acquisition. In Italian, the English possessive *whose* is represented as the prepositional phrase *di chi* (*of whom*). In declarative sentences, the preposition *di* is used to show possession as in (23). Comparing (23) with (24a) one can see that in the question in (24a) the verb *è* moves in front of the subject *la macchina* and triggers subject-verb inversion. The difference is that in Italian, the NP acts as the subject of the main *di chi* clause and the verb *essere* shows agreement features with the NP. (24b) shows the verb agreement with a plural NP *le scarpe*. Furthermore, in Italian *di chi* questions the complementizer *che* (*that*) introduces the relative clause. Similar to English, the possessive phrase *di chi* cannot be subextracted from the NP as in (25).

- (23) La macchina è di Luca.
The car is of Luca.
The car is Luca's.
- (24) a. Di chi è la macchina che ha guidato Gianni?
Of whom is the car that drove Gianni?
Whose car did Gianni drive?
- b. Di chi sono le scarpe che indossa Marta?
Of whom are the shoes that wears Marta?
Whose shoes does Marta wear?
- (25) a. *Di chi ha guidato Gianni la macchina?
b. *Di chi indossa Marta le scarpe?

The interrogative structures of the possessive English *whose* and Italian *di chi* are different in that in English the NP must be pied-piped to the possessive *whose* and following the NP subject-auxiliary inversion is required. In Italian, the NP is the subject of the main clause and *che* introduces the relative clause. The general structures are shown below in Table 2. As one can see, it is difficult to compare the two because they are very different. For this

reason, I expect to see difficulty in the acquisition of English *whose*-questions by L1 Italian speakers.

Table 2. *English 'whose' and Italian 'di chi' comparison*

English <i>whose</i>	Italian <i>di chi</i>
WH + NP + AUX + SUBJ + V	PP + V + SUBJ + CHE + relative clause
Whose keys did she find?	Di chi sono le chiavi che ha trovato lei?

1.5 L2 Acquisition of English questions

The complex structure of questions in English poses difficulties in acquisition, both as an L1 and as an L2. As mentioned, in English main questions there is an obligatory inversion of the subject and auxiliary. This inversion does not take place in embedded clause questions. The problem of acquiring English questions is knowing when to apply subject-auxiliary inversion. Within L1 acquisition, studies have shown that children commonly make word-order errors in the production of questions and that these errors seem to be inconsistent, with children producing inversion errors in certain types of questions while producing well-formed structures in others (Ambridge et al., 2006). Pozzan and Valian (2017) found that English-speaking children more often correctly invert in main clause *yes/no* questions than in main clause *wh*-questions. Furthermore, they found a positive correlation between correct inversion in main clause questions and incorrect inversion in embedded questions. This suggests that children generalize subject-auxiliary inversion from main clause contexts to embedded. Stromswold (1990) studied monolingual English-speaking children's acquisition of the complex auxiliary system. Most of the errors made, however, did not involve auxiliaries but were errors of *do*-support and inversion. Children quickly master the auxiliary system but have problems with *do*-support, which is unique to English. Learning when and what to invert in English questions is challenging for both L1 and L2 learners. Within L2 acquisition, errors of inversion persist even in highly proficient learners.

Pienemann (1998; Pienemann & Keßler, 2011) presents six stages that L2 learners move through when acquiring English questions (Table 3). According to Pienemann and the Processability Theory, learners move through generalizable developmental patterns when acquiring an L2, regardless of the learner's L1 and L2 learning situation. Therefore, Pienemann claims that all learners of English, regardless of their L1, will progress through the six stages represented in Table 3 when acquiring English questions.

Table 3. *Order of acquisition of questions in L2 English (Pienemann 1998; Pienemann & Keßler, 2011)*

Stage	Structure	Grammatical	Example
6	Cancel Inversion	Yes	She asked why Dan has left.
		No	She asked why has Dan left.
5	Aux 2nd	Yes	What are you reading?
		No	What you are reading?
4	Yes/no inversion	Yes	Has she called?
		No	She has called?
	Copula inversion	Yes	Is he late?
		No	He is late?
3	Wh-fronting	No	Why he is late?
2	SVO?	Possibly	You go home?
		Possibly	He is coming?
1	Single word question	Possibly	My money? (= is this...?)

The first stage is the production of single word questions with rising intonation. For example, “Mine?”. These questions can be multiple words but still considered single lexical entries, for example, “This one?”. Learners then move on to produce sentences with an SVO word order in Stage 2. The rising intonation is the only indication that the produced utterance is a question, as the learner is still unable to produce inverted structures at this stage. A learner in Stage 2 would produce a question like “He is home?”. Stage 3 is characterized by the fronting of the *wh*-word at the beginning of the utterance still in SVO order. For example, “Why he is home?”. Stage 4 is when the learner begins to invert in *yes/no* questions, moving the position of the verb in front of the subject in “Is he home?”. Stage 5 is characterized by the learner producing subject-auxiliary inversion in *wh*-questions such as, “Why is he home?”. The final stage in the L2 acquisition of English questions is the distinction of main and embedded questions. That is, the ability to apply cancel inversion to embedded questions such as, “He asked where she is.” Based on Pienemann’s stages of acquisition, it is clear that utilizing cancel inversion in the production of embedded questions is an advanced structure and difficult even for highly proficient learners of English.

Processability Theory claims that, while all learners go through the same developmental patterns, learner variation arises in the choices the learner makes when having to solve these developmental problems (Pienemann & Keßler, 2011). For example, during Stage 5 before learners have completely acquired the auxiliary second structure in *wh*-questions, they can form

wh-questions in two ways: either omitting the element that requires inversion (Where you lost it?) or utilizing the word order that they've acquired from Stage 2 (Where you have lost it?). These errors, then, would not be due to L1 interference but to the fact that learners can only process what they have already acquired.

Spada and Lightbown (1999) provide a similar account of developmental stages in English questions. They've based these stages on their research studying the role of classroom instruction in the acquisition of English questions. Stage 1 is characterized by the use of single words or fragments. In Stage 2 learners produce SVO utterances with a rising intonation. Stage 3 is when learners begin *do*-fronting and *wh*-fronting. During Stage 4 learners begin inverting in *wh*-questions with the auxiliary *be* and inverting in *yes/no* questions. Finally, Stage 5 learners produce *wh*-questions with the auxiliary in second position.

According to The Council of Europe's Common European Framework of Reference (CEFR) for English learners, learners at the A2 level are able to produce simple direct *wh*-questions. B1 level learners begin to produce more complex sentences. It is at this level that learners are able to produce indirect questions such as, "Guess where it is." While these structures are introduced to students at the beginner/intermediate levels, it does not mean that at these levels students master a given structure (e.g., *cancel inversion*). Referencing the English Profile compiled by researchers at Cambridge ESOL and Cambridge University Press (CUP), B2 learners should have acquired and should be able to produce main and embedded *wh*-question structures (Cambridge ESOL & CUP, 2011). However, it has been shown that even advanced learners of English continue to produce errors within these structures.

Other factors to keep in mind regarding L2 acquisition of English questions are the effects, if any, of the learner's L1 and different *wh*-words.

1.5.1 Effect of L1 transfer on L2 acquisition

Some studies have investigated the role of the L1 in a learner's acquisition of English questions. Pienemann (1998) and Pienemann and Keßler (2011) claim that the stages of acquisition presented above are independent of the L1. All English learners move through the stages regardless of their L1. However, the question of L1 influence on acquisition of L2 English questions still remains. Some studies have found an effect while others contribute the errors to factors independent of the learner's L1, for example, interference from the L2. Spada and Lightbown (1999) found an interaction between the developmental stages listed above and the learner's L1. They studied French speaking children learning English and found that, without explicit instruction, the children continued to make assumptions about the structure of English

questions based on the structure of French questions, incorrectly applying these L1 structures to English.

Pozzan (2011) and Pozzan and Quirk (2014) also studied the effect of the L1 on the production of English questions. The study included two elicited oral production experiments with L1 Spanish and L1 Chinese participants with intermediate/advanced levels of English. Experiment 1 elicited oral production of main *yes/no* and *wh*- English questions and Experiment 2 elicited oral production of embedded *yes/no* and *wh*- English questions. It was hypothesized that, if L1 transfer is not a factor, there will be no difference in errors produced by the two groups. If, however, L1 transfer influences production in intermediate/advanced learners, it was hypothesized that the L1 Spanish group would produce fewer errors and higher inversion rates in Experiment 1, due to the nature of the syntactic structure of Spanish questions in which the subject must follow the verb. Additionally, the L1 Chinese group was expected to produce fewer errors and higher non-inversion rates in Experiment 2, as Chinese is similar to English in that it does not allow for subject-verb inversion in embedded contexts.

The results of the study showed a significant effect of L1 in Experiment 1. L1 Chinese speakers produced higher rates of correct and target-like inverted responses than the L1 Spanish participants. However, no effect of L1 was found in Experiment 2, with both groups producing comparable rates of target-like non-inversion. The results of Experiment 2 are consistent with theories that claim no L1 transfer on the L2 and those that claim L1 transfer only in earlier stages of acquisition. The L1 effect results found in Experiment 1 are unexpected. Spanish learners were hypothesized to produce more target-like, inverted responses in Experiment 1, due to the syntactic structure of Spanish questions. Chinese allows for no inversion in either question type, and yet the L1 Chinese learners produced more correct, inverted responses in Experiment 1 and comparable rates of correct, non-inverted responses in Experiment 2. Several L2 acquisition theories claim that similarities between the L1 and the L2 (i.e., inversion in Spanish and English) interfere with, rather than facilitate, the acquisition of the L2 structure (Oller & Ziahosseiny, 1970; Ringbom, 1987). Similarly, Kleinmann (1977) claimed that an L2 structure that is completely different from that of the L1 (i.e., inversion in English but not Chinese) is easier to acquire due to its saliency. This could explain why the Chinese group showed higher rates of target-like inversion in Experiment 1. The authors hypothesize that in Experiment 2, both groups of participants overgeneralized the inversion rules from main questions to embedded questions and that there was an interference of the L2 not transfer of the L1.

In addition to eliciting oral production of English questions, Pozzan (2011) examined written production through the International Corpus of Learner English (ICLE). The corpus is made up of 6,085 texts written in English by undergraduate students of varying L1s who had all learned English in a non-English-speaking country. 16 different L1s were represented. A sample of the texts was rated for writing proficiency and 60% of the texts were rated as advanced. Regarding L1 transfer in *wh*-main questions, she found there to be a significant increase in the odds of an inverted response when the learner's L1 required inversion. Learners from an L1 that requires inversion produced more inverted questions than those whose L1 does not allow for inversion. When inversion was possible but not obligatory in the L1, she found a decrease in the odds that the learner would produce an inverted structure. These results support the theory that similarities (e.g., obligatory inversion and optional inversion) between the L1 and L2 make parameter resetting more difficult, interfering with, instead of facilitating, the acquisition of the L2 structure (Oller & Ziahosseiny, 1970; Ringbom, 1987). However, within embedded questions, Pozzan found no significant effect of L1 transfer on accuracy and non-inverted responses. Learners from L1s where inversion is possible showed slightly higher inversion rates, but this did not prove to be a significant predictor on accuracy or non-inversion rates.

Similarly, Gonzalez Fernandez (2015) found no effects of L1 transfer on L2 production of English embedded questions. L1 Spanish speakers of varying English proficiencies participated in various tasks in order to prompt them to produce embedded questions. The goal of the study was to find whether the source of ungrammatical subject-auxiliary inversion in embedded questions was due to transfer from the L1 or interference of the L2. The results showed no transfer of Spanish VP movement to English embedded contexts. However, ungrammatical subject-auxiliary inversion was found. The conclusion made was that the errors of inversion in embedded questions are due to interference from the L2, and incorrectly generalizing subject-auxiliary inversion from main to embedded contexts.

Looking at L2 comprehension of *wh*-questions, Hopp (2017) explored L2 processing of English subject and object *which*-questions, comparing the results with monolingual child learners and monolingual adults. The goal was to assess whether non-target-like performance is due to L1 transfer and/or proficiency. Hopp found effects of L1 transfer on L2 comprehension in intermediate L2 learners. Additionally, the results differed for each of the proficiency groups studied (intermediate, high-intermediate, advanced) with target-like comprehension increasing with proficiency. The results are compatible with L2 acquisition theories that claim the L1 grammar is the initial state.

Data from previous studies has supported L1 transfer theories that claim no transfer or transfer only initially in the beginning stages of acquisition. Additionally, it may be that learners have more difficulty acquiring the structure of English questions when inversion is possible but not obligatory in their L1 compared to learners whose L1 does not allow for inversion in any case. Regarding embedded questions, Gonzalez Fernandez (2015), Pozzan (2011) and Pozzan and Quirk (2014) attributed errors of cancel inversion to interference from the L2 and not the L1.

1.5.2 Effect of *wh*-word

Another factor to take into account when studying the acquisition of English questions is *wh*-word. Studies have shown that in both L1 and L2 acquisition, learners acquire *when*, *which*, and *whose*-question structures later than *what*, *where* and *who* (see Bloom et al., 1982; Park, 2000; Pozzan, 2011; Pozzan & Quirk, 2014; Spada & Lightbown, 1999; Stromswold, 1990). *Why* has also been documented as being acquired later and showing lower rates of inversion in L2 production. Pozzan (2011) and Pozzan and Quirk (2014) found both groups of Chinese and Spanish speakers to produce noticeably low inversion rates in main *why*-questions. The finding also appears in embedded contexts as *why* was associated with higher rates of target-like, non-inverted responses. Within the written production analysis of the ICLE, Pozzan (2011) found *why*-questions were associated with the lowest rates of inversion in main questions and there was a significant positive effect of producing an inverted response in main questions when the *wh*-word was *how* or *what*. Additionally, Stromswold (1990) and Pozzan and Valian (2017) showed L1 English-speaking children invert more frequently in main argument *wh*-questions than in adjuncts.

Lee (2008) also reported on the argument-adjunct asymmetry in L2 acquisition. She studied L1 Korean learners of English and, through a grammaticality judgment task, found that the learners made more correct judgments in non-inverted main argument *wh*-questions than in main adjunct *wh*-questions. The learners were more likely to judge the non-inverted questions as ungrammatical in *what* and *who* questions than in *why* and *how* questions. Non-inverted adjunct *wh*-questions were less frequently noted as ungrammatical. Additionally, the study investigated if the argument-adjunct asymmetry can be attributed to differences in the input (i.e., learners are exposed to arguments more frequently than adjuncts) or to differences in the syntactic structure of the two *wh*-types. Analyzing different English textbooks and scripts from American TV shows and movies, she found no asymmetry in the frequency of arguments and adjuncts. This supports the hypothesis that the argument-adjunct asymmetry doesn't come

from differences in the input but rather from differences in the syntactic properties of the two *wh*-types.

Park (2000) studied Korean children's acquisition of English *when*-questions. Citing Bloom et al. (1982)'s sequence of acquisition for L1 English questions, Park (2000) notes that *when* is the last *wh*-word to be acquired by L1 English children. This has been attributed to cognitive development that later L2 learners should not be affected by. However, Park (2000) found that the later acquisition of *when* holds in L2 learners in his study of L1 Korean children.

Thornton (2008) reports on an English-speaking child's production of *why*-questions over the course of four years, from age 2 to age 6. The child was found to lack subject-auxiliary inversion in main *why*-questions compared to other *wh*-questions. Thornton compares the child's early grammar to that of the Italian *perché* mentioned in section 1.3.1, hypothesizing that children may initially treat *why* similarly, merging it into the higher phrase structure SpecIntP rather than moving it to the SpecFocP. While Italian allows for *perché* to be merged, English does not, resulting in ungrammatical productions in early child grammar.

To summarize, past studies have shown that arguments are acquired and mastered earlier than adjuncts in both L1 and L2 acquisition. Furthermore, arguments have been associated with higher rates of inversion. *When*, *which* and *whose* are the last *wh*-question words to be acquired by L1 speaking children (Bloom et al., 1982) and *why*-questions are associated with low inversion rates in both L1 and L2 acquisition.

2. CURRENT STUDY

The goal of the present study is to adapt the Pozzan (2011) and Pozzan and Quirk (2014) elicited oral production study and explore how L1 Italian learners of English produce English *wh*-questions, both main and embedded. Keeping in mind the literature on L2 acquisition of English questions, the study aims to analyze production errors, specifically focusing on errors of inversion, attempting to identify when these structures are mastered by L2 English learners. The experiments are designed to elicit written production of both main and embedded *wh*-questions. Past studies have shown that L2 learners produce more inversion errors in *wh*-questions than in *yes/no* questions (Pozzan, 2011; Pozzan & Quirk, 2014). The present study focuses only on the production of *wh*-questions. A number of researchers have proposed stages of L2 acquisition of English questions independent of the learner's L1 (see Pienemann, 1998; Pienemann & Keßler, 2011; Spada & Lightbown, 1999). These stages claim that learners acquire the structure of *yes/no* questions before that of *wh*-questions and main questions before embedded questions. Keeping this in mind, the present study aims to answer the following questions:

1. *What types of errors do L2 learners of English make in the production of English wh-questions?*
2. *Do participants produce more correct responses in main questions than in embedded questions?*
3. *Is there an effect of wh-word on inversion rates?*
4. *Is there an effect of verb type on inversion rates?*
5. *Are non-inversion errors in embedded questions due to L1 transfer or L2 interference and overgeneralization from main to embedded contexts?*

2.1 Experiment 1: Main questions

The goal of Experiment 1 is to investigate how L1 Italian learners of English produce English main *wh*-questions and the errors they make, with a specific focus on subject-auxiliary inversion errors. Pozzan (2011) and Pozzan and Quirk (2014) investigated subject-auxiliary inversion errors in English main questions by adult learners of different L1s. They examined the elicited oral production of L1 Spanish and L1 Chinese learners of English with intermediate/advanced proficiency, looking at L1 as a possible factor influencing inversion rates. In addition to analyzing the L1, they looked at how question type (*wh*- vs. *yes/no*) and *wh*-word (*who*, *what*, *why*, *where*) affected the participants' inversion rates.

As a follow up study to Pozzan (2011) and Pozzan and Quirk (2014), the present study examines L1 Italian speakers with intermediate/advanced levels of English and their written production of *wh*-questions. Experiment 1 investigates English main *wh*-questions, also looking at the effect of *wh*-word on inversion rates. The present study includes the same four question words from the previous study (*who*, *what*, *why*, *where*) and adds the genitive *whose* as a fifth *wh*-word.

Whose was added as a fifth *wh*-word in order to investigate how Italian L1 learners acquire and master this structure, as there is no equivalent structure in the L1. Additionally, there have been few studies on how *whose* is acquired in L2 English. Due to the unique nature of *whose*, I expect to see participants having more difficulty producing correct *whose* main questions. Additionally, as past studies have shown (Pozzan, 2011; Pozzan & Quirk, 2014; Spada & Lightbown, 1999), I expect to see lower inversion rates associated with *why*. Although *why* has been shown to be associated with lower rates of inversion across L1s, Italian *why*-questions do not require obligatory subject-verb inversion. This could have an influence on Italian participants' inversion in English *why*-questions although, as only one L1 is included in the present study, no definitive conclusions regarding L1 transfer can be made.

2.1.1 Method

2.1.1.1 Participants

The goal of this study is to examine the production of main questions by L1 Italian intermediate/advanced learners of English. All participants were native speakers of Italian with intermediate or advanced levels of English. All participants were currently living in Italy, with the exception of four participants who were living in another non-English dominant speaking country (Switzerland, Spain, Germany, and France). Participants were recruited online via Facebook and sent a link to the Participant Contact Form in order to sign up to be included in the subject pool. After signing up, participant demographic and language background information was collected through the Language Profile Questionnaire implemented through Qualtrics. The full Language Profile Questionnaire can be found in Appendix F. Participants were sent both English and Italian versions of the questionnaire. During the questionnaire, participants were asked to self-rate their English proficiency level using the Common European Framework of Reference (CEFR) scale. The majority of participants (15) rated their English level a C1 on the CEFR scale. Two participants said they were a C2 level and five participants rated their level as B2. Additionally, participants were asked to rate how well they speak, read, write and understand English on a scale of 1 (not well at all) to 7 (extremely well). The mean

of these four scores was calculated to reflect the participants' mean English ability score. All participants had some level of higher education and the majority of participants (15) had completed or were in the process of completing a university master's degree. Two participants had completed or were completing a PhD and five participants had completed or were completing a bachelor's degree. Participants ranged greatly in how often they are exposed to and use English in an average week.

Four participants were excluded because they failed to complete all parts of the study. Data from 22 participants was analyzed. Table 4 shows participant demographic data.

Table 4. *Data from 22 L1 Italian participants (self-rate CEFR levels: B2 = 5, C1 = 15, C2 = 2)*

	Mean	Range	Standard deviation
Age	28	22 - 42	5.3
English ability rated 1 - 7	5.7	3 - 6.5	0.7
Weekly English exposure in hours	57.8	10 - 168	38.9
Weekly English use in hours	37.8	3 - 120	33.8
Age when started to learn English	6.4	2 - 9	1.9
Age when felt comfortable with English	17.5	11 - 29	4.7
Years of English courses taken	16.9	9 - 28	3.8

2.1.1.2 Materials

Experiment 1 was an elicited written production computerized task created through PsychoPy3 software and implemented online through Pavlovia. The experimental design was modeled after that of Pozzan (2011) and Pozzan and Quirk (2014) and adapted to elicit written production instead of oral production. Participants were introduced to Phil, a shy student who needed help asking his teacher, Ms. Smith, some questions. Participants were instructed to help Phil ask Ms. Smith questions. The first screen showed Phil and the experimental prompt in a speech bubble. In order to elicit *wh*-questions, participants were shown a prompt such as "Mark is complaining to the waiter. Ask Ms. Smith why." After being instructed to read the prompt aloud, participants were then told to advance to the next screen which showed an image of Ms. Smith in front of a blackboard. They were asked to type in the question to Ms. Smith. Participants were not able to return to the previous screen to see the prompt, ensuring they were not reading the prompt in order to produce the question. In the final screen, Ms. Smith answered, "I don't know." Photos of the experiment screens can be found in Appendix A.

The experiment included a demo in order to ensure that the task and the instructions were clear to the participants. Because the experimenter could not be present to answer any

questions or provide clarification, it was necessary to include a demo. The demo prompt, intended to elicit a *yes/no* question, was “Maybe Sue works today. Ask Ms. Smith.” The demo was timed, and on the second screen participants were shown the question “Does Sue work today?” as it was typed out. Photos of the demo screens can be found in Appendix A. After the demo, the participants reached the practice items. The practice items were the same used in Pozzan (2011) and Pozzan and Quirk (2014) and included six prompts, three intended to elicit subject *wh*-questions and three intended to elicit *yes/no* questions. A list of the practice items can be found in Appendix B.

Each participant was randomly assigned to either List 1 or List 2 and shown a total of 20 experimental items. The items were manipulated to include prompts with or without an auxiliary verb, with the lexical verbs requiring *do*-support in the produced question. The fully-within factor was the presence of an auxiliary. The two auxiliaries utilized were *be* and *have*. *Wh*-question word was also manipulated. *Wh*-word was a within-subjects but between-items factor. The five *wh*-words included were the adjuncts *why* and *where* and the arguments *who*, *what*, and *whose*. Half of the verbs were in past tense and the other half appeared in the present tense. The outline below shows the breakdown of items included in Experiment 1 and a full list is found in Appendix B.

20 *wh*- items

- 10 lexical verbs
 - 2 *why*, 2 *who*, 2 *where*, 2 *what*, 2 *whose*
- 10 auxiliary verbs
 - 5 auxiliary *be*
 - 1 *why*, 1 *who*, 1 *where*, 1 *what*, 1 *whose*
 - 5 auxiliary *have*
 - 1 *why*, 1 *who*, 1 *where*, 1 *what*, 1 *whose*

The order participants were shown the items was randomized in order to control for order effects. The same verbs appeared with the same *wh*-word across both lists. However, if a verb was used as a lexical verb in List 1, it appeared with an auxiliary in List 2.

2.1.1.3 Procedure

After filling out the Participant Contact Form, participants were sent an email with a link to the Language Background Questionnaire administered through Qualtrics. At least 24 hours after

completing the Language Background Questionnaire, participants were sent another email with a link to the Pavlovia online experiment. After reading the consent forms and giving their consent, participants were given detailed instructions and shown a demo before reaching the practice items. The demo was necessary to ensure that, because of the online nature of the experiment, participants understood what was being asked of them and the instructions were clear. After completing Experiment 1, participants were automatically redirected to Experiment 2. Experiment 1 took approximately 15 minutes to complete.

2.1.1.4 Transcription and coding

Two coding schemes were used to code the participants' responses – strict coding and lax coding. Within the strict coding scheme, responses were coded as either correct (native-like) or incorrect (non-native-like) in terms of word order, verb type (i.e., an auxiliary verb in the prompt was changed to a lexical verb in the response or vice versa), verb morphology, question type (i.e., a subject *wh*- or *yes/no* question instead of an object *wh*-) and target-like subject and objects. In order to be coded as correct, the semantics and structure of the response had to be the same as the target. Semantic substitutions, as in (26), were coded as correct.

(26) Why did your husband call the shop? (Target: Why did your husband call the store?)

Incorrect responses were coded into five further categories: non-inversion (27) and raising (28) errors, double tense or double auxiliary errors (29), lack of morphology errors (30), non-target auxiliary errors (31) and other. The category of other included non-target *wh*-word errors (32), omitted or incorrect prepositions, phrases, objects or subjects, production of a subject *wh*-question instead of an object *wh*-question (33), a *yes/no* question or an embedded question (34), VP movement (35), aborted or skipped trials, incorrect morphology, non-target-like lexical items, incorrectly adding a determiner after *whose* (36), and producing a *whose* non-target marked structure (37).

(27) Where sara has invited them?

(28) Why Mark complains about his job?

(29) Where does your sister usually travels to in June?

(30) Why Ms. smith husband call the store?

(31) What did Tom teach his son to do in the house? (Target: What has Tom taught his son to do in the house?)

(32) Who is marc complaining with the server? (Target: Why is Mark complaining to the server?)

(33) Who met Laura in France? (Target: Who did Laura meet in France?)

(34) Do you know why the boss had closed the office door?

(35) Whose baseball hat is wearing Jan?

(36) Whose the phone has marta found in her car?

(37) Whose is the shirt that Anna cleaned?

Within the lax coding scheme, responses were coded as either correct or incorrect in terms of word order, question type, and verb morphology. Responses that were grammatical and maintained a target-like meaning but were not target-like in terms of verb type (i.e., changing an auxiliary verb in the prompt to lexical verb in the response or vice versa), subject (38), and lexical items used were coded as correct. Responses that were missing phrases or objects but otherwise target-like were coded as correct (39).

(38) Who are you meeting for the job interview? (Target: Who is Laura meeting for a job interview?)

(39) What is Jane reading? (Target: What is Jane reading to the students?)

Incorrect responses were further categorized into four categories: errors of inversion or raising errors, double tense or double auxiliary errors, lack of morphology errors, and other errors. Other errors included aborted or skipped trials, an embedded, *yes/no* or subject *wh*-question, VP movement, non-target *wh*-word, incorrect or omitted preposition, incorrect morphology and incorrectly adding a determiner after *whose*.

2.1.2 Results and discussion

The data was analyzed in terms of correct and inverted responses. The percentage of correct productions was found by dividing the number of correct productions by the total number of productions. The percentage of inverted responses was calculated by dividing all inverted responses (correct or incorrect) by the amount of responses code-able for inversion. Responses that were incorrect in terms of, for example, lexical items or *wh*-word used, but applied subject-auxiliary inversion were included in the % inverted analysis. Responses not code-able for

inversion were, for example, those that contained two auxiliaries, embedded questions and skipped trials.

Table 5. *L2 learners' written production of main wh-questions by coding category (strict)*

Strict coding category			
Correct	258 (58.6%)		
Non-inverted	34 (7.7%)		
Double tense/aux	20 (4.5%)		
Lack of morphology	8 (1.8%)		
Non-target aux	44 (10%)		
Other	76 (17.2%)		
Embedded Question	5		
Incorrect / Omitted Object	4	Relative Clause	4
Incorrect / Omitted Phrase	16	Lexical Items	2
Incorrect / Omitted Preposition	5	Skipped	3
Incorrect Morphology	3	Subject Wh-question	5
Incorrect Possessor	2	VP Movement	3
Incorrect Subject	6	Whose Marked Structure	14
Incorrect Wh-Word	2	Yes/No Question	2
<i>Total</i>	<i>440</i>		

Table 5 shows the distribution of production errors following the strict coding scheme. 58.6% of productions were coded as correct and target-like. A total of 34 errors of inversion (7.7%) were found. 10% of errors made came from participants using a non-target auxiliary verb (i.e., *be* instead of target *have* or changing the verb from auxiliary to lexical and applying *do*-support). These non-target auxiliary errors were coded as correct in the lax coding level. 4.5% of responses contained a double tense or double auxiliary error. One could hypothesize that, in a majority of these responses, participants were copying the tensed main verb from the prompt and adding *do*-support in their response (e.g., “What does your son cooked for dinner?” and “Where does your sister usually travels to in June?”). Errors coded as *other* made up 17.2% of responses. Most of these other errors came from incorrect or omitted phrases or participants producing what I will refer to as a *whose* marked structure. This structure, marked as incorrect and non-native-like, included responses such as “Whose is the hat that Joe is wearing?” and “Whose is the shirt that Anna cleaned?” A more in-depth look at errors produced in *whose*-questions is found below in Section 2.1.2.3.

Table 6. *L2 learners' written production of main wh-questions by coding category (lax)*

Lax coding category			
Correct	333 (75.7%)		
Non-inverted	34 (7.7%)		
Double tense/aux	20 (4.5%)		
Lack of morphology	8 (1.8%)		
Other	45 (10.2%)		
Embedded Question	5	Skipped	3
Incorrect / Omitted Preposition	2	Subject Wh-question	5
Incorrect Morphology	3	VP Movement	3
Incorrect Wh-Word	2	Whose Marked Structure	14
Lexical Items	2	Yes/No Question	2
Relative Clause	4		
<i>Total</i>	<i>440</i>		

Table 6 shows the distribution of errors following the lax coding scheme. 75.7% of productions were coded as correct. Non-target auxiliaries and omitted phrases coded as incorrect following the strict coding scheme were coded as correct in the lax scheme.

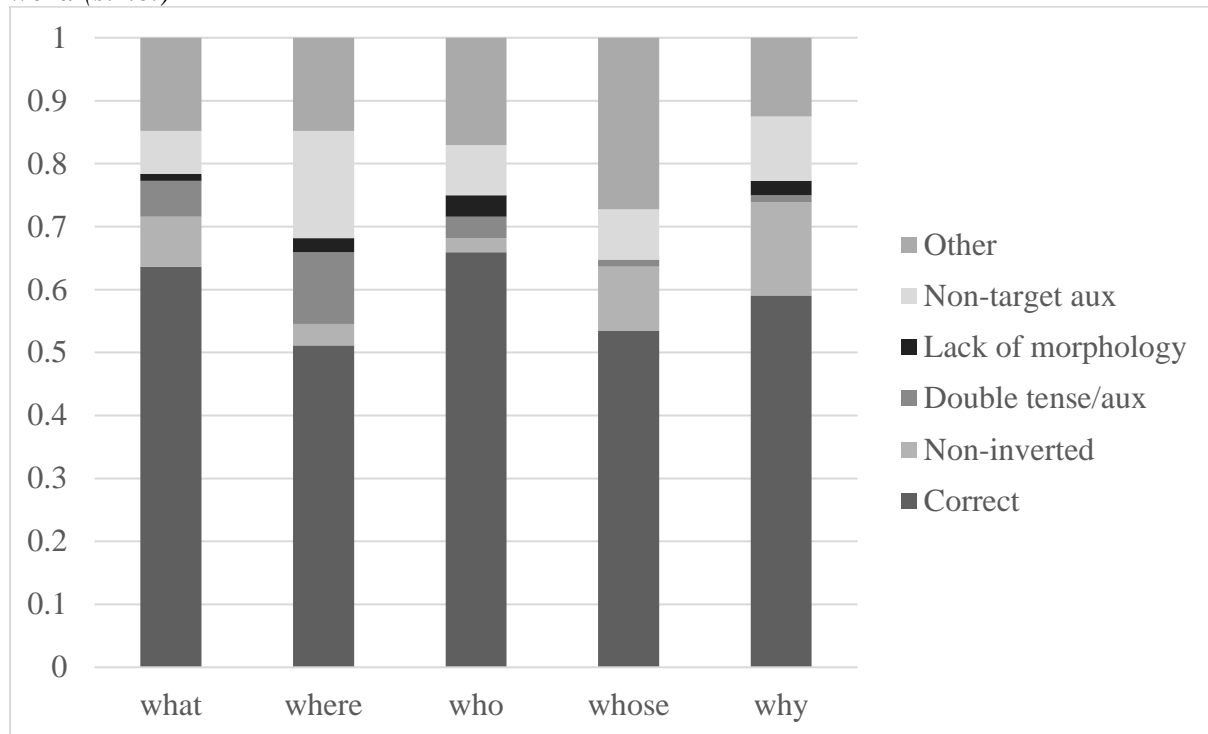
2.1.2.1 *Wh-word*

Table 7 and Figure 1 show Experiment 1 results by *wh-word*. As predicted, *why*-questions are associated with the lowest rate of inversion and *whose*-questions are associated with a lower percentage of correct and inverted responses. Within the strict coding scheme, *where*-questions are shown to have the lowest amount of correct responses. This is largely due to the non-target auxiliary errors coded as correct in the lax level. *Who* and *where* are associated with the highest inversion rates and *who* is associated with the highest number of correct responses.

Table 7. *L2 learners' written production of main wh-questions by wh-word*

Wh-word	Correct strict	Correct lax	Inversion errors (failure to apply)	% Inverted
what	56 (63.6%)	69 (78.4%)	7	91.3%
where	45 (51.1%)	68 (77.3%)	3	95.8%
who	58 (65.9%)	71 (80.7%)	2	97.3%
whose	47 (53.4%)	57 (64.8%)	9	86.6%
why	52 (59.1%)	68 (77.3%)	13	84.1%
<i>Total</i>	<i>258 (58.6%)</i>	<i>333 (75.7%)</i>	<i>34</i>	<i>91%</i>

Figure 1. *L2 learners' written production of main wh-questions by coding category and wh-word (strict)*



Comparing adjunct *wh*-words to arguments, argument *wh*-words were associated with slightly higher rates of inversion (Table 8). However, if data from problematic *whose*-questions is removed and the arguments *what* and *who* are compared to the adjuncts *where* and *why*, the distinction between the two is clearer, and one can see that, overall, arguments are associated with higher percent correct and inverted responses (Table 9). With that said, the adjunct *where* had a 95.8% inversion rate, compared to inversion rate of the argument *what* at 91.3%. Thus, there seems to be more of a *why*-asymmetry than an argument-adjunct asymmetry within the inversion rates of main questions.

Table 8. *L2 learners' written production of main wh-questions by wh-type*

Wh-type	Correct strict	Correct lax	% Inverted
adjunct	55.1%	77.3%	89.6%
argument	61.0%	74.6%	91.9%
<i>Total</i>	<i>58.6%</i>	<i>75.7%</i>	<i>91%</i>

Table 9. *L2 learners' written production of main wh-questions by wh-type (excluding whose)*

Wh-type	Correct strict	Correct lax	Inversion errors (failure to apply)	% Inverted
adjunct	55.1%	77.3%	16	89.6%
argument	64.8%	79.5%	9	94.2%
<i>Total</i>	<i>59.9%</i>	<i>78.4%</i>	<i>25</i>	<i>91.9%</i>

2.1.2.2 Verb type

Table 10 shows results by verb type. Comparable rates of inversion were found for auxiliary and lexical verbs. A large proportion of non-target auxiliary errors were found in the auxiliary verbs and coded as incorrect strict and correct lax. Pozzan (2011) and Pozzan and Quirk (2014) also found no significant effect of verb type on the rate of correct and inverted responses.

Table 10. *L2 learners' written production of main wh-questions by verb type*

Verb type	Correct strict	Correct lax	Inversion errors (failure to apply)	% Inverted
aux	126 (57.3%)	174 (79.1%)	18	90.9%
lexical	132 (60%)	159 (72.3%)	16	91.1%
<i>Total</i>	<i>258 (58.6%)</i>	<i>333 (75.7%)</i>	<i>34</i>	<i>91%</i>

2.1.2.3 Whose

Whose was added as an additional *wh*-word because of its unique structure in English. Furthermore, little research has been done to study L2 acquisition of the English possessive *whose*. Overall, looking at Table 7, one can see that *whose* is associated with a low number of correct responses as well as a lower rate of inversion. The most frequent error associated with *whose* was what was coded as a *whose* marked structure. 16% of productions in *whose* trials included this structure. The structure, in which the NP is not pied-piped to the possessive *whose*, was coded as incorrect and non-native-like. While attested in British English, this structure is infrequent in the input. It reflects a direct translation of the Italian *di chi* structure found in Table 2. (40) below shows some of the responses in which participants used this structure. Similar to Italian, the subject of the main clause is the NP and the complementizer *that* introduces the relative clause.

- (40) a. Whose are the keys that you found in the classroom?
 b. Whose is the bag that bill usually carries to the airport?
 c. Whose is the hat?

Participants also produced a similar relative clause structure in which the NP remained pied-piped to possessive *whose*. Again, these productions were coded as incorrect and non-native-like. (41) shows some examples.

- (41) a. Whose jacket was it that Joe wore to the party?
 b. Whose phone is the one Gemma found in the car?
 c. Whose jacket was the one J took to the party?

Table 11. *L2 learners' written production of whose-questions by CEFR self-rate level*

CEFR level	Correct strict	Correct lax	% Inverted
B2	5.0%	5.0%	16.7%
C1	66.7%	81.7%	92.6%
C2	75.0%	87.5%	100%
<i>Total</i>	<i>53.4%</i>	<i>64.8%</i>	<i>86.6%</i>

Table 11 shows *whose* data by participants' self-rated CEFR level. One can see that there is a positive correlation between proficiency level and percent correct and inverted responses. It seems that *whose* is mastered at later stages of acquisition but because participants' proficiency level was not balanced, more data from B2 and C2 level participants is necessary to fully understand when *whose* is mastered as a *wh*-word. The following section explores overall performance by self-rated proficiency level.

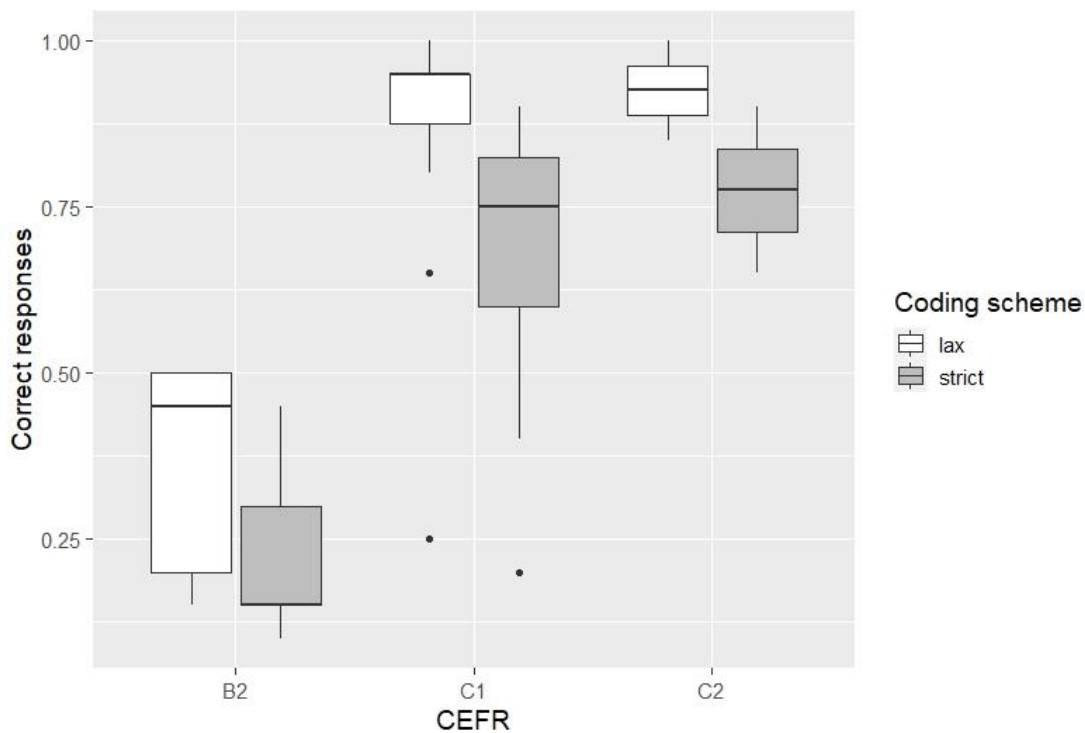
2.1.2.4 Proficiency level

Participants were asked to self-rate their English level according to the CEFR scale. Five participants rated their level a B2, 15 a C1 and two a C2. Additionally, they were asked to rate how well they speak, understand, read and write in English on a scale of 1 (not well at all) to 7 (extremely well). The mean of these four scores was taken to reflect their average English ability score. Table 12 and Figure 2 show performance by CEFR self-rate level. The CEFR level of participants was not balanced, with the majority of participants a C1 level. However, the data shows the expected trend of percent correct and inverted responses increasing as the proficiency level increases. Because the study was focused on performance of intermediate/advanced learners as a whole, the CEFR levels were not balanced. A follow-up study including more C2, B2 and even B1 level participants would be helpful in identifying trends of performance based on proficiency levels.

Table 12. *L2 learners' written production of main wh-questions by self-rated proficiency*

CEFR level	Mean English ability	Correct strict	Correct lax	% Inverted
B2	4.80	23.0%	36.0%	66.1%
C1	5.88	68.0%	86.7%	95.0%
C2	6.13	77.5%	92.5%	100%

Figure 2. *Correct responses of main wh-questions by CEFR self-rated level*



Experiment 1 was designed to elicit main *wh*-questions. Overall, the data shows a positive correlation between proficiency level and correct and inverted responses. Furthermore, participants tended to invert less in *why*-questions and showed difficulty producing *whose*-questions. After completing Experiment 1, participants completed Experiment 2 which was designed to elicit embedded *wh*-questions.

2.2 Experiment 2: Embedded questions

The goal of Experiment 2 is to investigate the participants' mastery of the structure of English embedded questions through an elicited written production task. The main objective of Experiment 2 is to measure the rate of cancel inversion, analyzing whether or not the participants incorrectly apply subject-auxiliary inversion in embedded question contexts. As mentioned before, studies have shown that the structure of English embedded questions is introduced and mastered at a later L2 acquisition level than that of English main questions. In Experiment 2, I expect to see more incorrect responses as well as a higher rate of participants incorrectly applying subject-auxiliary inversion (compared to the rate of participants failing to apply subject-auxiliary inversion in Experiment 1). Additionally, as past studies have shown, I expect to see a higher rate of correct, non-inverted responses in *why*-questions (see Pozzan, 2011; Pozzan & Quirk, 2014). Similar to Experiment 1, I also expect that participants will have

difficulty producing correct, non-inverted *whose*-embedded questions due to the unique nature of this structure in English. Past studies have attributed errors of non-inversion to either L1 transfer or L2 interference. If participants that correctly applied subject-auxiliary inversion in Experiment 1 are found to incorrectly apply subject-auxiliary inversion in Experiment 2, it can be argued that they are overgeneralizing the rule of inversion from main to embedded contexts, showing interference of the L2 itself. If participants invert the entire VP, mirroring the structure of Italian embedded questions, then it can be argued that they are transferring properties from the L1.

2.2.1 Method

2.2.1.1 Participants

The same participants that completed Experiment 1 completed Experiment 2. Experiment 2 was always completed after Experiment 1 because, following Pozzan (2011) and Pozzan and Quirk (2014), main questions were used to elicit the production of embedded questions.

2.2.1.2 Materials

Experiment 2 was administered in the same way as Experiment 1. Experiment 2 was created using PsychoPy3 software and implemented online through the Pavlovia platform. Experiment 2 directly followed Experiment 1 and was also modeled off of the embedded questions experiment performed in Pozzan (2011) and Pozzan and Quirk (2014). In Experiment 2 Phil was now asking Ms. Smith questions. Participants were introduced to a new character, Bridget, who was curious as to what Phil was asking. The participants were instructed to tell Bridget what Phil wanted to know. The questions used in Experiment 2 were the target productions from Experiment 1. Participants were shown the most different question in Experiment 2 than the one they would have produced in Experiment 1. If a participant was in List 1 in Experiment 1, she was exposed to the produced questions from List 2 in Experiment 2 and vice versa.

The first screen showed images of Ms. Smith and Phil, within which Phil was asking Ms. Smith a question. The participants were instructed to read the question aloud and then to advance to the next screen which showed Bridget asking, “What does Phil want to know?” The participants were instructed to type their response in full sentence form starting with “Phil wants to know” (e.g., “Phil wants to know why Mark is complaining to the waiter”). Participants could not return to the previous screen. Photos of the experiment screens can be found in Appendix C.

Since the experimenter could not be present, a demo was included in which Phil asked Ms. Smith, “Does Sue work today?” The experiment screen then changed to next screen showing Bridget. In the speech bubble Bridget asked, “What does Phil want to know?” The demo continued to type out “Phil wants to know if Sue works today.” Photos of the experiment demo can be found in Appendix C. After seeing the demo, participants were shown six practice items and 20 experimental items. The practice items included six main questions, three subject *wh*- questions and three *yes/no* questions. The practice items were again taken from Pozzan (2011) and Pozzan and Quirk (2014). A list of the practice items used is included in Appendix D. As in Experiment 1, the fully-within factor was presence of an auxiliary verb, with lexical verbs appearing with *do*-support in the main question prompt but not requiring it in the target embedded question. The auxiliaries used were *be* and *have*. *Wh*-word was a within-subjects but a between-items factor. The five question words were the adjuncts *where* and *why* and the arguments *who*, *what*, and *whose*. Half of the verbs appeared in past tense, with the other half in the present tense. The outline below shows the breakdown of the 20 experimental items intended to elicit embedded questions and a full list of the items can be found in Appendix D.

20 *wh*-questions

- 10 lexical verbs
 - 2 *why*, 2 *who*, 2 *where*, 2 *what*, 2 *whose*
- 10 auxiliary verbs
 - 5 auxiliary *be*
 - 1 *why*, 1 *who*, 1 *where*, 1 *what*, 1 *whose*
 - 5 auxiliary *have*
 - 1 *why*, 1 *who*, 1 *where*, 1 *what*, 1 *whose*

In order to ensure that participants were not simply quoting the main question but producing it in an authentic embedded form, a quarter of the prompts included a second person possessive subject (e.g., “Where does your sister usually travel in June?”). The intended target embedded productions were to include a third person possessive (e.g., “Phil wants to know where her/Ms. Smith’s sister usually travels in June”). The order participants were shown the items was randomized in order to control for order effects.

2.2.1.3 Procedure

Participants were automatically redirected to Experiment 2 after completing Experiment 1 and were given detailed instructions and shown a demo before proceeding to the practice items. Experiment 2 took approximately 15 minutes to complete.

2.2.1.4 Transcription and coding

Two coding schemes were used to code the participants' responses – strict coding and lax coding. Within the strict coding scheme, responses were coded as either correct (native-like) or incorrect (non-native-like) in terms of word order, verb type (auxiliary verb in the target was not changed to a lexical verb in the response), verb morphology, question type (subject *wh-* or *yes/no* questions instead of an object *wh-*) and target-like subject and objects. In order to be coded as correct, the semantics and structure of the response had to be the same as the target. Semantic substitutions, as in (42), were coded as correct.

(42) Phil wants to know whose cap Joe is wearing. (Target: Phil wants to know whose baseball hat Joe is wearing.)

Incorrect responses were coded into five further categories: incorrectly applying subject-auxiliary inversion (43) or *do*-support (44), double tense or double auxiliary errors (45), lack of morphology (46) or omitted auxiliary (47), non-target verb type (48) and other. The category of *other* included aborted or skipped trials, incorrect or omitted phrases, prepositions or objects, non-target *wh*-words (49), subject *wh*-questions (50), incorrect auxiliaries (51), incorrect morphology (52), non-target subjects, VP movement (53), non-target lexical items (54), incorrectly adding a determiner after *whose* (55), and producing a *whose* non-target marked structure (56).

(43) Phil wants to know who has Luke visited in hospital.

(44) Phil wants to know what does Tom teach at the university.

(45) Phil wants to know whose phone is Emma has found in her car.

(46) Phil wants to know where dan go every day at noon.

(47) Phil wants to know why her brother buying two bottles of wine.

(48) Phil wants to know where her daughter has hidden the phone. (Target: Phil wants to know where her daughter hid her phone.)

(49) Phil wants to know why Luke visits every Monday. (Target: Phil wants to know who Luke visits every Monday.)

(50) Phil wants to know who is complimenting molly in the hallway. (Target: Phil wants to know who Molly is complimenting in the hallway.)

(51) Phil wants to know where dan is gone on vacation.

(52) Phil wants to know whose jacket Joe weared at the party.

(53) Phil wants to know whose baseball hat is wearing Jane.

(54) Phil wants to know who Mark has invited to the hospital. (Target: Phil wants to know who Luke has visited in the hospital.)

(55) Phil wants to know whose the t shirt Sara has cleaned.

(56) Phil wants to know whose is the bag that Billy carries to the airport.

Within the lax coding scheme, responses were coded as either correct or incorrect in terms of word order, question type, and verb morphology. Responses that were grammatical and maintained a target-like message but that were not target-like in terms of verb type (i.e., an auxiliary verb in the prompt changed to a lexical verb in the response or vice versa), subject (57), and lexical items used were coded as correct. Responses that were missing phrases or objects but otherwise target-like were coded as correct (58).

(57) Phil wants to know why my husband called the store. (Target: Phil wants to know why Ms. Smith's husband called the store.)

(58) Phil wants to know what Tom teaches. (Target: Phil wants to know what Tom teaches at the university.)

Incorrect responses were further categorized into four categories: incorrectly applying subject-auxiliary inversion or *do*-support, double tense or double auxiliary errors, lack of morphology or omitted auxiliary and other. The category of *other* included aborted or skipped trials, incorrect or omitted prepositions or objects, non-target *wh*-words, subject *wh*- questions, incorrect auxiliaries, incorrect morphology, no subject, VP movement, non-target lexical items, and incorrectly adding a determiner after *whose*.

2.2.2 Results and discussion

The data was analyzed in terms of correct and non-inverted responses. As was the case in Experiment 1, the percentage of correct productions was found by dividing the number of

correct productions by the total number of productions. The percent of non-inversion was calculated by dividing all non-inverted responses (correct or incorrect) by the amount of responses code-able for non-inversion. Responses that were incorrect in terms of, for example, lexical items or *wh*-word used, but still applied cancel inversion were included in the % non-inverted analysis. Responses not able to be coded for non-inversion were, for example, those that contained omitted auxiliaries, or aborted and skipped trials.

Table 13. *L2 learners' written production of embedded wh-questions by coding category (strict)*

Strict coding category			
Correct	207 (47%)		
Inverted	83 (18.9%)		
Double tense/aux	1 (0.2%)		
Lack of morphology	22 (5%)		
Non-target aux	45 (10.2%)		
Other	82 (18.6%)		
Aborted	3	Lexical Items	7
Incorrect / Omitted Object	2	No Subject	2
Incorrect / Omitted Phrase	3	Passive	3
Incorrect / Omitted Preposition	7	Skipped	6
Incorrect Aux	3	Subject Wh-question	3
Incorrect Morphology	2	VP Movement	2
Incorrect Subject	18	Whose + Determiner	5
Incorrect Wh-Word	8	Whose Marked Structure	8
<i>Total</i>	<i>440</i>		

Table 13 shows the distribution of errors per the strict coding scheme. 47% of trials were coded as correct and native-like. A total of 83 trials (18.9%) incorrectly applied subject-auxiliary inversion. Compared to Experiment 1 in which 4.5% of responses contained a double tense or double auxiliary, only one of these errors occurred in Experiment 2. This makes sense if, as hypothesized, in Experiment 1 participants copied the tensed verb from the prompt and inserted *do*-support, resulting in a double tensed question. In Experiment 2 there was no chance of copying a tensed lexical verb from the prompt. 5% of responses showed a lack of morphology and 10.2% of responses contained a non-target auxiliary. These types of non-target aux errors were coded as correct in the lax coding scheme. 18.6% of errors were coded as *other*. Within these, the most frequent error was producing an incorrect subject. These largely included failures to change the 2nd person possessive *your* in the prompt to 3rd person *her/Ms. Smith's* (e.g., failure to produce “Phil wants to know what her/Ms. Smith’s son is cooking”

with the prompt “What is your son cooking?”) Incorrect subject errors were coded as correct in the lax coding scheme. Again, participants produced the *whose* marked structure in Experiment 2, although less frequent with just eight instances. Section 2.2.2.3 discusses errors associated with *whose*.

Table 14. *L2 learners’ written production of embedded wh-questions by coding category (lax)*

Lax coding category			
Correct	279 (63.4%)		
Inverted	83 (18.9%)		
Double tense/aux	1 (0.2%)		
Lack of morphology	22 (5%)		
Other	55 (12.5%)		
Aborted	3	No Subject	2
Incorrect / Omitted Object	1	Passive	3
Incorrect / Omitted Preposition	4	Skipped	6
Incorrect Aux	3	Subject Wh-question	3
Incorrect Morphology	2	VP Movement	2
Incorrect Subject	1	Whose + Determiner	5
Incorrect Wh-Word	8	Whose Marked Structure	8
Lexical Items	4		
<i>Total</i>	<i>440</i>		

Table 14 shows the distribution of errors per the lax coding scheme. 63.4% of all responses were coded as correct and native-like. 12.5% of errors were coded as *other*. Within the *other* errors, the most frequent errors were the use of a non-target *wh*-word and the *whose* marked structure.

Participants made a larger number of inversion errors (incorrectly applying) in Experiment 2 (83) compared to failure to apply inversion in Experiment 1 (34). The rate of inversion in Experiment 1 was higher at 91% compared to the rate of cancel inversion in Experiment 2 at 79.5%. Furthermore, the total number of correct responses was lower in Experiment 2 than Experiment 1. 58.6% of responses were strictly coded as correct in Experiment 1 compared to 47% of responses in Experiment 2. Overall, these results seem to support the theory that the structure of main *wh*-questions is acquired and mastered before that of embedded *wh*-questions. These results are consistent with the hypothesis that participants make fewer errors in main questions compared to embedded questions.

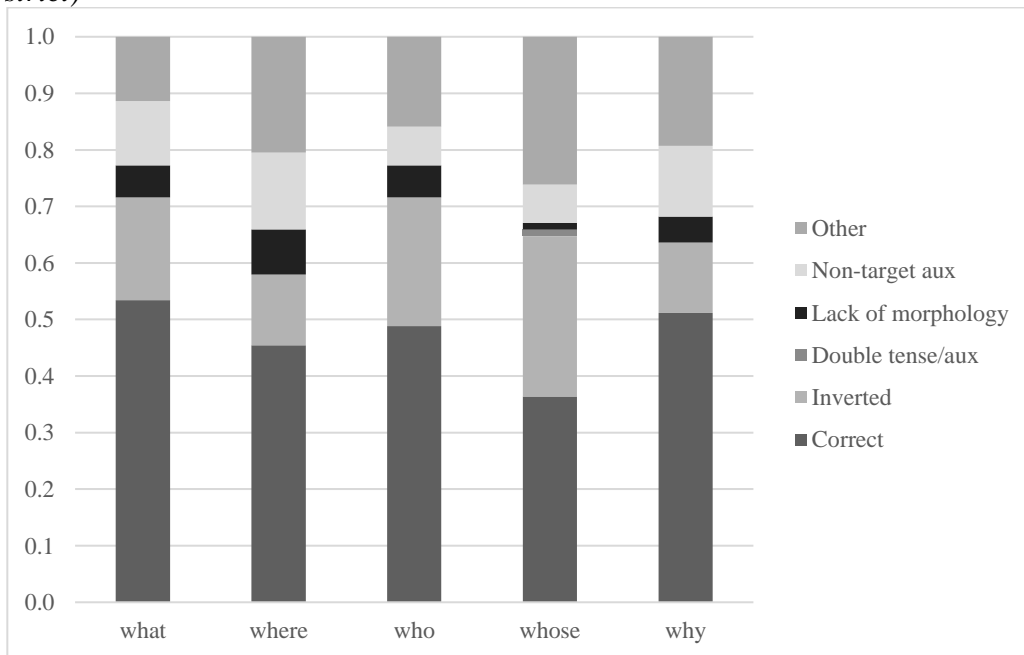
2.2.2.1 *Wh-word*

Table 15 and Figure 3 show data from embedded questions by *wh*-word. Overall, *whose* is associated with the lowest number of correct, native-like responses and the lowest percent non-inverted responses. The results show that *why* is associated with the highest percent non-inverted responses. This was also the case in Experiment 1. As mentioned, previous studies have found *why* to be associated with lower rates of inversion in both main and embedded contexts and in both L1 and L2 acquisition. The present study shows this *why*-asymmetry in which in main questions, *why* is associated with fewer correct, inverted responses while in embedded questions, *why* is associated more correct, non-inverted responses.

Table 15. *L2 learners' written production of embedded wh-questions by wh-word*

Wh-word	Correct strict	Correct lax	Inversion errors	
			(incorrectly applying)	% Non-inverted
what	47 (53.4%)	63 (71.6%)	16	81.4%
where	40 (45.5%)	59 (67%)	11	86.9%
who	43 (48.9%)	52 (59.1%)	20	75.6%
whose	32 (36.4%)	38 (43.2%)	25	62.7%
why	45 (51.1%)	67 (76.1%)	11	87.1%
<i>Total</i>	<i>207 (47%)</i>	<i>279 (63.4%)</i>	<i>83</i>	<i>79.5%</i>

Figure 3. *L2 learners' production of embedded wh-questions by coding category and wh-word (strict)*



While an argument-adjunct asymmetry was not found in Experiment 1, Experiment 2 results show a more apparent asymmetry in which arguments are associated with higher rates

of inversion. Specifically, *who*, which in Experiment 1 had the highest rate of inversion, is found to have the second lowest rate of cancel inversion in Experiment 2. Table 16 shows that participants applied cancel inversion more frequently in adjuncts than in arguments. In Experiment 1, the adjunct *where* showed higher rates of correct, inverted responses and in Experiment 2 *where* shows higher rates of correct, non-inverted responses. Thus, there seems to be an argument-adjunct asymmetry found in embedded contexts but not in main contexts, where only an asymmetry with *why*-questions was found.

Table 16. *L2 learners' written production of embedded wh-questions by wh-type*

Wh-type	Correct strict	Correct lax	% Non-inverted
adjunct	48.3%	71.6%	87.0%
argument	46.2%	58.0%	74.0%
<i>Total</i>	<i>47%</i>	<i>63.4%</i>	<i>79.5%</i>

2.2.2.2 Verb type

Table 17 shows results by verb type. Overall, lexical verbs were associated with a higher amount of correct responses and a higher rate of non-inverted responses. In 37 (44%) of the total inversion errors made, participants incorrectly applied *do*-support and in the remaining 47 (56%), participants incorrectly inverted the auxiliary and the subject. Participants were more likely to incorrectly invert the auxiliary than to apply *do*-support in embedded contexts where it is not needed.

Table 17. *L2 learners' written production of embedded wh-questions by verb type*

Verb type	Correct strict	Correct lax	Inversion errors	
			(incorrectly applying)	% Non-inverted
aux	92 (41.8%)	137 (62.3%)	48	76.0%
lexical	115 (52.3%)	142 (64.5%)	35	82.8%
<i>Total</i>	<i>207 (47%)</i>	<i>279 (63.4%)</i>	<i>83</i>	<i>79.5%</i>

2.2.2.3 Whose

Data from Experiment 1 showed *whose*-questions were associated with a lower percentage of correct responses and rate of inversion. Experiment 2 shows similar results. As noted in Table 15, *whose* is associated with the lowest number of correct responses and the lowest rate of cancel inversion. Below are some examples of *whose* productions. In Experiment 2 participants produced the *whose* marked structure (59) less frequently than in Experiment 1 (8 compared to 14).

- (59) a. Phil wants to know whose is the book that bill is carrying to the class.
b. Phil wants to know whose is the shirt anna cleaned.

Some participants seemed to apply different strategies in order to cope with having to produce the unacquired structure. A few produced passive structures as in (60), while some used a non-target *wh*-word or phrase (61). Two participants incorrectly produced *whose* alongside a determiner (62).

- (60) a. Phil wants to know whose bag was carried at the airport.
b. Phil wants to know whose phone was found in the car by Kate.

- (61) a. Phil wants to know what kind of jacket Dan wear at the party.
b. Phil wants to know who owns the basketball hat.

- (62) a. Phil wants to know whose the key the teacher found in the classroom.
b. Phil wants to know whose the shirt anna has cleaned.

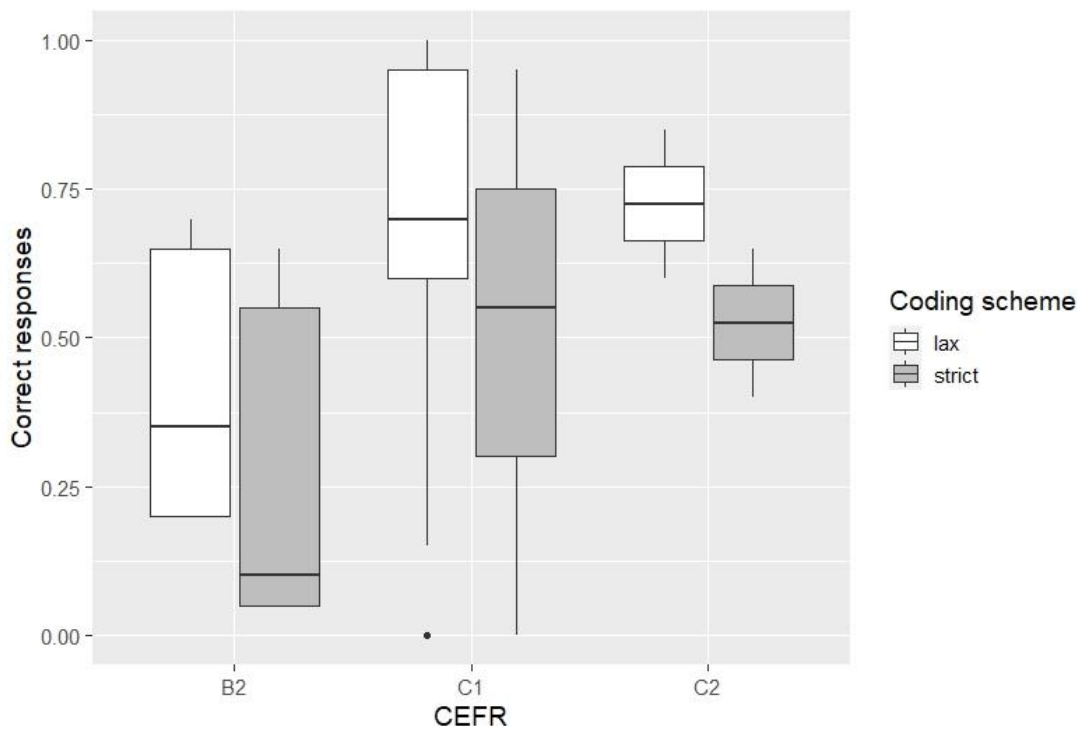
2.2.2.4 Proficiency level

Table 18 and Figure 4 show the results of Experiment 2 by self-rated proficiency level. As expected, the rate of correct responses increases as the level of proficiency increases. The B2 level participants, however, were unexpectedly shown to produce the highest rate of non-inverted responses. Again, the sample size of B2 participants is low and more participants are needed in order to identify how proficiency affects rates of non-inversion.

Table 18. *L2 learners' written production of embedded wh-questions by self-rated proficiency*

CEFR level	Mean English ability	Correct strict	Correct lax	% Non-inverted
B2	4.80	28.0%	42.0%	92.1%
C1	5.88	51.7%	69.3%	75.6%
C2	6.13	52.5%	72.5%	83.8%

Figure 4. *Correct responses of embedded wh-questions by CEFR self-rated level*



In order to expand upon the initial proficiency results, additional participants were recruited to investigate how living in an immersive context affects acquisition. Below the pilot data of L2 learners living in an English-speaking environment is presented.

2.3 L2 speakers in an immersive context

To further explore L1 Italian speakers' production of English main and embedded questions, Experiments 1 and 2 were distributed to L1 Italian speakers with a highly proficient level of English who had been living in an English-speaking country for at least two years. The goal was to gather initial pilot data to explore how living in an English immersive context affects the acquisition and mastery of English interrogative structures. I expected to see fewer errors in both experiments and few, if any, inversion errors in main questions.

2.3.1 Method

2.3.1.1 Participants

Participants were recruited online via Facebook groups for Italian expats. Participant requirements were that they were Italian native speakers who had been living in a dominant English-speaking country, uninterrupted, for at least two years. After filling in the Participant Contact Form, participants were sent the Bilingual Speaker Language Profile Questionnaire.

The questionnaire, which was adapted to apply to bilingual speakers, asked participants to self-rate their CEFR level, and their reading, writing, speaking and comprehension abilities in both English and Italian. Additionally, participants were asked about their average exposure and use of both English and Italian in a given week. A total of nine participants completed the study. Six participants were living in the UK at the time of completion and three were living in the US. Two participants had completed high school, three had completed or were completing a university master's degree, and four had completed or were completing a PhD. The majority (7) of the participants rated their level of English a C2, one rated herself a C1 and one rated herself a B2. Due to difficulties recruiting participants, participants' age, time spent living abroad and weekly exposure to and use of Italian and English varied greatly. However, participants stated they were more frequently exposed to English and used English more in an average week than Italian. Table 19 shows participant demographics.

Table 19. L1 Italian participants living in an immersive environment demographic data

	Mean	Range	Standard deviation
Age	35.3	25 - 59	12.2
English ability rated 1 - 7	6.5	5.5 - 7	0.6
Time spent abroad (in years)	11.1	2.1 - 34	11.3
Weekly English exposure in hours	114.4	37 - 168	47.5
Weekly English use in hours	107.4	12 - 168	53.4
Weekly Italian exposure in hours	62	2 - 168	56.1
Weekly Italian use in hours	40.4	5 - 116	44.5
Age when started to learn English	9.3	3 - 16	4.5
Age when felt comfortable with English	20.2	16 - 28	4.5
Years of English courses taken	11.3	5 - 19	4.8

2.3.1.2 Materials

Participants completed both Experiment 1 and 2. The materials were the same used in the study with non-immersive L2 learners.

2.3.1.3 Procedure

Participants were recruited online through Facebook groups for Italian expats (e.g., *Italians in London, Italians in Cambridge*) and filled out the Participant Contact Form. Participants were then sent an email with a link to the Bilingual Speaker Language Profile Questionnaire. At least 24 hours after consenting and completing the questionnaire, participants were sent another email with a link to the study on Pavlovia. Participants completed Experiment 1 and then were redirected to Experiment 2. Both experiments took approximately 15 minutes to complete.

2.3.1.4 Transcription and coding

Responses were coded as correct (native-like) or incorrect (non-native-like). The same coding schemes used to code the data from the non-immersive learners were used to code the data from the immersive learners.

2.3.2 Results and discussion

Results were analyzed in terms of correct and non-inverted responses, using the same criteria as the previous study.

2.3.2.1 Experiment 1: main questions

Tables 20 and 21 show results by error type for both strict and lax coding schemes. 66.7% of productions were coded as strict correct and 91.7% of responses were coded as lax correct. Only three inversion errors were found, and they all were produced by the same participant.

Table 20. *L2 immersive learners' written production of main wh-questions by coding category (strict)*

Strict coding category			
Correct	120 (66.7%)		
Non-Inverted	3 (1.7%)		
Double tense/aux	4 (2.2%)		
Non-target aux	22 (12.2%)		
Other	31 (17.2%)		
Aborted	1	Incorrect subject	1
Embedded Questions	4	Lexical Items	3
Incorrect / Omitted object	7	Skipped	1
Incorrect / Omitted phrase	8	Subject Wh-	1
Incorrect / Omitted preposition	1	Whose Subextraction	1
Incorrect aux	1	Yes/No Question	1
Incorrect morphology	1		
<i>Total</i>	<i>180</i>		

Table 21. *L2 immersive learners' written production of main wh-questions by coding category (lax)*

Lax coding level	
Correct	165 (91.7%)
Non-Inverted	3 (1.7%)
Double tense/aux	4 (2.2%)
Other	8 (4.4%)
Aborted	1
Incorrect / Omitted preposition	1
Incorrect morphology	1
Lexical Items	1
Skipped	1
Subject Wh-	1
Whose Subextraction	1
Yes/No Question	1
<i>Total</i>	<i>180</i>

Table 22 shows data by *wh*-word. Not surprisingly, all inversion errors occurred in *why* or *whose*-questions. *Whose* is associated with slightly lower percent correct responses, although less dramatic than the results with L2 non-immersive learners showed. The low percent correct strict responses in *where*-questions largely comes from participants using a non-target auxiliary and omitting phrases. Inversion rates were expected to be high and all but one participant inverted in 100% of their responses. The participant who produced inversion errors rated himself a C2 level and his English abilities a 7. Additionally, he stated that he had lived in the US for 30 years. These results are unexpected, and it is unclear why someone with such a high proficiency in English would continue to make inversion errors in main *wh*-questions.

Table 22. *L2 immersive learners' written production of main wh-questions by wh-word*

Wh-word	Correct strict	Correct lax	Inversion errors (failure to apply)	% Inverted
what	77.8%	94.4%	0	100%
where	58.3%	91.7%	0	100%
who	66.7%	94.4%	0	100%
whose	61.1%	88.9%	1	97.1%
why	69.4%	88.9%	2	94.3%
<i>Total</i>	<i>66.7%</i>	<i>91.7%</i>	<i>3</i>	<i>98.3%</i>

2.3.2.2 Experiment 2: embedded questions

Tables 23 and 24 show results by error type for both strict and lax coding schemes. Per the strict coding scheme, 58.3% of responses were correct and 81.1% of responses were coded as

correct per the lax coding scheme. Inverted responses represented 10% of total responses, compared to 18.9% found within the L2 non-immersive learners. Four of the nine participants produced at least one incorrect, inverted response.

Table 23. *L2 immersive learners' written production of embedded wh-questions by coding category (strict)*

Strict coding category			
Correct	105 (58.3%)		
Inverted	18 (10%)		
Lack of morphology	3 (1.7%)		
Non-target aux	23 (12.8%)		
Other	31 (17.2%)		
Aborted	3	Incorrect wh-word	2
Incorrect / Omitted object	2	Lexical Items	3
Incorrect / Omitted phrase	1	No Subject	1
Incorrect aux	1	Skipped	1
Incorrect subject	14	Whose Marked	2
Incorrect morphology	1		
<i>Total</i>	<i>180</i>		

Table 24. *L2 immersive learners' written production of embedded wh-questions by coding category (lax)*

Lax coding category			
Correct	146 (81.1%)		
Inverted	18 (10%)		
Lack of morphology	3 (1.7%)		
Other	13 (7.2%)		
Aborted	3	Incorrect wh-word	2
Incorrect / Omitted object	2	Lexical Items	3
Incorrect / Omitted phrase	1	No Subject	1
Incorrect aux	1	Skipped	1
Incorrect subject	14	Whose Marked	2
Incorrect morphology	1		
<i>Total</i>	<i>180</i>		

The total rate of non-inverted responses was 89.5%, compared to 82.8% from the L2 non-immersive learners' results. Table 25 shows results by *wh*-word. *Whose* continues to be associated with low rates of correct and non-inverted responses. Furthermore, *why* is associated with the highest rates of non-inversion. *Who* is also associated with low non-inversion rates. Table 26 shows the results for arguments and adjuncts. Overall, participants produced incorrect, inverted responses more frequently in arguments than in adjuncts. In fact, adjuncts

were associated with a high percent non-inverted rate of 95.7%. However, the argument *what* showed higher rates of non-inversion at 94.3%, indicating difficulties in applying cancel inversion to *who* and *whose*-questions.

Table 25. *L2 immersive learners' written production of embedded wh-questions by wh-word*

Wh-word	Correct strict	Correct lax	Inversion errors (incorrectly applying)	% Non-inverted
what	63.9%	86.1%	2	94.3%
where	63.9%	91.7%	2	94.3%
who	63.9%	80.6%	6	83.3%
whose	38.9%	55.6%	7	76.7%
why	61.1%	91.7%	1	97.1%
<i>Total</i>	<i>58.3%</i>	<i>81.1%</i>	<i>18</i>	<i>89.5%</i>

Table 26. *L2 immersive learners' written production of embedded wh-questions by wh-type*

Wh-type	Correct strict	Correct lax	% Non-inverted
adjunct	62.5%	91.7%	95.7%
argument	55.6%	74.1%	85.1%
<i>Total</i>	<i>58.3%</i>	<i>81.1%</i>	<i>89.5%</i>

While appearing less frequently than in the data of non-immersive L2 learners, non-inversion errors in embedded questions were found with highly proficient L2 English speakers who have spent a significant amount of time living in an English-speaking country. Of the four participants who inverted in embedded questions, all self-rated their English level a C2 and had lived in an immersive context for at least six years. One participant had lived in an English-speaking country for 34 years and another for 30 years. The data shows that errors in the production of English *wh*-questions persist throughout very advanced learners who receive significant levels of L2 input.

A final pilot study was done with native English speakers to explore their acceptability of produced structures from the L2 participants, specifically focusing on the *whose* marked structure.

2.4 Acceptability judgment task on English questions

In order to judge the acceptability of the L2 learners' productions, an acceptability judgment task was administered to native speakers of English. Specifically, I wanted to measure the acceptability of the *whose* marked structures previously discussed. The structure seems to be attested in British English and, while unlikely that a native speaker would produce it in the

elicited written production study, I wanted to gauge how acceptable native speakers rated this structure.

2.4.1 Method

2.4.1.1 Participants

Participants were recruited online and sent the link to the Qualtrics questionnaire. All participants were native English speakers of varying varieties. 10 of the participants were speakers of American English, nine British English and two Australian English. Data from a total of 21 native English speakers was analyzed. Table 27 shows participant age data.

Table 27. Native English speakers' mean age by English variety

English variety	Age		
	Mean	Range	Standard deviation
American (n = 10)	25	22 – 27	2.05
Australian (n = 2)	23.5	23 – 24	0.5
British (n = 9)	23.7	21 – 34	3.75

2.4.1.2 Materials

The task was administered online through Qualtrics. Participants were asked to rate the sentences on a scale from 1 (No one would say this) to 5 (This is perfect). Participants were told that a rating of 3 is “I wouldn’t say this, but I hear people say this.” The participants were told to not focus on grammaticality, but whether or not the sentence was natural and something a native speaker would produce. Additionally, the instructions told the participants to not focus on lexical differences in their English variety (i.e., to not mark a sentence a 1 because it uses the word *mall* instead of *shopping center*). The task included both main and embedded questions and the experimental items were the *whose* marked structures produced by the L2 learners in the elicited written production experiments. Both grammatical and ungrammatical *what*, *why*, *where* and *yes/no* questions were included as filler items. The ungrammatical questions included errors of inversion, double tense errors, and no auxiliary errors. The embedded questions all included the main clause “Chris wants to know.” The order the participants were shown the items was randomized to control for order effects. The task included a total of 32 experimental items and 128 fillers. Below is the breakdown of the items included in the task.

I. Experimental items:

32 *whose*-questions

- 16 main questions
 - 8 *whose* marked
 - 8 *whose* unmarked
- 16 embedded questions
 - 8 *whose* marked
 - 8 *whose* unmarked

II. Fillers:

32 *what*-questions

- 16 main questions
 - 8 grammatical
 - 8 ungrammatical
- 16 embedded questions
 - 8 grammatical
 - 8 ungrammatical

32 *where*-questions

- 16 main questions
 - 8 grammatical
 - 8 ungrammatical
- 16 embedded questions
 - 8 grammatical
 - 8 ungrammatical

32 *why*-questions

- 16 main questions
 - 8 grammatical
 - 8 ungrammatical
- 16 embedded questions
 - 8 grammatical
 - 8 ungrammatical

32 *yes/no* questions

- 16 main questions
 - 8 grammatical

- 8 ungrammatical
- 16 embedded questions
 - 8 grammatical
 - 8 ungrammatical

2.4.1.3 Procedure

Participants were recruited online and given the link to the task administered through Qualtrics. After reading the consent form and consenting to participate, participants were asked to state their age, their native language and if English, their variety, any other languages known, their country of origin and their current country. After completing the demographic questions, the task began. The task took approximately 15 minutes to complete.

2.4.2 Results and discussion

Table 28 shows average ratings of *whose* marked structures and unmarked structures by English variety and sentence type, main or embedded. As previously stated, participants were instructed to rate each sentence on a scale from 1 (No one would say this) to 5 (This is perfect). A 3 rating was “I wouldn’t say this, but I hear people say this.” Overall, participants rated the *whose* marked structures lower than the *whose* unmarked structures. On average, the *whose* marked structure was rated higher when appearing as a main question versus an embedded question. Speakers of American English strongly rejected the *whose* marked structures as acceptable in both main and embedded contexts. In main questions, both speakers of Australian and British English rated the structures higher, 3.33 and 3.44 respectively. In embedded questions, Australian and British participants rated the marked structure lower compared to their average rating in main questions, but still higher than American speakers’ average ratings for both main and embedded *whose* marked structures.

Table 28. *Native English speakers’ average rating of whose marked and unmarked structures by English variety and sentence type*

English variety	<i>Whose</i> structure and sentence type			
	Marked main	Unmarked main	Marked embedded	Unmarked embedded
American	1.58	4.54	1.49	3.96
Australian	3.33	4.81	2.09	4.2
British	3.44	4.81	2.39	4.36
<i>Total</i>	<i>2.53</i>	<i>4.68</i>	<i>1.91</i>	<i>4.16</i>

Table 29 compares average ratings of the *whose* marked structures and the ungrammatical filler sentences included in the task. Interestingly, the American speakers rated the *whose* marked structures lower than the ungrammatical sentences in both main and embedded contexts. In main ungrammatical questions, the American speakers rated non-inverted *yes/no* questions higher. This is not surprising, as non-inverted *yes/no* main questions are attested in the speech of native English speakers. If ungrammatical *yes/no* question data is removed from the analysis, the American speakers' average rating of *whose* marked main questions becomes slightly higher than their rating of ungrammatical main questions. However, the results for embedded questions remain the same with *whose* ungrammatical structures having a lower average rating than all other ungrammatical embedded structures. On the other hand, the Australian and British speakers rated the *whose* marked structures higher than the ungrammatical structures in both main and embedded contexts, with the difference in main questions much greater.

Table 29. Average ratings of *whose* marked structures and ungrammatical structures by English variety and sentence type

English variety	Structure and sentence type			
	<i>Whose</i> marked main	Ungrammatical main	<i>Whose</i> marked embedded	Ungrammatical embedded
American	1.58	1.63	1.47	1.69
Australian	3.33	1.98	2.09	2.04
British	3.44	1.79	2.39	1.93
<i>Total</i>	<i>2.53</i>	<i>1.73</i>	<i>1.91</i>	<i>1.82</i>

A final analysis was performed to explore which *whose* marked structures were most acceptable for the native speakers. Data from the American speakers was not included in this analysis as they were found to reject the acceptability of the *whose* marked structures in all contexts. The highest rated marked structures by the Australian and British participants were “Whose is the hat that Joe is wearing?” and “Whose is the hat?” both receiving a rating of 4 out of 5. Furthermore, “Whose are they keys that you found in the classroom?” and “Whose is the shirt that Anna cleaned?” both had an average rating of 3.9. Within the embedded questions, the most accepted sentence was “Chris wants to know whose is the shirt Anna cleaned.” Its average rating was lower, however, at 2.89. A table with the average ratings of all the *whose* marked structures by Australian and British speakers is found in Appendix G.

To conclude, the data shows that, on average, participants of all varieties rated the acceptability of the *whose* marked structures lower than the *whose* unmarked structures.

Speakers of American English strongly rejected the acceptability of the *whose* marked structures in both main and embedded contexts, rating them lower than the filler ungrammatical sentences. The Australian and British English speakers, however, rated the *whose* marked structures in main questions as more acceptable. Based on their average ratings, it can be inferred that this structure is not necessarily something they would produce but something they hear others say. Keeping the data of this acceptability judgment task in mind, it is still appropriate to code the L2 learners' productions of the *whose* marked structure as incorrect and non-native-like. The native speakers, overall, were shown to prefer the *whose* unmarked structures to the marked structures.

3. DISCUSSION

The goal of the present study was to investigate the errors made by L1 Italian learners of English in the production of English *wh*-questions. The current study was adapted from the elicited oral production study in Pozzan (2011) and Pozzan and Quirk (2014). A direct comparison between the two studies would not be helpful for a few reasons. Firstly, Pozzan's study elicited oral production and the study presented in this thesis elicited written production. While participants were told to type their answers as fast as possible, written production allows for more time to think and is less spontaneous than oral production. For these reasons, it is also less challenging. Secondly, Pozzan reported on data from L1 Spanish and L1 Chinese participants living in the US. The present study included L1 Italian participants living in Italy. Therefore, one can assume that the L1 Italian participants have, on average, lower exposure to English and use English less frequently in a given week compared to the participants in Pozzan's study. Finally, the present study investigated an additional *wh*-word, *whose*. With that said, however, there are patterns found in both studies. As mentioned, both studies found *why* to be associated with lower inversion rates in both main and embedded questions. Furthermore, both studies found a lower percentage of correct responses in embedded *wh*-questions than in main *wh*-questions. This result is expected and follows the stages of L2 acquisition of English questions presented in Section 1.5 (see Pienemann, 1998; Pienemann & Keßler, 2011; Spada & Lightbown, 1999).

3.1 Question type

It is clear that the structure of embedded questions in English is more complex and acquired after that of main questions. In order to examine which participants had acquired and mastered each structure, an analysis examining rates of inversion in each experiment was done. Table 30 shows a list of participants and their inversion rates in each experiment. Using 90% as the threshold (i.e., a participant inverted 90% of responses in Experiment 1 and applied cancel inversion 90% of the time in Experiment 2) eight participants were shown to have mastered both main and embedded *wh*-questions. Nine participants mastered main questions but not embedded questions. Two were shown to have mastered neither and, unexpectedly, three participants were shown to have mastered embedded questions but not main questions. Of these three, participant 22 showed a low inversion rate in Experiment 1 (28.6%) and applied cancel inversion 100% of the time in Experiment 2. It seems that this participant was not consistently

applying inversion in either experiment and it's unclear why this was the case. The other two, participants 19 and 20, do not follow the stages of acquisition and are exceptions to the rule.

Table 30. *L2 learners' inversion rates in main wh- and embedded wh-questions by participant*

Participant	CEFR rating	Mean English ability	Inversion rate MQ	Non-inversion rate EQ	MQ mastered (90%)	EQ mastered (90%)
1	C2	6.5	100%	90%	+	+
2	C1	5.5	100%	95%	+	+
3	C1	6	100%	0%	+	-
4	C1	5.5	100%	100%	+	+
5	B2	4.5	75%	89.5%	-	-
6	B2	5.75	100%	76.5%	+	-
7	C1	6	100%	100%	+	+
8	C1	6	100%	100%	+	+
9	C1	5.75	100%	75%	+	-
10	C1	6	100%	87.5%	+	-
11	C1	6.5	29.4%	72.2%	-	-
12	C1	6	100%	20%	+	-
13	C1	6	95%	100%	+	+
14	B2	5.5	100%	83.3%	+	-
15	C1	5.5	94.7%	80%	+	-
16	C1	5.75	100%	26.3%	+	-
17	C1	6	100%	80%	+	-
18	C1	6	100%	100%	+	+
19	B2	5.75	83.3%	100%	-	+
20	B2	5.25	50%	91.7%	-	+
21	C1	5.75	100%	95%	+	+
22	B2	3	28.6%	100%	-	+

A similar analysis was done with the pilot data from the participants living in an English immersive context and is shown in Table 31 below. Six participants out of nine were shown to have mastered both main and embedded questions. Two participants mastered main questions but not embedded while one participant mastered neither main nor embedded questions. Therefore, all but one participant was shown to have mastered main questions. These results are expected in that they pattern with previously mentioned hypotheses claiming the acquisition of main questions before embedded, and unexpected in that even some highly proficient L2 learners living in an immersive context are shown to have not mastered the structures of English questions.

Table 31. *L2 immersive learners' inversion rates in main wh- and embedded wh-questions by participant*

Participant	CEFR rating	Mean English ability	Time abroad in years	Inversion rate MQ	Non-inversion rate EQ	MQ mastered (90%)	EQ mastered (90%)
1	C2	5.75	6	100%	72.2%	+	-
2	B2	5.5	5	100%	100%	+	+
3	C2	7	34	100%	84.2%	+	-
4	C1	6.5	5	100%	100%	+	+
5	C2	7	6	100%	100%	+	+
6	C2	6.75	6	100%	90%	+	+
7	C2	7	30	83.3%	57.9%	-	-
8	C2	7	6	100%	100%	+	+
9	C2	5.75	2.08	100%	100%	+	+

3.2 L1 transfer

The present study included speakers from the same L1 and therefore, conclusions regarding L1 transfer in main questions cannot be made. In embedded questions, errors of cancel inversion have been argued as being the result of L1 transfer or L2 interference. Learners can be said to overgeneralize what they know about English main questions to embedded questions, resulting in ungrammatical inverted structures. However, Italian embedded questions require inversion. So, it is unclear in this case if the errors stem from interference of the L2 or transfer from the L1.

With that said, I hypothesize that the non-inversion errors in embedded questions are due to L2 interference and not L1 transfer. As mentioned in Chapter 1, English main questions allow for only the auxiliary to move while the main verb remains in the lower position. In Italian questions, both the auxiliary and main verb must move together, and the subject cannot intervene between the two. Therefore, direct L1 transfer from Italian to English would be characterized by movement of the entire VP in English. VP movement was found in just two embedded productions while subject-auxiliary inversion errors occurred 83 times. Thus, it seems that participants applied the structure of English main questions to embedded questions rather than transferring the structure of Italian embedded questions.

Additionally, participants produced what was called a *whose* marked structure that directly reflects the structure of Italian *di chi* questions. This structure was coded as incorrect and non-native-like. The results of the acceptability judgment task with native English speakers showed interesting results. The *whose* marked structure was strongly rejected by American English speakers in both main and embedded contexts while speakers of Australian and British

English rated it as more acceptable, notably in main questions. I'd hypothesize that the L2 speakers' tendency to produce this *whose* marked structure is due to L1 transfer simply because the frequency of this structure in the input would be extremely low. *Whose*-questions are noted as being produced infrequently, in comparison to other *wh*-questions, and exposure to this *whose* marked structure would be even less frequent. However, more data from speakers with different L1s is needed to test this L1 transfer hypothesis but it could be that learners transfer the L1 structure when having to produce an unacquired structure in the L2.

3.3 *Wh*-word

Similar to Pozzan (2011) and Pozzan and Quirk (2014), no direct argument-adjunct asymmetry was found in main *wh*-questions. The adjunct *where* was associated with high inversion rates in main questions. In embedded questions, both the adjuncts *why* and *where* showed higher rates of non-inversion than the arguments *who*, *what* and *whose*. Thus, an argument-adjunct asymmetry was more apparent in embedded questions than in main questions and a *why*-asymmetry was found in both, as L2 learners tended to invert *why*-questions less than other *wh*-questions. This has been shown to be the case in L1 acquisition and in L2 learners across different L1s. Although *perché* does not require inversion in Italian questions, the low inversion rates found in this study cannot be attributed solely to L1 influence because this trend has been found cross-linguistically with L1s that don't require inversion at all, such as Korean (Lee, 2008) and Chinese (Pozzan, 2011). To review, Rizzi (2001) claims that *perché* in Italian behaves like a higher adverbial and, unlike other adjuncts, can be generated outside the VP. Other *wh*-words in Italian occupy FocP, whereas *perché* occupies a higher position, IntP. Within L1 acquisition, Berk (2003) claims that English-speaking children treat *why* like *how come* which, briefly mentioned in Section 1.2.1, does not allow for inversion. *How come* occupies SpecIntP and does not trigger inversion. Therefore, according to Berk (2003), children who fail to invert in *why*-questions do so because they treat *why* as *how come* and regard it as a non-inverting and non-moving *wh*-phrase (Conroy & Lidz, 2007). Another hypothesis, mentioned in Section 1.5.2, brought forward by Thornton (2008), is that English-speaking children treat *why* like *perché*, merging it into SpecIntP but also allowing for it to move to SpecFocP, therefore permitting movement of *why*. Children produce *why* inversion errors when they mistakenly merge it into IntP rather than move it to FocP. These two hypotheses do not account for the results of the current study, but it is clear that *why* is syntactically represented differently than other *wh*-words cross-linguistically. Whether the low inversion rates of L1

Italian participants can be traced to the properties of *perché* in the L1 or to a property of UG that treats *why* differently cannot be determined from the present study.

The present study also found *whose* to be associated with lower rates of correct and inverted responses. As noted in Section 1.4.1, the structures of the English *whose* and Italian *di chi* are different and *whose* behaves differently than other *wh*-words in that it cannot move on its own. Past studies have shown *whose* acquisition difficulties in monolingual English children and L2 child learners (see Gavrusseva, 1997; Gavrusseva & Thornton, 2001) but more studies investigating L2 acquisition of *whose* across different L1s are needed to understand how and when learners acquire *whose*-questions. If referring back to the English Profile, it is not clear when the *wh*-word *whose* is taught to L2 learners. The ability to use the relative pronoun *whose* is a characteristic of B1 level students, but there doesn't seem to be a focus on using *whose* as a *wh*-word in questions (Cambridge ESOL & CUP, 2011). *Whose* as a *wh*-question word is used infrequently in writing and arguably even less so in spoken English. The ICLE that Pozzan (2011) analyzed included over 6,000 texts. Within these texts, the most common *wh*-word produced in main questions was, unsurprisingly, *what*, appearing 1,264 times. The second most frequent *wh*-word was *how*, appearing 699 times. In comparison, *whose* appeared just seven times within the ICLE. Within embedded questions, *whose* appeared just once in over 6,000 texts. The learners that did produce *whose*, produced it correctly 100% of the time and showed no errors of inversion. It can be inferred, and as some of the data from the current study showed, some learners employ strategies in order to avoid using *whose* if they have not fully acquired and mastered this structure.

3.4 Proficiency

The current study provided initial insight into how proficiency affects the production of *wh*-questions. The results showed the number of correct responses increases as the level of proficiency increases. In Experiment 1 this was also the case with percent inverted responses. In Experiment 2, however, it was found that B2 level participants produced higher rates of non-inversion than C1 and C2 level participants. In order to better understand this relationship, more participants of varying levels are needed. Furthermore, the present study asked participants to self-rate their proficiency level, but an objective measure of proficiency is needed in order to provide an accurate analysis of the role proficiency plays in the production of English questions. Additionally, the pilot data from the learners living in an English-speaking country showed that errors of inversion and difficulties acquiring *whose* persist

throughout highly proficient bilinguals. It seems that exposure to these structures is not enough for full acquisition and there is a need for explicit instruction of these structures.

4. CONCLUSIONS

4.1 Answered questions

The study presented in this thesis aimed at answering the following questions:

1. *What types of errors do L2 learners of English make in the production of English *wh*-questions?*
2. *Do participants produce more correct responses in main questions than in embedded questions?*
3. *Is there an effect of *wh*-word on inversion rates?*
4. *Is there an effect of verb type on inversion rates?*
5. *Are non-inversion errors in embedded questions due to L1 transfer or L2 interference and overgeneralization from main to embedded contexts?*

The focus of the study presented in this thesis was to analyze production errors, mainly errors of inversion. Overall, participants produced more non-inversion errors in embedded questions than inversion errors in main questions. In main questions, participants produced more double tense errors, presumably because they were copying the tensed verb from the prompt. These errors were virtually non-existent in embedded questions. Participants also produced non-target auxiliary errors somewhat frequently in both experiments and tended to omit phrases in their response. Both of these types of errors could be attributed to the task difficulty or to memory constraints. Few lack of morphology or omitted auxiliary errors were found in main questions. In embedded questions, more productions lacking morphology were found and one could hypothesize that, similar to double tense errors in main questions, participants were copying the untensed lexical verb from the main question prompt to their embedded response.

The data showed participants produced more errors in embedded *wh*-questions than in main *wh*-questions. This finding supports the hypothesis that L2 learners of English acquire and master the structure of main questions before the structure of embedded questions. The present study also found a *why*-asymmetry within the inversion rates in main questions and the non-inversion rates in embedded questions. *Why* is associated with lower rates of inversion in both structures. Additionally, adjuncts were associated with a higher rate of non-inverted responses in embedded questions. The intermediate and advanced L2 English participants also showed difficulties in producing correct, native-like *whose*-questions. *Whose* was added as a

wh-word to the study because of its particular structure and as hypothesized, *whose* was associated with low percent correct responses in both main and embedded questions.

The present study found no apparent effect of verb type on inversion rates. In main questions, inversion rates for lexical and auxiliary verbs were comparable. In embedded questions, participants incorrectly inverted more auxiliary verbs than lexical. However, it is not surprising that participants would be more likely to invert the auxiliary verb in embedded questions rather than applying *do*-support within lexical verb conditions.

Finally, the present study found inversion in embedded questions to be a result of L2 interference rather than L1 transfer. In embedded questions, participants produced English subject-auxiliary inversion and not Italian VP movement, showing that they had overgeneralized the rule of subject-auxiliary inversion from English main questions to English embedded questions.

4.2 Unanswered questions and opportunities for follow-up studies

The data presented in this thesis supports the hypotheses that main questions are acquired before embedded questions and that *why*-questions show low inversion rates in both contexts. The study, however, had its limitations. In order to better understand how L2 learners acquire English *wh*-questions, future studies including more participants with varying L1s would be helpful. Additionally, a range of proficiency levels would be beneficial in identifying clear trends of how and when certain structures are acquired. The initial pilot data showed that the immersive participants produced more correct responses in both main and embedded questions than the non-immersive participants. In order to better understand how living in an immersive environment affects acquisition, a larger sample of L2 participants living in an English-speaking country is needed. Furthermore, few studies have investigated the L2 acquisition of *whose*-questions. Gavrusseva and Thornton (2001) studied monolingual English-speaking children's acquisition of *whose*-questions while Gavrusseva (1997) studied two L1 Russian child L2 learners of English. The present study showed the L1 Italian learners produced different types of errors than the L1 English learners and L1 Russian learners in the previously mentioned studies. Additionally, they were shown to employ different learner strategies to avoid using an unacquired structure. Further research is needed to investigate why learners make these errors and when learners begin to correctly produce *whose*-questions. Additionally, more learners with varying L1s are needed to understand if L1 transfer plays a role in the production of *whose*-questions.

4.3 Implications for teaching English as a second language

The English Profile states that B2 level learners should have acquired and be able to produce main and embedded *wh*-questions. This study has shown that intermediate/advanced speakers make errors in both main and embedded *wh*-questions and highly advanced, bilingual speakers continue to make errors in embedded questions. Furthermore, both groups of participants showed difficulties producing *whose*-questions. *Whose* as a *wh*-word is not a focus in the foreign language classroom. In fact, the English Profile does not mention *whose* apart from its use as a relative pronoun (Cambridge ESOL & CUP, 2011). The ICLE corpus showed that *whose*-questions are produced infrequently by English speakers (Pozzan, 2011). However, because of their specific syntactic nature, there should be an explicit focus on teaching these structures to L2 English learners.

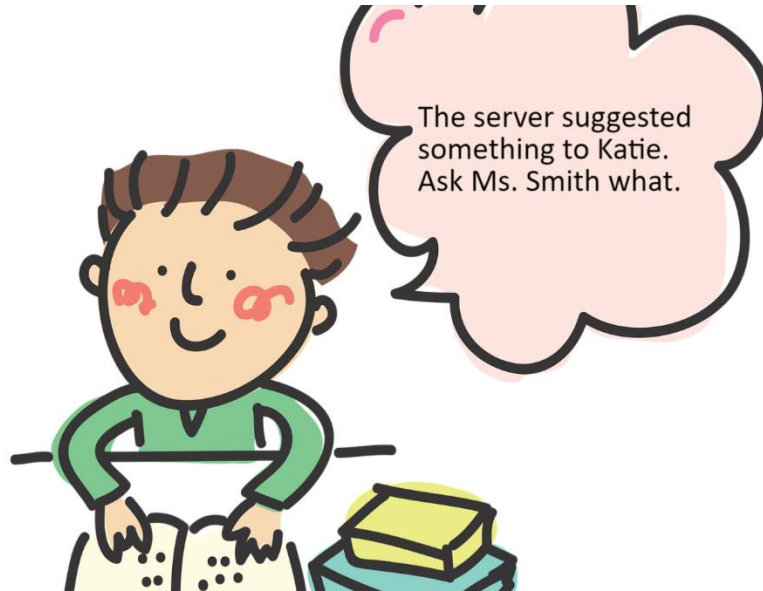
Highly advanced, bilingual English speakers with plenty of exposure to the input, were still shown to produce errors in embedded questions. As Spada & Lightbown (1999) found, L1 French learners who had not received explicit, grammar-focused, instruction continued to apply rules from their L1 to English main questions. This study found that errors of cancel inversion were due to L2 interference and overgeneralization of the rule of inversion from main questions to embedded questions. It seems that explicit instruction of English *wh*-embedded questions is needed, or learners will continue to apply the rules of main *wh*-questions in embedded contexts.

APPENDICES

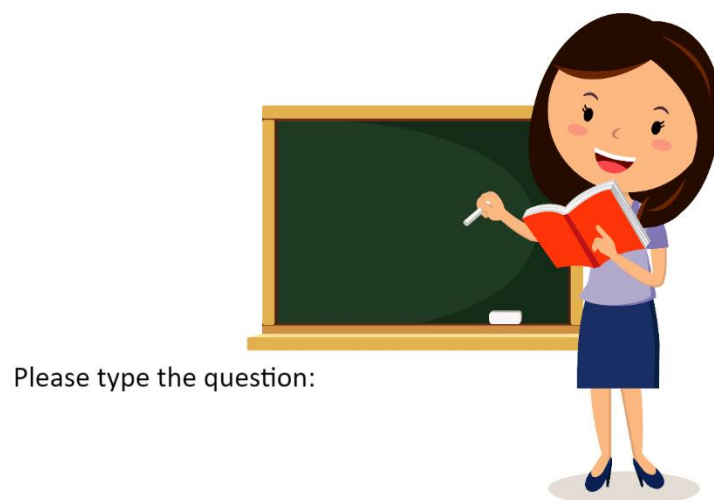
Appendix A

Procedure for main *wh*-questions

Screen 1



Screen 2



Screen 3



Experiment 1 demo

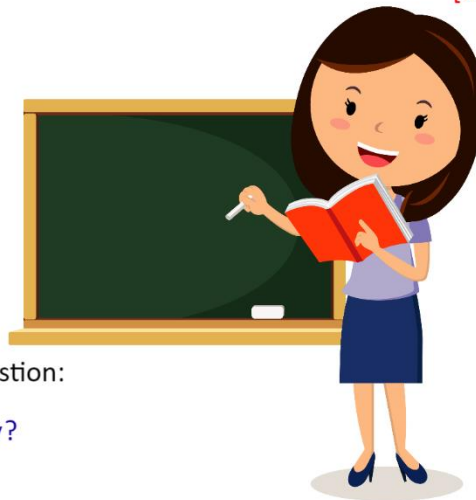
Screen 1

[Start a caption](#)



Screen 2

[demo]



Please type the question:

Does Sue work today?

Screen 3



Appendix B

Practice Items

Someone is yelling upstairs. Ask Ms. Smith who.
Someone usually plays the piano next door. Ask Ms. Smith who.
Somebody ran ten miles. Ask Ms. Smith who.
Maybe Ms. Smith's son is painting a picture. Ask her.
Maybe Ms. Smith's husband usually drives her home. Ask her.
Maybe Ms. Smith's husband has left for work. Ask her.

List 1

Ms. Smith's son is cooking something. Ask Ms. Smith what.
Jane always reads something in the morning. Ask Ms. Smith what.
The server suggested something to Katie. Ask Ms. Smith what.
Tom has taught his son to do something in the house. Ask Ms. Smith what.
Dan has gone somewhere new on vacation. Ask Ms. Smith where.
Ms. Smith's daughter hid her phone somewhere. Ask Ms. Smith where.
Sarah has invited them somewhere for her birthday. Ask Ms. Smith where.
Ms. Smith's sister usually travels somewhere in June. Ask Ms. Smith where.
Molly complimented someone after the meeting. Ask Ms. Smith who.
Olivia is driving someone to the mall. Ask Ms. Smith who.
Laura is meeting someone for a job interview. Ask Ms. Smith who.
Luke visits someone every Monday. Ask Ms. Smith who.
Bill is carrying someone else's book to class. Ask Ms. Smith whose.
Anna cleans someone else's house on Wednesdays. Ask Ms. Smith whose.
Emma has found someone else's phone in her car. Ask Ms. Smith whose.
Joe wore someone else's jacket to the party. Ask Ms. Smith whose.
Ms. Smith's brother buys a bag of apples every week. Ask Ms. Smith why.
Ms. Smith's husband called the store. Ask Ms. Smith why.
The boss has closed her office door. Ask Ms. Smith why.
Mark is complaining to the server. Ask Ms. Smith why.

List 2

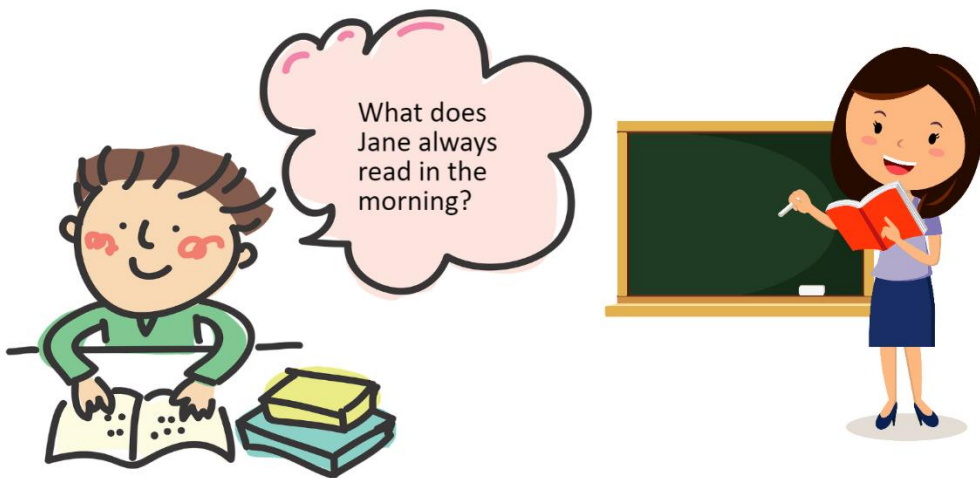
Ms. Smith's son cooked something for dinner. Ask Ms. Smith what.
Jane is reading something to the students. Ask Ms. Smith what.
The consultant has suggested something to Ben. Ask Ms. Smith what.
Tom teaches something at the university. Ask Ms. Smith what.
Dan goes somewhere every day at noon. Ask Ms. Smith where.
Ms. Smith's daughter is hiding somewhere. Ask Ms. Smith where.
Sarah invited Tash to play somewhere. Ask Ms. Smith where.
Ms. Smith's sister has traveled somewhere in Asia. Ask Ms. Smith where.
Molly is complimenting someone in the hallway. Ask Ms. Smith who.
Olivia drives someone to work on Mondays. Ask Ms. Smith who.
Laura met someone in France. Ask Ms. Smith who.
Luke has visited someone in the hospital. Ask Ms. Smith who.
Bill usually carries someone else's bag to the airport. Ask Ms. Smith whose.
Anna has cleaned someone else's shirt. Ask Ms. Smith whose.

The teacher found someone else's keys in her classroom. Ask Ms. Smith whose.
Joe is wearing someone else's baseball hat. Ask Ms. Smith whose.
Ms. Smith's brother is buying two bottles of wine. Ask Ms. Smith why.
Ms. Smith's husband has called Jim. Ask Ms. Smith why.
The teacher closed the window in the classroom. Ask Ms. Smith why.
Mark complains about his job a lot. Ask Ms. Smith why.

Appendix C

Procedure for embedded *wh*-questions

Screen 1



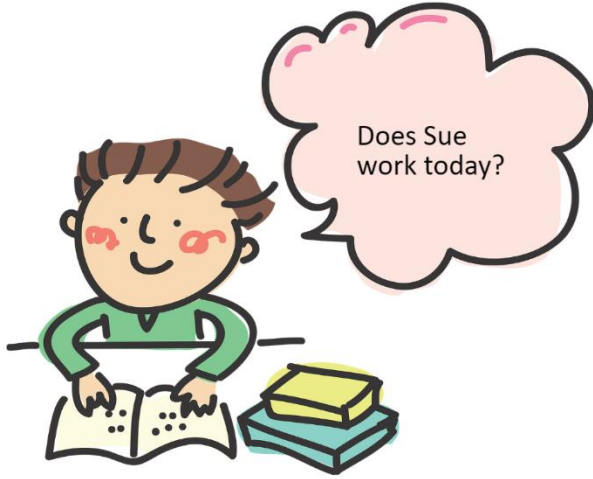
Screen 2



Please type your answer below:

Experiment 2 demo

Screen 1



[demo]



Screen 2

Start a session



[demo]

Please type your answer below:

Phil wants to know if
Sue works today.

Appendix D

Practice Items

Who is yelling upstairs?
Who usually plays the piano next door?
Who ran ten miles?
Is your son painting a picture?
Does your husband usually drive you home?
Has your husband left for work?

List 1

What did your son cook for dinner?
Why does Mark complain about his job a lot?
Where has your sister traveled in Asia?
Why has your husband called Jim?
Why did the teacher close the window in the classroom?
Who is Molly complimenting in the hallway?
Where does Dan go every day at noon?
What does Tom teach at the university?
Who does Olivia drive to work on Mondays?
Whose baseball hat is Joe wearing?
Where is your daughter hiding?
Why is your brother buying two bottles of wine?
Where did Sarah invite Tash to play?
Whose bag does Bill usually carry to the airport?
What has the consultant suggested to Ben?
Who has Luke visited in the hospital?
Whose keys did the teacher find in her classroom?
Whose shirt has Anna cleaned?
Who did Laura meet in France?
What is Jane reading to the students?

List 2

What has Tom taught his son to do in the house?
Whose book is Bill carrying to class?
Where did your daughter hide your phone?
Where has Sarah invited them for her birthday?
Why has the boss closed her office door?
Why does your brother buy a bag of apples every week?
Why did your husband call the store?
Who did Molly compliment after the meeting?
What did the server suggest to Katie?
Where does your sister usually travel in June?
Who does Luke visit every Monday?
Why is Mark complaining to the server?
Where has Dan gone on vacation?
Whose phone has Emma found in her car?

Whose jacket did Joe wear to the party?
What is your son cooking?
Who is Olivia driving to the mall?
Who is Laura meeting for a job interview?
What does Jane always read in the morning?
Whose house does Anna clean on Wednesdays?

Appendix E

Consent form

Gentile partecipante,

Il presente studio è condotto da Jessie Quinn, laureanda magistrale in Scienze del Linguaggio presso il Dipartimento di Studi Linguistici e Culturali Comparati dell'Università Ca' Foscari – Venezia sotto la supervisione di prof.ssa Giulia Bencini. Accettando questo modulo, esprime il suo consenso alla partecipazione allo studio e alle attività in esso incluse.

Lo studio è rivolto a soggetti maggiorenni che conoscono la lingua inglese. L'interesse principale è quello di indagare le strutture interrogative in inglese ed eventuali problematiche nel produrre correttamente queste strutture. Le verrà chiesto di completare due esperimenti della durata di circa 15 minuti in cui dovrebbe leggere delle domande in Inglese su argomenti semplici che non richiedono nessuna conoscenza specifica. In base alla domanda le verrà chiesto di formulare la risposta scritta pertinente. Le verranno mostrati esempi prima dell'inizio dello studio.

La partecipazione a questo studio è volontario e potrà decidere di abbandonare lo studio in qualsiasi momento senza alcun tipo di conseguenza negativa. Esprimendo il suo consenso, autorizzerà i ricercatori a conservare in formato digitale e a trattare in maniera confidenziale per tutta la durata del progetto di ricerca i dati personali acquisiti. Al fine di tutelare la privacy, tutti i dati raccolti non saranno mai riconducibili alla sua persona, secondo quanto previsto da Codice etico e di comportamento dell'Università Ca' Foscari – Venezia e dalla normativa nazionale vigente. I dati saranno trattati in forma deidentificata ai sensi del Regolamento (UE) 2016/679 e del D.Lgs. n. 196/2003 e saranno presentati in forma aggregata a livello di gruppo. Potrà chiedere di modificare, ritirare o eliminare il consenso alla partecipazione allo studio e tutti i dati forniti in qualsiasi momento contattando il/la responsabile della raccolta dati. I risultati delle analisi dei dati in forma aggregata e anonima potranno essere pubblicati sotto forma di tesi, libri o articoli per riviste scientifiche. Lo studio e i moduli che le viene chiesto di compilare hanno ricevuto l'approvazione della Commissione Etica di Ateneo in data 05.02.2020, verbale n. 1/2020 (per ulteriori informazioni: commissione.etica@unive.it).

Per qualsiasi domanda relativa alle procedure dello studio, ora o in futuro, può contattare:

- Supervisore della ricerca: prof.ssa Giulia Bencini, giulia.bencini@unive.it

- Ricercatrice/responsabile della raccolta dati: Jessie Quinn, 874173@stud.unive.it

- Eventuali altri recapiti: Staff BemboLab, bembolab@unive.it, 041/2345738 - 041/2345748

Il/La sottoscritto/a dichiara di aver letto con attenzione e compreso le informazioni contenute nel presente documento. Dichiara di esprimere il proprio consenso a partecipare allo studio qui descritto e autorizzare i ricercatori a trattare, gestire e archiviare tutti i dati personali con le modalità sopracitate. Il consenso potrà essere modificato/revocato in qualsiasi momento.

- Esprimo il consenso alla partecipazione allo studio (1)
- Non esprimo il consenso alla partecipazione allo studio (2)

Appendix F

Language Profile Questionnaire

We would like to ask you to help us by answering the following questions concerning your language history, use, attitudes, and proficiency. This survey was created to better understand the profiles of L2 learners of English. The survey consists of 30 questions and will take less than 10 minutes to complete. The Language Profile Questionnaire was created referencing the Bilingual Language Profile: English-French created by Birdsong and colleagues at the University of Texas at Austin and the Language History Questionnaire created by the Language Acquisition Research Center at Hunter College CUNY.

This is not a test, so there are no right or wrong answers. Please answer every question to the best of your ability. You will have an opportunity to clarify and explain any of your responses regarding questions that were unclear or difficult to answer. Thank you very much for your help.

I. Biographical Information

Name _____ Date _____

Age _____ Male / Female / Other

Current place of residence: city/state _____ country _____

Country of origin: _____

If your country of origin is different than your country of residence, when did you move to the country where you currently live? _____

Highest level of formal education (your current or most recent education level, even if you have not finished the degree).

- Graduate school (PhD/MD/JD)
- Graduate school (MA/MS)
- College (BA/BS)
- High School

- Middle School
- Other (specify): _____

II. Language history *In this section, we would like you to answer some factual questions about your language history.*

1. Please list all the languages you know in order of dominance. If you are equally dominant in two languages, please pick an order for them.

- 1)
- 2)
- 3)
- 4)

2. At what age did you start learning English?

Since birth 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20+

3. At what age did you start to feel comfortable using English?

As early as I 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20+ not yet | can't remember

4. How many years of English language classes have you had (pre-school through university)?

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20+

5. How many years of classes (grammar, history, math, etc.) have you had in English (pre-school through university)?

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20+

6. Please indicate the age at which you started using English in each of the following environments.

At home	With friends	At school	At work	Language learning software	Online games	Social media

7. If you have lived or travelled in countries where you used English for three or more months, please indicate the name of the country, the length of your stay, and how often you used English for each country, using the following scale.

Never	Rarely	Sometimes	Regularly	Often	Usually	Always
1	2	3	4	5	6	7

*You may have been to the country on multiple occasions, each for a different length of time. Add all the trips together.

Country	Length of stay* [month(s)]	Frequency of use (1 – 7)

8. How much time have you spent in a family or home environment where English was spoken?

Indicate months or years _____

9. How much time have you spent in a work or school environment where English is spoken?

Indicate months or years _____

III. Language use *In this section, we would like you to answer some questions about your language use.*

10. In an average week, how many hours do you use English with friends?

Indicate hours _____

11. In an average week, how many hours do you use English with family?

Indicate hours _____

12. In an average week, how many hours do you use English at school/work?

Indicate hours _____

13. Please estimate the number of hours per week that you are exposed to English.

Indicate hours_____

14. How often do you use English to speak to the following groups of people? Please enter the number in the table according to the scale below. *Include significant others in this category if you did not include them as family members (e.g., married partners). **Include anyone in the work environment in this category (e.g., if you are a teacher, include students as coworkers).

Never Rarely Sometimes Regularly Often Usually Always

1 2 3 4 5 6 7

	Scale 1 - 7
Family members	
Friends*	
Classmates and Coworkers**	
People on the Internet	

15. How often do you use English for the following activities? Please enter the number in the table according to the scale below.

Never Rarely Sometimes Regularly Often Usually Always

1 2 3 4 5 6 7

*This includes shouting, cursing, showing affection, etc.

**This includes counting, calculating tips, etc.

***This includes telephone numbers, ID numbers, etc.

	Scale 1 - 7
Thinking	
Talking to yourself	
Expressing Emotion*	
Dreaming	
Arithmetic**	
Remembering numbers***	

16. How often do you spend engaged in the following activities in English?

Never Rarely Sometimes Regularly Often Usually Always

1 2 3 4 5 6 7

	Scale 1 - 7
Watching television	
Listening to radio or podcasts	
Reading for fun	
Reading for school/work	
Writing emails	
Writing for school/work	
Listening to music	

IV. Language proficiency *In this section, we would like you to rate your language proficiency.*

1=not well at all 7=very well

17. a. How well do you speak English? 1 2 3 4 5 6 7

18. a. How well do you understand English? 1 2 3 4 5 6 7

19. a. How well do you read English? 1 2 3 4 5 6 7

20. a. How well do you write English? 1 2 3 4 5 6 7

21. If you have taken any standardized language proficiency tests (e.g., TOEFL, IELTS, PET), please write the name of each test and the score you received. If you do not remember the exact score, then indicate an "Approximate score" instead.

Test	Score	(Approximate score)	Date Taken

22. Using the CEFR, what would you self-rate your level of English, whether or not you have a certification?

A1 A2 B1 B2 C1 C2

V. Language attitudes *In this section, we would like you to respond to statements about language attitudes.*

1=disagree 7=agree

- 23. I feel like myself when I speak English. 1 2 3 4 5 6 7
- 24. I identify with an English-speaking culture. 1 2 3 4 5 6 7
- 25. It is important to me to use (or eventually use) English like a native speaker. 1 2 3 4 5 6 7
- 26. I want others to think I am a native speaker of English. 1 2 3 4 5 6 7
- 27. Please enter the language you feel the most comfortable in when listening, speaking, reading, and writing in each of the contexts listed below.

	Listening	Speaking	Reading	Writing
At home				
With friends				
At school				
At work				
On the Internet				
On social media				

28. Please rate your language learning skill. In other words, how good do you feel you are at learning new languages, relative to your friends or other people you know? Pick one.

Very Poor	Poor	Limited	Average	Good	Very Good
1	2	3	4	5	6

29. Please comment below to indicate any additional answers to any of the questions above that you feel better describe your language background or usage.

30. Please comment below to provide any other information about your language use.

Appendix G

Australian and British English native speakers' average ratings of whose marked structures

<u>Sentence</u>	<u>Average rating on a scale from 1 – 5</u>
Whose is the hat that Joe is wearing?	4.00
Whose is the hat?	4.00
Whose are the keys that you found in the classroom?	3.90
Whose is the shirt that Anna cleaned?	3.88
Whose is the t-shirt Anna has cleaned?	3.25
Whose is the phone that Emma has found in her car?	3.13
Whose jacket was the one Kevin took to the party?	3.10
Whose is the bag carried by Will?	2.11
Chris wants to know whose is the shirt anna cleaned.	2.89
Chris wants to know whose were the keys that the teacher found.	2.75
Chris wants to know whose is the house that Anna cleans on Wednesday.	2.63
Chris wants to know whose are the keys that the teacher found.	2.50
Chris wants to know whose is the book that Chris is carrying to the class	2.38
Chris wants to know whose is the book Chris is carrying to the classroom.	2.22
Chris wants to know whose is the bag that Billy carries to the airport.	2.11
Chris wants to know who is the key the teacher found in the classroom.	1.55

REFERENCES

- Ambridge, B., Rowland, C., Theakston, A., & Tomasello, M. (2006). Comparing different accounts of inversion errors in children's non-subject wh-questions: What experimental data can tell us? *Journal of Child Language*, *33*, 519–557.
- Bencini, G., & Valian, V.V. (2008). Abstract sentence representations in 3-year-olds: Evidence from language production and comprehension. *Journal of Memory and Language*, *59*(1), 97–113. <https://doi.org/10.1016/j.jml.2007.12.007>
- Berk, S. (2003). Why “why” is different. In B. Beachley, A. Brown, & F. Conlin (Eds.), *Proceedings of BUCLD 27* (127–137). Somerville, MA: Cascadilla Press.
- Bianchi, V., Bocci, G., & Cruschina, S. (2017). Two types of subject inversion in Italian wh-questions. *Revue Roumaine De Linguistique*, *LXII*(3), 233–252.
- Bley-Vroman, R. (1989). The Logical Problem of Foreign Language Learning. *Linguistic Analysis*, *20*, 3–49.
- Bloom, L., Merkin, S., & Wootten, J. (1982). Wh-questions: Linguistic factors that contribute to the sequence of acquisition. *Child Development*, *53*, 1084–1092.
- Bocci, G., & Pozzan, L. (2014). Some questions (and some experimental answers) about Italian subjects: Subject positions in main and embedded questions in L1 and attrition. In C. Contemore & L. D. Pozzo (Eds.), *Inquiries into Linguistic Theory and Language Acquisition. Papers Offered to Adriana Belletti* (28–44). Siena: CISCL Press.
- Calabrese, A. (1982). Alcune ipotesi sulla struttura informazionale della frase in italiano e sul suo rapporto con la struttura fonologica. *Rivista di Grammatica Generativa*, *13*, 489–526.
- Cambridge ESOL, & Cambridge University Press. (2011). *English Profile: Introducing the CEFR for English*. Cambridge: Cambridge University Press.

- Chomsky, N. (1965). *Aspects of the theory of Syntax*. Cambridge, MA:MIT Press.
- Chomsky, N. (1986). *Knowledge of language its nature, origin, and use*. New York: Praeger.
- Conroy, A., & Lidz, J. (2007). Production/comprehension asymmetry in children's why questions. In A. Belikova, L. Meroni, & M. Umeda (Eds.), *Proceedings of GALANA 2007* (73–83). Somerville, MA: Cascadilla Press.
- Epstein, S.D., Flynn, S., & Martohardjono, G. (1996). Second language acquisition: Theoretical and experimental issues in contemporary research. *Behavioral and Brain Sciences, 1*, 677-758.
- Fisher, C., Gertner, Y., Scott, R. M., & Yuan, S. (2010). Syntactic bootstrapping. *Wiley Interdisciplinary Reviews: Cognitive Science, 1*(2), 143–149.
<https://doi.org/10.1002/wcs.17>
- Gavruseva, E. (1997). *The Syntax of Possessive Phrases and Acquisition of Whose-Questions in Child LI and L2 English* [Doctoral dissertation, Georgetown University].
- Gavruseva, E., & Thornton, R. (2001). Getting It Right: Acquisition of Whose-Questions in Child English. *Language Acquisition, 9*(3), 229–267.
https://doi.org/10.1207/s15327817la0903_2
- González Fernández, C. (2015). *The syntax of embedded questions and its acquisition by L2 learners of English* [master's thesis, Universitat Autònoma de Barcelona].
- Guasti, M. T. (1996). Acquisition of Italian interrogatives. In H. Clahsen (Ed.), *Generative perspectives on language acquisition* (241–270). Amsterdam/Philadelphia: John Benjamins.
- Håkansson, G., & Pienemann, M. (2002). Transfer and typological proximity in the context of second language processing. *Second Language Research, 18*(3), 250–273.
<https://doi.org/10.1191/0267658302sr206oa>

- Hopp, H. (2017). The processing of English which-questions in adult L2 learners: Effects of L1 transfer and proficiency. *Zeitschrift Fur Sprachwissenschaft*, 36(1), 107–134.
<https://doi.org/10.1515/zfs-2017-0006>
- Kleinmann, H. (1977). Avoidance behavior in adult second language acquisition. *Language Learning*, 27, 93–107.
- Lee, S.-Y. (2008). Argument-Adjunct Asymmetry in the Acquisition of Inversion in Wh-Questions by Korean Learners of English. *Language Learning*, 58, 625–663.
<https://doi.org/10.1111/j.1467-9922.2008.00452.x>
- Oller, J., & Ziahosseiny, S. (1970). The contrastive analysis hypothesis and spelling errors. *Language Learning* 20, 183-189.
- Park, H. (2000). When-questions in second language acquisition. *Second Language Research*, 16, 44-76. <https://doi.org/10.1191/026765800666268444>
- Pienemann, M. (1998). Developmental dynamics in L1 and L2 acquisition: Processability Theory and generative entrenchment. *Bilingualism: Language and Cognition*, 1(1), 1–20. <https://doi.org/10.1017/s1366728998000017>
- Pienemann, M. (2015). An Outline of Processability Theory and Its Relationship to Other Approaches to SLA. *Language Learning*, 65(1), 123–151.
<https://doi.org/10.1111/lang.12095>
- Pienemann, M., & Keßler, J.-U. (Eds.). (2011). *Studying processability theory: Introductory textbook*. Amsterdam: John Benjamins.
- Pozzan, L. (2011). *Asking Questions In Learner English: First and Second Language Acquisition of Main and Embedded Interrogative Structures* [Doctoral dissertation, The City University of New York].

- Pozzan, L., & Quirk, E. (2014). Second language acquisition of English questions: An elicited production study. *Applied Psycholinguistics*, 35(6), 1055–1086.
<https://doi.org/10.1017/S0142716412000690>
- Pozzan, L., & Valian, V.V. (2017). Asking questions in child English: Evidence for early abstract representations. *Language Acquisition*, 24(3), 209–233.
<https://doi.org/10.1080/10489223.2016.1187615>
- Ringbom, H. (1987). *The role of the first language in foreign language learning*. Clevedon: Multilingual Matters.
- Rizzi, L. (1996). Residual verb second and the Wh-Criterion. In A. Belletti & L. Rizzi (Eds.), *Parameters and Functional Heads* (63–90). Oxford/New York: Oxford University Press.
- Rizzi, L. (2001). On the position “int(errogative)” in the left periphery of the clause. In G. Cinque & G. Salvi (Eds.), *Current studies in Italian syntax: Essays offered to Lorenzo Renzi* (287–296). Amsterdam: Elsevier.
- Schacter, J. (1990). On the issue of completeness in second language acquisition. *Second Language Research*, 6, 93–124.
- Schwartz, B. D., & Sprouse, R. A. (1994). Word Order and Nominative Case in Non-native Language Acquisition: A longitudinal study of (L1 Turkish) German interlanguage. In T. Hoekstra & B.D. Schwartz (Eds.), *Language acquisition studies in generative grammar: Papers in honor of Kenneth Wexler from the 1991 GLOW workshops* (317–368). Amsterdam: Benjamins. <https://doi.org/10.1075/lald.8.14sch>
- Schwartz, B. D., & Sprouse, R. A. (1996). L2 cognitive states and the Full Transfer/Full Access model. *Second Language Research*, 12(1), 40–72.
<https://doi.org/10.1177/026765839601200103>

- Spada, N., & Lightbown, P. M. (1999). Instruction, first language influence, and developmental readiness in second language acquisition. *The Modern Language Journal*, 83(1), 1–22. <https://doi.org/10.1111/0026-7902.00002>
- Stromswold, K. (1990). *Learnability and the acquisition of auxiliaries* [Doctoral dissertation, MIT].
- Thornton, R. (2008). Why continuity. *Natural Language & Linguistic Theory*, 26, 107–146. <https://doi.org/10.1007/s11049-007-9031-z>
- Vainikka, A., & Young-Scholten, M. (1994). Direct access to X'-Theory: evidence from Korean and Turkish adults learning German. In T. Hoekstra & B.D. Schwartz (Eds.), *Language acquisition studies in generative grammar* (265–316). Amsterdam: John Benjamins. <https://doi.org/10.1075/lald.8.13vai>
- Vainikka, A., & Young-Scholten, M. (1996). The early stages in adult L2 syntax: additional evidence from Romance speakers. *Second Language Research*, 12(2), 140–176. <https://doi.org/10.1177/026765839601200202>
- Valian, V.V., & Casey, L. (2003). Young children's acquisition of *wh*-questions: The role of structured input. *Journal of Child Language*, 30(1), 117–143. <https://doi.org/10.1017/S0305000902005457>
- White, L. (2003). *Second language acquisition and Universal Grammar* (Cambridge Textbooks in Linguistics). Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511815065>