Universidad de Valladolid

FACULTAD de FILOSOFÍA Y LETRAS
DEPARTAMENTO de FILOLOGÍA INGLESA
Grado en Estudios Ingleses

TRABAJO DE FIN DE GRADO

A GRAMMATICAL STUDY ON HOW TYPICALLY DEVELOPED CHILDREN AND CHILDREN WITH SPECIFIC LANGUAGE IMPAIRMENT PRODUCE *WHAT* AND *WHEN*

Mónica Montón Catalina

Tutora: Raquel Fernández Fuertes

2017-2018

ABSTRACT

This undergraduate dissertation presents a grammatical study on the production of *what* and *when* by English native speakers. It presents a double comparison regarding two child groups that have been MLU-matched (i.e. typically developed children (TD) and children with specific impairment (SLI)) and two data types (i.e. narrative and spontaneous). The analysis deals with the production of *what* and *when* both as isolated words and as part of *wh*-questions and *wh*-relative clauses. The results of the study reveal that the production of *when* children is higher than that of *what* in the case of the TD children, as opposed to the SLI children who present more instances of *what*. Likewise, a correlation may be set in regards to the use of these *wh*-words in the two target structures: the TD children produce more *wh*-relative clauses, whereas the SLI children produce more *wh*-questions.

KEYWORDS: *what*, *when*, *wh*-questions, *wh*-relative clauses, typically developed children, children with specific language impairment.

RESUMEN

Este trabajo ofrece un estudio gramatical sobre la producción de *what* y *when* en datos de niños ingleses nativos. Presenta una doble comparación en relación a dos grupos de niños que han sido agrupados teniendo en cuenta el criterio LME (i.e. niños sin trastornos asociados al lenguaje, niños con un trastorno específico del lenguaje (TEL)) y a dos tipos de datos (i.e. de narrativa y espontáneos). El análisis realizado en el presente estudio tiene que ver con la producción de las palabras *what* y *when* tanto cuando se usan por separado como cuando forman parte de oraciones *wh*- y oraciones de relativo. Los resultados de este estudio muestran que los niños que se desarrollan con normalidad producen más casos de *when* que de *what,* frente a los niños con TEL que presentan más ejemplos de *what.* Así mismo, se puede establecer una relación entre el uso de estas palabras y las estructuras gramaticales en cuestión: los sin trastornos asociados al lenguaje producen más oraciones de relativo, mientras que los niños con TEL presentan más casos de oraciones *wh*-.

PALABRAS CLAVE: *what*, *when,* oraciones *wh*-, oraciones relativas de *wh*-, niños sin trastornos específicos del lenguaje, niños con trastornos asociados al lenguaje.

CONTENTS

FOREWORD: CONTEXTUALIZATION OF THE UNDERGRADUATE DISSERTATION

This undergraduate dissertation is the last formal requirement to complete the degree in English Studies at the University of Valladolid. Among all the contents of the degree, this dissertation falls under the A2 category "Scientific Description of the English Language", according to the current description of the degree (Universidad de Valladolid 2017-2018).

This dissertation deals with the study of two grammatical units (i.e. *what* and *when*). I have selected this topic since it is a great opportunity to work on a different area of grammar and undertake a complete research study. Moreover, I find these two *wh*-words as an appealing topic as they are highly frequent and consequently, research on how English speakers produce them really attracted me. Therefore, thanks to this study I will have an extensive knowledge of these grammatical units.

Furthermore, working on this area will provide me with different competences and knowledge about this topic and related grammatical issues. Firstly, I will be more familiar with different properties dealing with *wh*-words (i.e. functions, uses, *wh*-movement), and more specifically with those of *what* and *when*. Secondly, by doing an empirical analysis I will be able to work deeply with data (i.e. extracting, classifying, analyzing, etc.) and different software packages (i.e. CLAN). Finally, this dissertation will help me to investigate on the topic and to establish my own research questions.

Besides, in this undergraduate dissertation I have applied general and specific competences that I have acquired over these past four years. Therefore, the particular competences I have used to work on and to write this dissertation are described below in regards to what the current description of the degree establishes.

The first thing I had to do was to search for related and relevant information on the target topic. Hence a bibliographical investigation has been carried out by looking up different available resources (i.e. UVa library books, on-line articles, internet sources, etc.). Consequently, the following competences have been reinforced:

- Ability to analyze and synthesize, conceptualize and abstract.
- Ability to manage information.
- Comprehension of knowledge related to the structural and systematic properties of the English language.
- Fluency in the use of common means and technological resources.
- Capacity to understand and show the acquired knowledge.

Furthermore, apart from these competences by doing my own research analysis, which is presented in this dissertation, I have been able to develop some specific competences which include the following:

- Personal and autonomous learning.
- Creativity.
- Aptitude to solve problems.
- Maturity, discipline, and intellectual, academic and expressive thoroughness.

In addition, since this dissertation presents a grammatical study on the production of two *wh*-words, and not only on a grammatical description, several competences related to this field have been applied, too. These include the following:

- Ability to comprehend the English grammar and its description.
- Ability to control the English language in a formal and academic register, both oral and written.
- Capacity to use specific means and technological resources (i.e. CHILDES, CLAN).

1. INTRODUCTION

The acquisition and production of a first language have been widely studied by many linguists and other researchers that have been concerned with the different language fields (i.e. vocabulary, semantics, syntax, etc.). In order to acquire their first language, children go through different stages along the acquisition process in which different structures (e.g. grammatical and ungrammatical, simple and complex) are used in accordance to their ages and MLU (Mean Length of Utterance) values. The way children develop through the different chronological stages (age) and developmental stages (MLU) is affected by the properties of the different grammatical units that make up a language. For instance, in the case of a specific type of grammatical unit, *wh*-words, different issues could be taken into account. The production of *wh*-words occurs gradually since some of them (i.e. *who, what,* or *where*) are easier to acquire than others (i.e. *when, which,* or *whose*) (e.g. Bloom et *al.* 1982, Rowland et *al.* 2003). Furthermore, while children are acquiring *wh*-words, they need to know how to use them properly in the different grammatical structures in which they may appear (i.e. *wh*-questions or *wh*-relative clauses). Thus, once they have acquired these properties, something that typically occurs at the age of 5, children's production is primarily adult-like. However, non-adult like forms may persist for a little bit longer which makes the analysis of children's production after the initial stages of acquisition an interesting field of study.

Hence, the current dissertation presents an empirical study based on the production of two *wh*-words*, what* and *when,* by typically developed (TD) children and by children with specific language impairment (SLI), all of them English native speakers. Three issues are targeted in this respect and in relation to these *wh*-words. Firstly, the production of *what* and *when* is considered, separately. Secondly, the distribution of these *wh*-words (*what* and *when*) across sentence types is analyzed, that is, whether these are used in *wh*-relative or in *wh*-questions. And thirdly, a parallel comparison according to the type of data (i.e. spontaneous or narrative) is done in regards to each aspect under investigation. Therefore, the target of this study is to analyze whether the production of *wh*-words, in relation to

different grammatical issues (i.e. *wh*-word type, sentence type, and data type), is similar for TD and SLI children when they are MLU-matched.

This dissertation is divided into seven main sections which are in turn subdivided into different subsections. This first section, the introduction, presents a brief explanation and contextualization of the topic of the study. The second section deals with a brief grammatical review of *wh*-words regarding their main grammatical properties. These include the definition and classification of *wh*-words, the uses and clause functions that the target *wh*-words (i.e. *what* and *when*) could have in a sentence (i.e. interrogative clause marker and relativizer), the types of clauses in which these *wh*-words can appear (i.e. *wh*-questions and *wh*-relative clauses), and the type of movement that is required (i.e. *wh*-movement). In the third section, some previous empirical studies dealing with the target topic are presented in order to contextualize the one that is offered in this dissertation and that is presented in the following section. Afterwards, the fourth section introduces the main research questions and objectives of the present study. The next section contains an explanation of the methodology followed in order to obtain all the necessary data to carry out this empirical study. Likewise, this section is divided into three main subsections: firstly, how the data have been selected is explained; secondly, the procedure used to extract the data is presented; and finally, how the data have been classified is illustrated. Furthermore, in the sixth section the results of this empirical study, in relation to the research questions, are presented and explained. Finally, the last section presents the conclusions reached after having analyzed the data. Additionally, a bibliography is included, where all the sources used in order to write this dissertation are included.

Besides, in order to carry out this study, a corpus has been compiled in order to classify and later on analyze all the data obtained. This has been done on an excel spreadsheet, which is attached to this dissertation in an electronic format.

2. LITERATURE REVIEW

In any language, "the essence of grammatical units is that they are meaningful and combine with each other in systematic ways" (Biber et *al.* 1999:50). According to Biber et *al.* (1999), these grammatical units can be divided into seven major subgroups: discourse, sentence, clause, phrase, word, morpheme, and phoneme/grapheme. Amongst these subgroups, this dissertation is concentrated on a specific word type, and on the sort of clauses in which this word type may appear: *wh*-words and *wh*-clauses. A theoretical description of these grammatical units is provided with a focus on the two that constitute the target of this work: *what* and *when*.

Therefore, this section is divided into three main subsections. Firstly, what *wh*-words are and how they are classified is explained. Next a specific analysis of uses and clause functions of *what* and *when,* is carried out. This leads to the next subsection, an account of the types of syntactic clauses in which these two *wh*-words may appear. Finally, as it is explained later, the fact that *wh*-words are used in *wh*-questions, implies the necessity of having a *wh*-movement. Therefore, the last part of this section outlines what *wh*-movement is and the bases of this phenomenon.

2.1. *Wh*-words
2.1.1. *Wh*-words: definition and classification

According to Biber et *al.* (1999:55), "to the ordinary language user, words are the basic elements of language". Words can be divided into three major categories: lexical words, functional words, and inserts. *Wh*-words are functional words whose main function is to introduce either *wh*-relative clauses, as in (1), or *wh*- interrogative clauses, as in (2).

> (1) *[He warned the public not to approach the men, [**who** are armed and dangerous]*$_{CP2}$*]*$_{CP1}$ (Biber et *al.*, 1999:195)

(2) *[**Which** one do you mean?]*CP1   (Biber et *al.*, 1999:87)

This subgroup of words is composed of nine *wh-words* and their corresponding compounds: "*Who, whom, which, whose, what, where, when, why, and how, or their compound in -ever: whoever, whatever, etc*." (Huddleston, 1984:366). The following examples, from (3) to (11), show an instance of each of them, either as part of an interrogative or a *wh*-relative clause.

(3) *[**Who** are you talking about?]*CP1

(4) *[For **whom** would I be working?]* CP1

(5) *[**Which** photos are we going to look at?]* CP1

(6) *[**Whose** turn is it tonight?]*CP1

(7) *[**What** are they doing?]* CP1

(8) *[I could lead you [to the shop **where** I bought it]CP2]*CP1

(9) *[I can't think of a think [**when** I would be going by myself]CP2]* CP1

(10)*[There's no reason [**why** you shouldn't go out for a drink with him] CP2] CP1*

(11) *[**How** was your trip, Nick?]* CP1

(Biber et *al.*, 1999: 204, 626)

As shown in these examples, *wh*-words can be used in two different ways, as interrogative markers, as in examples (3), (4), (5), (6), (7), and (11); and as relativizers, as in (8), (9), and (10).

Furthermore, Biber et *al*. (1999) discuss that these *wh*-words have an elicit function and that they are employed in order to provide further information which is absent. In other words, the receiver is waiting for specific linguistic information, which differs depending on the *wh*-word used. Consequently, the use of one *wh*-word or another depends on the required, or needed linguistic information (e.g. person as in (3), (4), and (6); location as in (8), time as in (9)).

2.1.2. <u>Uses and clause functions of *when* and *what*</u>

These grammatical units, can be used in different sentence types (e.g. *wh*-questions, as in (12); and *wh*-relative clause, as in (13)) and can have different syntactic roles, accordingly (e.g. interrogative clause marker, as in (12); and relativizer, as in (13)).

(12) *[**When** is Mary leaving home?]*~CP1~
(13) *[The day [**when** she arrived]*~CP2~ *you were not in the city]*~CP1~

In example (12), the question has been formulated in order to know the time when Mary is leaving home and, therefore, to complete a piece of information that is missing (i.e. time). Likewise, in example (13), the relative pronoun has been used as a link between the subordinate and the main clause and as a way of providing the required information. Even if *when* is used to supply information about time in both examples, they differ in an important way: *wh*-words, in general, and *when,* in the examples above, can be used either as interrogative clause markers, as in (12), or as relativizers, as in (13).

2.1.2.1. <u>*When* and *what* as interrogative clause markers</u>

According to Biber et *al.* (1999:87): "interrogative clause markers are used as pronouns (*who, whom, what, which*), determiners (*wha*t*, which, whose*), or adverbs (*how, when, where*, *why*)". Taken this statement as the starting point, and already focusing on the target *wh*-words (e.g. *when* and *what)*, their different uses are illustrated in (14) to (16).

(14) *[**When** are you leaving?]* ~CP1~
(15) *[**What** do they want?]*~CP1~   (Biber et *al*., 1999:87)
(16) *[**What** schoolchild's imagination could fail to be stimulated by such a challenge?]* ~CP1~   (Huddleston, 1984:369)

As example (14) shows, a *wh*-adverb, *when*, has been used in order to introduce a *wh*-question, an adverbial one. Therefore, as it can be inferred, *when* is used as a time adverb and, as it is explained later on, it functions as an adjunct of time. Hence, it indicates time and has the grammatical feature [+time].

On the other hand, in example (15), *what* has been used as a *wh*-pronoun and is, therefore, the lexical head of the DP. In this sense, *what* is characterized by having a [-human] grammatical feature (Huddleston, 1984). Besides, as example (16) shows, *what* can be used as a determiner, too. In those cases, the *wh*-word could be the determiner accompanying a [+/- human] noun: *what* could refer to a [+human] noun and consequently would have a [+ human] feature, as in example (17); or it could refer to a [-human] noun, having a [-human] feature, as reflected in (18).

(17) *[**What candidate** will you vote for?]$_{CP1}$*

(18) *[**What party** are you in favour of?]$_{CP1}$*   (Quirk et *al.,*1985:370)

2.1.2.2. *When* and *what* as relativizers

As previously explained, *wh*-words can be either used as interrogative clause markers or as relativizers. In regard to what Biber et *al.* (1999:87) explain: "relativizers are used as pronouns (*who, whom, which, that*), determiners (*which, whose*), or adverbs (*when, where, why*)". In (19) an instance of one of the two target *wh*-words is shown.

(19) *[That summer marked the time [**when** their carefree childhood really ended] $_{CP1}]_{CP2}$*   (Biber et *al*., 1999:628)

Regarding *when* as a relativizer, its principal function is to serve as a link between the main clause (complementizer phrase one (CP$_1$) and the subordinate clause (complementizer phrase two (CP$_2$)). In this way, *when* is used to relate both clauses by means of time reference, thus introducing a temporal relationship, as in example (20).

(20) *[He was born in another age, the age [**when** we played not for a million dollars in prize money.*]$_{CP2}$]$_{CP1}$ (Biber *et al.*, 1999:608)

As it can be appreciated from example (20), *when* refers to the age *when we played not for a million dollars in prize money*. Thus, the *wh*-relative clause (CP$_2$) introduces an adverbial clause which functions as an adjunct and has *the age* as its antecedent.

Moreover, when these grammatical units introduce *wh*-relative clauses, as in (20), the *wh*-word, apart from its linking function as complementizer, has a grammatical function within the CP$_2$, as addressed in the next subsection. The *wh*-word in these relative clauses can be a relative adverb, too, as in (21).

(21) *But the bit [**when** he's finished that]*   (Biber *et al.*, 1999:628)

The other target *wh*-word, *what*, is not included under this classification, due to its being barely employed as a relativizer. Nonetheless, it can appear as such in conversational texts, as in example (22). As this use is less frequent and restricted to informal conversational contexts, it is, therefore, not analyzed under the classification of standard relativizers (Biber *et al.*, 1999).

(22) *Gotta make sure she's got the book [**what** I had last week]* (CONV).
(Biber *et al.*, 1999: 609)

2.1.3. Syntactic functions of *when* and *what*

Wh-words, as any other type of word, play a syntactic function within the structure they belong to, that is, within the clause. In this respect, the previous distinction of these *wh*-words as pronouns, or adverbs, is a determining factor.

On the one hand, the principal function of *when* is to be a time adjunct in the clause, as reflected in example (23).

(23) *[**When** did you see Mark?]*CP1   (Biber et al. 1999: 204)

On the other hand, according to Huddleston (1984), when *what* is used as a pronoun, it can function either as the subject of the clause, as shown in example (24), or as the direct object, as in (25). When it is used as a determiner, it functions as part of a DP, as shown in (16) above, and in this case, the DP can function either as subject or object of the clause.

(24) *[**What** happened?]*CP1   (Huddleston, 1984:369)
(25) *[**What** kind of novels do you enjoy reading?]*CP1

(Quirk et *al.*, 1985: 822)

2.2. <u>Types of *wh*-clauses</u>

Wh-words as explained in subsection 1.1 can be used as interrogative clause markers or as relativizers. Hence, *wh*-clauses could be either direct interrogative clauses (independent clauses), as shown in example (26), or *wh*-relative clauses (dependent clauses), as in (27).

(26) *[**When** are you leaving?]*CP1   (Biber et *al.*, 1999:87)
(27) *[It occurs at a time [**when** abolitionist leaders hoped for improved treatment of slaves]*CP2*]*CP1.   (Biber et *al.*, 1999:268)

Therefore, firstly *wh*-clauses as independent clauses or main clauses (i.e. *wh*-questions) are discussed and secondly, *wh*-clause as dependent or subordinate clauses (i.e. *wh*-relative clauses).

2.2.1. _Wh_-clauses as independent or main clauses

Independent clauses, as the term suggests, do not rely on another larger clause, as Biber et _al._ (1999) indicate. What this means in grammatical terms is that these clauses function as CP_1, as main clauses, as shown in (25) above. Independent interrogative clauses can be classified in three different subgroups: _wh_-question, as in example (28); yes/no questions, as in (29); and alternative questions, as in (30). Nevertheless, the target of this dissertation is to study _wh_-questions, therefore only this sort of independent clauses is studied.

(28) ***What*** *do they want?*	[+Q] [+WH]
(29) *Do you think he'll be any better?*	[+Q]
(30) *Do you know one or two?*	[+Q]

<div align="right">(Biber et al., 1999:205-207)</div>

W_h_-main clauses are grammatically characterized by having a [+Q] and a [+WH] feature. The [+Q] feature reflects that it is a question and not, for instance, a statement (31 _versus_ 32). The [+WH] feature shows that it is a clause that contains a _wh_-word (31 _versus_ 33).

(31) *Shall we go by bus or train?*	[+Q] [-WH]
(32) *He was driving on one side of the road.*	[-Q] [-WH]
(33) ***What*** *side of the road was he driving on?*	[+Q] [+WH]

<div align="right">(Quirk et al., 1985:821-823)</div>

Thus, (33) is an independent _wh_-question which has two grammatical features [+Q] and [+WH]. The former indicates that this clause is a question, and the later represents that it is a _wh_-question as it is introduced by a _wh_-word: _what_. Besides, according to Quirk et _al_. (1985:817), _wh_-questions, in order to be grammatically correct, are required to follow two syntactic criteria:

- "The *wh*-element (i.e. the clause element containing the *wh*-word) comes first in the sentence".
- "The *wh*-word itself takes the first position in the *wh*-element".

In example (18), [*what party*]$_{DP}$ *are you in favor of?,* these two criteria are clearly exemplified. In this case, it can be seen that the *wh*-element, *what party*, comes first in the clause, and within that DP, the *wh*-word, *what,* occupies the first place, the specifier position.

2.2.2. *Wh*-clauses as dependent or subordinate clauses

Wh-clauses are dependent or subordinate if they are embedded in a superior structure (Biber et *al.*, 1999:192). In grammatical terms, these clauses are CPs$_2$ that depend on a higher CP, normally a CP$_1$, regardless of the type of dependency that is established (i.e. direct object of CP$_1$, adjunct of CP$_1$, etc.). There is a wide variety of clauses belonging to this group of subordinate sentences and, an instance of these are *wh*-relative clauses, as shown in example (34).

(34) *[He has born in another age, the age [**when** we played not for a million dollars in prize money]*$_{CP2}$*]*$_{CP1}$ (Biber et *al.*, 1999:608)

Wh-clauses as dependent or subordinate clauses are grammatically characterized by having a [+WH] feature which indicates that it is a relative clause introduced by a *wh*-word and not, for instance, by the zero relativizer (35 *versus* 36).

(35) *[I can't think of a time [**when** I would be going by myself]*$_{CP2}$*]*$_{CP1}$
(36) *[It's time [they paid the money back]*$_{CP2}$*]*$_{CP1}$ (Biber et *al.*, 1999:628)

Besides, the semantic relationship which can be found between the *wh*- relative clause and its antecedent could be either restrictive, as in (37), or non-restrictive, as in (38). This

distinction leads to distinguish between restrictive, as in (37), and non-restrictive *wh*-relative clauses, as in (38), (Quirk et *al.*, 1985).

(37) *[The period [**when** I was in London]$_{CP2}$ was the decisive one in my life]$_{CP1}$*
(38) *[Last year,[**when** I was in London]$_{CP2}$, was an incredible one]$_{CP1}$*

Hence, on the one hand, it can be said that the main difference between these types of *wh*-relative clauses is that, in restrictive ones, the *wh*-relative clause and the antecedent are very intensely related semantically speaking. As can be appreciated from example (37), the *wh*-relative clause, *when I was in London*, is essential to understand the whole meaning of the clause, otherwise it would not be clear to which period of time the speaker is referring to. Nonetheless, on the other hand, in (38), if the *wh*-relative clause, *when I was in London*, is deleted the clause keeps its meaning, although a piece of information has been lost. Therefore, non-restrictive *wh*-relative clauses add extra information that the speaker wants to highlight with respect to the whole sentence.

2.3. *Wh*-movement

Subsection 1.2 explains that both *wh*-questions and *wh*-relative clauses need to have the [+WH] feature, the former in the specifier position of CP_1, as in (39), whereas the latter on the specifier position of CP_2, as in (40).

(39) *[**What** are they doing?]$_{CP1}$*
    Spec.   (Biber et *al.*, 1999:204)
(40) *[That was the period [**when** she lived here]$_{CP2}$]$_{CP1}$*
                 Spec.   (Quirk et *al.*, 1985:1254)

So, in order to satisfy that requirement, a mandatory syntactic movement has to be done, that is called *wh*-movement. The target of this movement is to move the *wh*-element to the specifier position of CP_1, in the case of interrogative clauses, as in (39), and to the specifier

position of CP$_2$, in *wh*-relative clauses, as in (40). Furthermore, this movement leaves a trace in the original place of the *wh*-element, which indicates that a movement has been done. The trace (*t*) occupies the original position of the *wh*-element. Thus, by means of co-indexation the trace (t$_i$) is connected with the moved *wh*-element (*when$_i$*).

(41) *[**What$_i$** are they doing t$_i$?]$_{CP1}$*   (Biber et *al.*, 1999:204)

(42) *[That was the period [**when$_i$** she lived here t$_i$]$_{CP2}$]$_{CP1.}$*

(Quirk et *al.*, 1985:1254)

In (41) the original position of the *wh*-word *what* is right after the verb *doing*. However, it cannot be there because the [+WH] feature is not satisfied. Therefore, a movement, in order to take the *wh*-element up to the specifier position of the CP, has to be done. As a consequence, a trace *t* is used to fill up the empty space that the movement has left. Both the trace and the *wh*-word are co-indexed (*i*). The same happens in example (42), but, in this case, the *wh*-word moves to the specifier position of CP$_2$ as it is a subordinate clause (i.e. *wh*-relative clause).

3. PREVIOUS STUDIES

In this section of the dissertation, some previous studies dealing with the acquisition and production of *wh*-words, *wh*-questions, and *wh*-relative clauses, in general and in more detail, are presented. It is divided into several subsections. The first one deals with a brief overview of the acquisition and production of syntax by TD and SLI children. In the second subsection, the order of acquisition of *wh*-words is explained, first in TD and then in SLI children. Finally, in the last one, the focus is on *wh*-relative clauses as they are acquired by both in TD and SLI children.

Before going into further detail, it is important to mention that one of the principal aims of this dissertation, as presented in section 1, is to establish a double comparison in order to

determine whether TD and SLI children show the same linguistic behavior and to observe how they use *what* and *when* in relative and interrogative clauses. However, the second comparison has not been previously studied in SLI works and, consequently, in this section, both grammatical structures are discussed separately.

3.1. TD *versus* SLI acquisition and production of syntax

The acquisition and production of different grammatical structures may be regarded as problematic for SLI children. In fact, this is a fairly recent field of study, compared to that of TD children, and, consequently, there are not many investigations and conclusions reached about it and much needs to be done. Nevertheless, what seems to be clear for the researchers in the field is that language acquisition follows a slower path in SLI children, and that the grammatical structures produced by them are not equal to the ones produced by their typically developed peers. This could be the consequence of the lack of some grammatical features and abilities (Guasti, 2002:380).

Moreover, Guasti supports the idea that language acquisition can be defined as the act of attaining a language in a natural way which happens under diverse conditions, in an specific time period, spontaneously, in the same way in all languages, and by obtaining constructive support from the surrounding area. Using Guasti's idea, the language development process undergone by children can be established. To begin with, it is important to distinguish between two stages: when children are able to identify speech (which happens from birth) and when they are capable of producing speech (which does not happen at least until they are between 6 and 8 months old). Guasti calls this first stage *babbling* since they solely produce unvarying sequences such as *bababa*. During the next months, they try to figure out what the relationship is between what they listen and what they see, and it is not until they are one year old when they produce their first words. Besides, until the age of 6, they keep on learning words and matching them in a word-to-world mapping procedure, which means that they try to find a reference in the real world for each meaning (the concept). Therefore, although they begin to acquire grammatical

structures when they are around 2 years old, it is not until a late stage when they use them properly. This is further illustrated in the subsections below that deal with the target structures in this dissertation, that is, *wh*-words, *wh*-questions and *wh*-relatives.

3.2. Acquisition and production of *wh*-words and *wh*-questions

This section on how children acquire *wh*-words, particularly *what* and *when,* is structured into two subsections: firstly, how TD children acquire *wh*-words, based on Bloom et *al*.'s (1982) and Rowland et *al*.'s (2003) studies; and secondly, how *wh*-questions are acquired by SLI children, based on Van der Lely and Battell's (2010) study.

Additionally, as previously mentioned, the study of grammatical properties in SLI children is a fairly recent field. Consequently, in this subsection, two different issues are explained, relying on what has been previously studied. These are: how TD children acquire *wh*-words and how SLI children acquire *wh*-questions. Although the acquisition of these two issues would be different, by analyzing how SLI children acquire *wh*-questions could give an insight about the order of acquisition of *wh*-words.

3.2.1. The order of acquisition of *wh*-words in TD children

In the same line as Guasti (2002), Yule (2006) argues that a child goes through different stages until he properly develops the adult syntax of the language he is exposed to. The last stage occurs, approximately, when the child is 5 years old, although in some cases the acquisition process can take longer than that. Additionally, as pointed out before, generally SLI children acquire language later than TD children; and even once SLI children have acquired the different linguistic properties, their productions sometimes are still ungrammatical (Schuele and Dykes, 2005).

There are several studies which aim to explain whether there is a specific order in the acquisition of *wh*-pronouns (i.e. Bloom et *al*. 1982, Rowland et *al*. 2003). These two

studies are related since the latter is an attempt to prove the conclusions reached by the former.

Bloom et *al.* (1982) made a longitudinal study of 7 children, from 2 to 3 years old, from New York City, in order to examine how *wh*-words are acquired in the formulation of questions. They departed from the idea that once the *wh*-pronoun has been used at least three times properly, it is considered to be acquired. Therefore, they concluded that an order of acquisition can be established, as in table 1.

Table 1. Developmental order for *Wh*-words	
Wh- word	Average Age of Acquisition
Where, What	26 months
Who	28 months
How	33 months
Why	35 months
Which, Whose, When	36 months

Source: O'Grady (1997:130, table 7.1.)

Regarding table 1, the *wh*-words *where* and *what* are the first ones to be acquired (at the age of 26 months) followed by *who* (28 months) and by *how* (33 months). After these, at the age of 35 months a child acquires the *wh*-word *why* and lastly when a child is 36 months he produces the other three *wh*-words: *which, whose,* and *when.*

Besides, as Bloom et *al*. (1982) argue, the order of acquisition presented in table 1 could be related to the degree of abstraction of the *wh*-word itself. This means that the *wh*-words *what* and *where* refer to more concrete entities than, for instance, *when* or *whose*. Thus, children acquire earlier those *wh*-words which refer to more concrete units, and in a later stage those which are related to abstract ones. Consequently, in relation to the target *wh*-words in this dissertation, more instances of *what* than *when* are expected to find, as children could find more problematic the use of *when* at an early stage.

Furthermore, Bloom et *al*. (1982) subdivide *wh*-words into three types: *wh*-pronominals, *wh*-sententials, and *wh*-adjectivals. The first type is comprised by *what* and *where*, which

are supposed to be the easiest ones as they refer to things or places which are concrete realities for children. The second type, which includes *how* and *when,* is considered to be more complex, as the expected answer has a greater complexity than the previous ones; for example, an instance is expected in the case of *how*. Lastly, the *wh*-adjectivals *which* and *whose* are classified as the most difficult ones since the answer requires something which makes reference to an object constituent (i.e. *which ball?, whose dinner?).*

As a follow up to Bloom et *al*.'s work, Rowland et *al.* (2003) explain the degree of complexity that the acquisition of several grammatical structures has. They take as their departure point the fact that the order of *wh*-words, as in table 1, appears to be heavily reliable. Furthermore, such order depicts that the *wh*-words which present simple syntactical relations (i.e. *what* and *where*) are acquired earlier than those which portray more abstract or complex entities or contexts (i.e. *why, how,* and *when*).

Their study examines the production of twelve children in a longitudinal way. All of them are from middle-class backgrounds, six from Manchester and six from Nottingham, England. Their ages are between 1.08;22 and 2;0.25 at the beginning of the study and from 2;9.10 to 3;0.10 at the end of it.[1] Additionally, in terms of their MLUs (Mean Length of Utterance), these initially range from 1.06 to 2.22 and from 2.85 to 4.12 at the end of the study period. The procedure used is to audio-record the children playing in their homes with their mother in two different ways: half the time playing with their own toys whereas the other half is the investigator the one in charge of giving toys to them. This process takes place throughout one year, having one session every three weeks.

Furthermore, since there are a large number of structures regarding *wh*-questions, they decide to base on Bloom et *al*.'s (1982) study and focus their research, exclusively, in those *wh*-questions composed by a main or copula verb (i.e. *Where've you gone?, What are you doing?*). They conclude that many of the children have not acquired all the *wh*-words yet, and they reassert the established order of *wh*-questions proposed by Bloom et *al.*(1982), as in table 2.

---

[1] In this dissertation, the age is represented using the following format: years.months;days.

Table 2. Order of acquisition of *wh*-words

Child	Stage				
	I	II	III	IV & V	Not acquired
Anne	where (2) what (3)		who (12)	which (27)	**why/how/when/** whose
Aran		where (5)	what (12) who (13)	why (24) when (26) how (29) which (32)	whose
Becky	what (3) where (6)	who (11)	how (15)	why (18) **which** (32)	**when**/whose
Carl		what/where (1)			who/how/why/ which/whose/when
Dominic	what (4)	where (12)			who/how/why/ when/which/whose
Gail	what (1)	where (5)	**how** (16) **who** (20)		why/when/which/ whose
Joel		what (9) where (10)	**how** (17) **who** (18) **whose** (28)		**why/when**/which
John		what/where (2)			who/how/when/ why/which/whose
Liz	what/where (4)	who (9)		**which** (29)	**how/why/when/** whose
Nicole	where (11)	what (18) who (23) why (31)			how/when/which/ whose
Ruth			what (26) **why** (28) where (30)	**who** (31)	how/which/whose/ when
Warren	where (1)			what (17) who (21)	which/why/how/ when/whose

Source: Rowland et al.'s (2003:618, table 3)

As table 2 shows, children are classified into five different stages (according to their MLUs). In the *not acquired* column it is clear that in many cases some of the *wh*-words have not been acquired yet. However, what it is clearly demonstrated is the order of acquisition previously established by other researchers. Therefore, from these results, they conclude that *wh*-pronominals (*what, where,* and *who)* are the first ones in being produced followed by *wh*-sententials (*when, why* and *how*). *Wh*-adjectivals (*which, and whose*) are, therefore, the latest in being produced by children.

3.2.2. Acquisition of *wh*-questions by SLI children

Most studies carry out about *wh*-questions and *wh*-relative clauses have had as their focus TD children. According to Guasti (2002), the development of research dealing with SLI participants is very recent in the psycholinguistic field. This could explain, the few accounts covering the acquisition process and development of these type of participants. In fact, Van der Lely and Battell (2003) conduct a survey regarding the acquisition of *wh*-questions in SLI children, and they start by acknowledging that although there are few studies dealing with *wh*-movement in SLI children, those are really relevant in linguistic theory and language acquisition (e.g. Crain and Thornton 1999, de Villiers and Roeper 1995, or Rizzi 1990).

Van der Lely and Battell (2003) consider a specific subgroup of SLIs, Grammatical (G)-SLI participants. The impairment of G-SLI children comes from the syntactic computational system and, in their study, they focus on *wh*-questions and on how G-SLI and TD children acquire these structures. They base their study in the *Representational Deficit for Dependent Relations* (RDDR) hypothesis. According to what Van der Lely and Battell (2003:154) point out: "the RDDR contends that the core deficit responsible for G-SLI children's grammar is in 'Movement' (Chomsky 1995), and more specifically, that whereas the basic operation/rule 'Move' in normal grammar is obligatory, in G-SLI grammar it is optional". Hence, following this hypothesis, they made two predictions. Firstly, they argue that G-SLI children would present difficulties in the production of *wh*-movement and T/Q-feature movement.[2] And secondly, they point out that G-SLI children would experiment fewer problems in the production of *wh*-movement in subject questions than in object ones. That is because object *wh*-questions comprise two different movements, firstly, as explained above, the *wh*-element needs to move to the specifier position of the CP_1 leaving a trace in its original place; and secondly, the T/Q feature movement has to be satisfied. In the case of subject *wh*-questions, the last movement, T/Q

---

[2] The T/Q-feature movement means that those questions which have an auxiliary verb have to undergo movement, in which the auxiliary verb is moved to the head of the CP. After the movement, the T/Q feature is satisfied.

feature, does not take place, and therefore, the *wh*-element can move directly to the specifier position of CP$_1$.

The study is carried out amongst three groups of teenagers, one group of G-SLI participants, and the other two of younger control participants. They decide to use two control groups in order to conduct different tests, in regard to diverse language abilities. The G-SLI group is formed by 15 teenagers whose ages range from 11;5.0 to 18;2.0. The other two groups of TD children have 12 participants each and are selected in a random way in a school of London. Their ages range from 5;3.0 to 7;4.0 in the language ability control group 1; and from 7;4.0 to 9;1.0 in the second control group. Moreover, these three groups are matched regarding vocabulary and morphosyntactic abilities.

In order to compile the data, the method used is based on a game with the intent to elicit subject and object questions, mainly focuses on three *wh*-words: *who, what,* and *which.* Once they have all the data collected, they observe that the TD children properly produce 70% of *wh*-words for subject and object questions. However, this percentage is reduced to less than 51% in the case of SLI children with the exception of *who* that is properly produced in 80% of the cases. The production of the different groups and *wh*-words is represented in table 3.

Table 3. Correct production for the three participants groups

	G-SLI			LA1			LA2		
	Mean (n = 6)	%	SD	Mean (n = 6)	%	SD	Mean (n = 6)	%	SD
WORD TYPE				SUBJECT					
who	5.08	85	1.00	5.33	89	1.30	6.00	100	0.00
what	3.17	51	1.75	4.92	82	1.38	5.42	90	0.67
which	3.08	51	1.31	4.75	79	1.71	5.92	98	0.29
SUBTOTAL (n = 18)	11.33	63	3.31	15.00	83	3.81	17.33	96	0.78
WORD TYPE				OBJECT					
who	1.33	22	1.23	3.25	54	2.34	5.50	92	0.67
what	2.67	45	1.72	4.25	71	1.82	5.83	97	0.39
which	2.00	33	2.17	5.00	83	1.76	5.75	95	0.62
SUBTOTAL (n = 18)	6.25	35	4.29	12.50	69	5.13	17.08	95	1.00
TOTAL (n = 36)	17.75	49	5.83	27.50	76	8.69	34.42	96	1.51

Source: Van der Lely and Battell (2010:164, table 2)

Table 3 depicts the mean average of correctness, of both subject and object *wh*-questions. This table portrays the accurate production of *wh*-questions looking at different issues. The results are compiled separately according to the *wh*-word (i.e. *who, what*, and *which*), the child group (i.e. G-SLI, LA1, and LA2), and the type of *wh*-question (i.e. subject or object).

Moreover, with this study they also prove that there is a difference in the production of the two control groups, as the production of the older group is clearly better. Furthermore, Van der Lely and Battell (2010) findings lend support to the RDDR hypothesis and conclude that G-SLI participants' performance is worse than the one of the two other control groups. They fail in the proper production of *wh*-movement and T/Q feature movement. Additionally, this difference in the performance is highly remarkable when it deals with the production of *what* and *which,* but it is slightly so in the case of *who.* This fact indicates that G-SLI children find less problematic the use of *who* over the rest of words.

So far, these previous studies suggest several issues that are relevant for this dissertation: firstly, it is clear that there is an order of acquisition regarding *wh*-words: *wh*-pronominals (*what, when,* and *who),* followed by *wh*-sententials (*when, why* and *how*), followed by *wh*-adjectivals (*which* and *whose*). Secondly, TD and SLI children show a preference for some *wh*-words over others, for instance; they prefer *what* over *when.* Lastly, and in general terms, it is evident that SLI children start producing *wh*-questions later than TD children.

3.3. Acquisition and production of *wh*-relative clauses

Two studies are reviewed here: one on TD children (Shcuele and Dykes 2005) and one on SLI children (Sheldon 1974).

3.3.1. Acquisition of *wh*-relative clauses by TD children

The analysis of *wh*-relative clauses in the production of TD children has been previously studied by different authors (e.g. Limber 1973, Sheldon 1974, Hamburger 1980, Flynn and

Lust 1980). They come to the conclusion that the first instances of *wh*-relative clauses produced by children usually do not have an explicit *wh*-word, as in (43), or tend to overgeneralize the use of *what,* as in example (44) (O'Grady 1997).

(43)   *Look it [_ Mommy have on]* (29 months)

(44)   *Look-a [**what** I made]* (28 months)   (O'Grady 1997)

Sheldon (1974) distinguishes between two types of *wh*-relative clauses: the self-embedded *wh*-relative clauses, as in (45), and the right branching *wh*-relative clauses, as in (46). The former type of *wh*-relative clauses seems to be harder for children since there is an interruption of the main clause, and consequently children need to memorize part of it. Thus, as in (45), the *wh*-pronoun *which* functions as the object of CP_2. However, this fact is not seen in right branching *wh*-relative clauses, as in (46), which in general terms are considered the easiest ones since they do not interrupt the main clause.

(45) *[I will read the memo [which Pat hopes[that John will send you]$_{CP3}$]$_{CP2}$]$_{CP1}$*

(46) *[We'll go to my parent's place [when you're ready]$_{CP2}$]$_{CP1}$*

(Quirk et al., 1985:1038,1298)

Consequently, Sheldon remarks that if children's acquisition abilities mainly rely on short term memory, they find problematic the acquisition of *wh*-relative clauses as embedded ones, since these require a longer term memory.

Additionally, Sheldon points out that in order to observe children's attitude towards these syntactic structures, information about how they understand *wh*-relative clauses must be collected. Hence, his main objectives are to prove what information is needed and to demonstrate the statements established by three hypotheses (the *Interruption Hypothesis,* the *Word Order Hypothesis,* and the *Parallel Function Hypothesis*) about the acquisition of *wh*-relative clauses.

Sheldon explains that the first hypothesis, the *Interruption Hypothesis*, predicts that children find more constraints in the acquisition of subject *wh*-relative clauses, as in 47,

than in that of object *wh*-relatives, as in (48). This is because in subject *wh*-relative clauses, as in (47), the *wh*-relative clause (CP$_2$) interrupts the main clause (CP$_1$). Conversely, in the case of object *wh*-relative clauses, as in (48), this does not happen, as the *wh*-relative clause appears at the end of the sentence.

(47) *[The boy [who hit the girl]*$_{CP2}$ *saw the man]*$_{CP1}$

(48) *[The man saw the boy [who hit the girl]*$_{CP2}$*]*$_{CP1}$ (Sheldon, 1974:274)

Meanwhile, the *Word Order hypothesis* establishes that those clauses in which the relativized constituent is the subject DP and not the object DP are easier to acquire. Lastly, Sheldon explains that the *Parallel Function Hypothesis* asserts that children infer that the relative pronoun has the same grammatical function as its antecedent. Therefore, this suggests that those *wh*-relative clauses in which the antecedent and the relative pronoun share the same grammatical function are acquired more easily by children, as illustrated in examples (49) and (50).

(49) *[The man saw **the boy** [**who** the girl hit]*$_{CP2}$*]*$_{CP1}$

(50) *[The man saw **the boy** [**who** hit the girl]*$_{CP2}$*]*$_{CP1}$ (Sheldon, 1974:275)

As in (49), both the relative pronoun *who* and its antecedent *the boy* have the same grammatical function, the object of their corresponding clauses. However, in example (50), the relative pronoun *who* and its co-referential DP *the boy* do not share the grammatical function, as the former is the subject of CP$_2$, whereas the latter is the object of CP$_1$.

Sheldon's study is based on a toy-manipulation task in which different types of *wh*-relative clauses, are provided to the children. Afterwards, a control test is used which contains a version of these clauses but instead of including subordination they are presented as coordinated sentences. The participants are 33 English monolingual speakers, aged from 3;8.00 to 5;5.00 from Texas Nursery school, and they are divided into three groups according to their ages.

After carrying out the study, Sheldon concludes that the difference in the production of the two types of *wh*-relative clauses is not significant. He finds out that neither the *Interruption Hypothesis* nor the *Word Order Hypothesis* are in line with the results. Conversely, this study lends support to the *Parallel Function Hypothesis* in which the relative pronoun and the antecedent should have the same grammatical function. The last conclusion he reaches is that for children coordinate sentences are easier than *wh*-relative clauses which is expected given that subordination is syntactically more complex than coordination.

Additionally, as Guasti (2002) suggests, other studies claim that the knowledge children have about *wh*-relative clauses is limited and that the comprehension of these structures is delayed until children are at least 6 years old.

3.3.2. Acquisition of *wh*-relative clauses by SLI children

Schuele and Dykes (2005) depart from the idea that there is not enough information about how SLI children acquire and produce complex syntactic structures. In their study, they classify as complex syntax the use of infinitive clauses, *wh*-relative clauses, and subordinate clauses. However, only those issues concerned with the relative ones are explained as the others are not targeted in this dissertation.

They base their research on Leonard's (1995) study and argue that complex grammatical structures appear both in TD and SLI children when they have the same MLU rates. However, SLI children show several deficiencies in the accurate production of complex syntax that are not found in their TD MLU-matched partners.

Schuele and Dykes (2005) with their study want to support two ideas: firstly, that SLI children do acquire certain complex grammatical constructions although later than their corresponding peers; and secondly, that some requirements of *wh*-relative clauses seem to be more problematic for SLI children.

Their research consists of a longitudinal study of one single SLI child. The data they use are collected from a database and they use 12 language samples from the same child aged from 3;3.0 to 7;10.0. The data are spontaneous, since the study is based on conversations between the investigator and the child and, on some occasions, on interactions between the child and the different family members. As the child gets older the topics of the conversations change accordingly to his age, which allows for more variety in his production.

According to the results, they prove that the child was 4;8.0 (MLU of 3.12) when he used the first *wh*-relative clause. Moreover, by analyzing the results, they find out that the acquisition and production of *wh*-relative clauses vary depending on their type. They affirm that subject *wh*-relative clauses, as in (51), are the first ones to appear, followed by nominal *wh*-relative clauses, as in (52), while the other *wh*-relative clause type, for instance, adjunct *wh*-relative clauses, as in (53), appear the latest.

(51) *[The man[**who** crashed the car]$_{CP2}$ is in jail]$_{CP1}$*
(52) *[The man [**who** Mary invited]$_{CP2}$ is here]$_{CP1}$*
(53) *[I wrecked the car the time [**when** I went to the store]$_{CP2}$]$_{CP1}$*

(Schuele and Dykes, 2005:303)

All in all, they conclude that for this child the acquisition of *wh*-relative clauses (seen as an instance of complex syntax) happens late, an issue that may be linked to his language impairment. Nevertheless, this needs more research since this study is just based on one single case; although as they point out there are some other studies dealing with morphosyntactic difficulties in SLI children which reach similar conclusions (i.e. Leonard 1998).

In short, these studies basically support two issues: firstly, that both TD and SLI children experience problems in the acquisition of *wh*-relative clauses as they are an instance of complex syntax. Therefore, these structures are acquired in a late period, when children are 6 years old in the case of TD children, and a little bit later in the case of SLI children. Secondly, for both types of children there seem to be some *wh*-relative clauses which seem

to be easier than others. TD children prefer right branching *wh*-relative clauses and those in which the relative pronoun and the antecedent have the same grammatical function. In the case of the SLI children, they clearly prefer subject *wh*-relative clauses.

4. PURPOSE OF THE STUDY AND RESEARCH QUESTIONS

This dissertation deals with the production of *what* and *when* by TD and SLI English children when their MLUs are similar, that is, with MLU-matched TD and SLI children. Moreover, it aims to present the distribution of these *wh*-words in relative and in interrogative sentences. Furthermore, a comparison of these issues in regard to two different types of data, spontaneous and narrative, is carried out, too. Taking these targets into account as well as the review of previous works conducted in the previous section, this one presents the research questions that have guided this dissertation.

Three research question sets are formulated and the three offer a double comparison between child groups (TD *versus* SLI) and data type (spontaneous *versus* narrative).

4.1. Question set 1: general performance

This first question set has to do with the collapsed production of the two *wh*-words considered in this dissertation (*what* and *when*), that is, without making a distinction between the two of them.

4.1.1. Research question 1: production of *wh*-words

Are there any relevant differences in the production of *wh*-words by the TD and the SLI children? According to previous studies that support that SLI children acquire language, in general terms, later than TDs (e.g. Yule 2006), a higher production in the case of TD children is expected.

4.1.2. Research question 2: production of *wh*-words per data type

In this second research question, the target is to see whether the production of *what* and *when* varies in relation to the data type (i.e. spontaneous *versus* narrative data). Since previous studies have not compared between spontaneous and narrative data, the main objective is to see what happens when the type of data is taken into account.

4.2. Question set 2: performance per *wh*-type

4.2.1. Research question 3: production of *what* and *when*

Regarding the studies done by Bloom et *al.* (1982) and Rowland et *al.* (2003), TD children acquire and produce *what* earlier than *when*. Hence, is there any difference in the amount of production of *what* and *when* by the TD children and by the SLI participants in this study as well? Furthermore, is there any correlation between *wh*-word (i.e. *what* or *when*) and participant type (i.e. TD or SLI)?

4.2.2. Research question 4: production of *what* and *when* per data type

The previous research question, studies what happens when the data are collapsed. But, is there any difference when the data are analyzed separately per type (i.e. spontaneous *versus* narrative data)? Do the results differ or do similar results arise?

4.3. Question set 3: performance per sentence-type

4.3.1. Research question 5: distribution of cases across sentence-type

As previous studies (Schuele and Dykes 2005, and Leonard 1995) point out, the acquisition and production of *wh*-relative clauses, understood as complex syntax, seem to be harder not only for SLI children but also for TDs. Thus, that both the TD and the SLI children would produce more *wh*-questions than *wh*-relative ones is expected.

So, do they produce *what* and *when* more in *wh*-relative clauses or in *wh*-questions? Is there any specific correlation between target *wh*-words (i.e. *what* and *when*) and sentence type (i.e. *wh*-relative clauses and interrogative questions) by the TD children and by the SLI children? Do they present the same preference in their productions?

4.3.2. <u>Research question 6: distribution of cases across sentence and data type</u>

Following what has been addressed in research questions 2 and 4, this last one deals with the data separated per sentence and data types. Therefore, would the TD and the SLI children rather use these *wh*-words in *wh*-questions and *wh*-relative clauses and in spontaneous and narrative data? Is there any relevant variation across sentence and data types?

5. DATA METHODOLOGY

This section of the dissertation, deals with the process followed in order to obtain the required data for the present study. Three main subsections are included. In the first subsection, the data selection process is explained, in relation to the corpora and participants chosen, and the MLU-matching criteria followed. The second subsection contains an explanation on how the data are extracted in regard to the Child Language Analysis (CLAN) programs used. Lastly, the criteria follow to select and classify the pertinent data for this study are presented.

5.1. <u>Data selection</u>

In order to conduct this research, the necessary data have been extracted from the CHILDES project (MacWhinney, 2000). This database includes different corpora that comprise data, from a diverse range of languages, which are video and/or audio recorded, and later on transcribed.

The data used for this study have been extracted from three different corpora. All of them belong to the Clinical-MOR folder which includes data of atypically developed children with specific disorders as well as typically developed ones to serve as a point of comparison. From Clinical-MOR, three corpora haven been used which include the two different groups of children to be examined: the ENNI corpus, from which 24 participants have been selected; the Conti-Ramsden 3, from which 4 children have been chosen; and another 72 children from the Conti- Ramsden 4 corpus. From the first two corpora, only SLI children have been selected, whereas from the Conti-Ramsden 4 corpus both TD (i.e. 50 children) and SLI participants (i.e. 22) have been picked.

Besides, for all the corpora the selection criteria which have been taken into account are the following: the type of data (i.e. narrative and spontaneous), the number of participants, the child group (i.e. TD and SLI), the age range, and the MLU range. The relevant properties are represented in tables 4 to 6 below.

Table 4. The ENNI corpus				
Type of data	MLU range	# of participants	Child group	Age range
Narrative	1.08-4.43	77	SLI	4;2.4-9;9.22
Narrative	---	300	TD	4;0.0-9;11.0

As table 4 shows, the ENNI corpus contains the narrative data of 77 impaired children and 300 control, all monolingual English speakers from Canada. From this corpus, only the impaired children has been analyzed. This group comprises 77 children aged between 4;2.4 and 9;9.22. Although biological gender is not a factor to be considered, there are 48 boys and 29 girls. Additionally, their MLU range is from 1.08 to 4.43. The data from this corpus are experimental and non-longitudinal, which means that there is just one file per participant. Consequently, the total amount of files taken out from this corpus are 77, although, as it is explained later on, not all them have been finally analyzed.

These children are recorded in their corresponding schools, preschools, or nurseries. In order to compile all the desirable data, the participants are divided into 6 groups (rising up

the level in each one). Once each child is assigned to a group, a different set of pictures is given to them, through which they have to tell an invented story about what they see or appreciate in the pictures.

Table 5. The Conti-Ramsden 3 corpus				
Type of data	MLU range	# of participants	Child group	Age range
Spontaneous	1.00-4.34	4	SLI	2;6.0-5;0.0

As in table 5, the Conti-Ramsden 3 corpus contains spontaneous data from 4 children from a longitudinal study. These 4 participants are English speakers from the UK, aged between 2;6.0 and 4;0.0 with and MLU range from 1.06 to 2.85. This corpus has between 22 and 23 files per participant. However, for this dissertation, only one file has been selected from each child, according to the MLU-matched criterion.

The 4 children are examined through a period of 16 months, and during this period of time, since the beginning of the study, they are attending nursery schools. However, the study is carried out at their homes, in a silent place, while the participants are playing generally with their mothers, although at some point other relatives may appear.

Table 6. The Conti-Ramsden 4 corpus				
Type of data	MLU range	# of participants	Child group	Age range
Spontaneous	1.7-11.2	99	TD	13;0.19-15;11.16
Narrative	2.38-7.07	99	TD	13;0.19-15;11.16
Spontaneous	1.19-7.83	19	SLI	13;1.09-15;4.01
Narrative	3.19-6.82	19	SLI	13;1.09-15;4.01

Table 6 conveys the properties of the last corpus used, the Conti-Ramsden 4. This corpus comprises two different sets of data: narrative and spontaneous. All the participants, 99 TDs and 19 SLIs, are from the UK and both types of data are collected from them. Their ages vary from 13;0.19 to 15;11.16 years old, and their MLUs range between 1.7 and 7.07 in the case of the TDs, and between 1.19 and 7.83 for the SLI children.

This corpus is the only one used in order to extract data from TD children, although the data from SLI participants are analyzed, too. All the participants are from the UK and their data are compiled through a storytelling method and a conversational program.

In short, these four corpora have been selected in order to satisfy the necessities of the study. As it can be observed, one corpus has been used to obtain data from TD children (i.e. Conti-Ramsden 4), whereas in order to extract SLI data three corpora have been needed (i.e. ENNIN, Conti-Ramsden 3, Conti-Ramsden 4).

5.2. Participants' selection

As briefly mentioned above, not all the children from the four corpora have been used for the analysis of this study. Thus, this subsection includes information on how the participants have been selected and the criteria used.

The first aspect that has been taken into account is their production. The four corpora used have a great number of files and participants. Therefore, the first step in the participants' selection has been to see who produces the relevant *wh*-words, *what* and *when*. Afterwards, all those participants who do not present at least one instance of *what* or *when* have been automatically discarded.

Once all the participants who produce these two *wh*-words have been selected, their MLUs have been calculated (as it is explained in the next point), in order to classify them according to their MLUs. By taking into account their MLUs, somehow, it is assured that all of them have acquired the same level of production and comprehension, as their production of morphemes or words per utterance is quite similar. Therefore, after having all the MLUs of the participants, these have been analyzed and a comparison between the ones of the TD and the SLI children has been carried out. In order to select the final number of participants, the MLU-matching criterion has been followed. In this way, each SLI children has been matched with a TD peer whose MLUs closely resemble. This ensures that, within

each TD-SLI pair, a comparison of children that are at the same developmental stage is done.

Additionally, as there are more TD children who produce either *what* or *when* or both, the number of SLI children has been taken into account when performing the MLU-matching criterion. This means, that there are as many TD peers as SLI children who produce these *wh*-words, and not the other way around.

After these criteria were applied, as tables 7 and 8 show, two classifications have been done. On the one hand, and in the case of narrative data, 35 SLIs and their 35 MLU-matched TDs have been selected. On the other hand, and in the case of spontaneous data, 15 SLIs and their corresponding 15 MLU-matched TDs have been picked out. All in all, a total amount of 100 children have been analyzed.

Table 7. MLU-matched TDs and SLIs: narrative data		
MLU RANGE	GROUP	#Of participants
2.0-2.9	SLI	5
	TD	4
3.0-3.9	SLI	18
	TD	19
4.0-4.9	SLI	9
	TD	9
5.0-5.9	SLI	1
	TD	1
6.0->	SLI	2
	TD	2

As in table 7, the narrative data from the TD and SLI participants have been classified according to their MLUs, which in this table are clustered in 5 developmental groups according to their MLU range. The last column represents the number of participants from both groups who have been analyzed in each MLU range.

Table 8. MLU-matched TDs and SLIs: spontaneous data		
MLU RANGE	GROUP	#of participants
2.0-2.9	SLI	3
	TD	3
3.0-3.9	SLI	4
	TD	4
4.0-4.9	SLI	1
	TD	2
5.0-5.9	SLI	4
	TD	2
6.0->	SLI	3
	TD	4

In this case, table 8 conveys exactly the same information as table 7, but with regards to spontaneous data. Additionally, as it can be observed in both tables (table 7 and table 8), the MLU stages cover a wide period (from MLU 2 to MLU 6 and above). This is useful to depict at the end of the analysis how their production changes when their MLUs are higher.

In short, two main points have been taken into account, firstly whether they produce *what* and *when* or not, and, secondly from those who present some instances of these *wh*-words, their MLUs have been extracted and matched between the two child groups. Hence, only those whose MLUs can be matched with a child from the opposite group have been selected for the analysis.

Additionally, place of origin has not been taken into account, that is, where they came from, as long as their mother tongue was English; considering that the property that is being studied is not affected by regional conditions. So, whether their English is British or American has no further relevance for the present analysis.

5.3. Data extraction

In order to obtain the necessary information from the corpora selected, the CLAN program package, a tool which allows to codify and analyze data, has been used. It enables to carry

out different analyses of the data available in CHILDES (MacWhinney 2000). So, this subsection describes each CLAN program used: *FREQ, MLU*, and *KWAL*.

Each of the programs just mentioned has a specific function. Therefore, each one has been used in order to obtain specific information from the data. Initially, who produces the *wh-*words *what* and *when* needed to be determined. Thus, in order to obtain that information, the CLAN program *FREQ* has been used. This program displays the number of cases in which the element in question appears in each file. Hence, in this case, by using *FREQ* the number of *what* and *when* instances in each child's production has been obtained.

Besides, it is important to highlight that at the beginning of the whole process, the program *FREQ* was used without narrowing the search to *what* and *when*. Rather, it was used to see all the instances of all *wh*-words, including *how*. Hence, once the number of instances of each *wh*-word was obtained, to focus exclusively on the two most used, in this case, *what* and *when*, was decided. Thereby, the rest of the analysis was restricted to just these two *wh-*words. In order to run these programs, a syntax line, based on what it was expected to extract was introduced. Here, the following ones were used: *<freq +t*CHI +s"what*" @>* to search for the number of *what* cases produced; and *<freq +t*CHI +s"when*" @>* for the cases of the *wh*-word *when*.

Subsequently, in order to classify the participants according to their MLUs, these were calculated. Hence, the *MLU* program which provides the rate of morphemes or words that a child produces per utterance was used. In this case, the corresponding syntax line introduced was *<mlu +t*CHI @>* which calculates the average number of morphemes per utterance.

Lastly, in order to determine whether these *what* and *when* instances are relativizers or interrogative clause markers, the context in which they appear needs to be output. Therefore, the *KWAL* program, which gives the context of each instance, was used. Depending on how much contextual information one is interested in, the syntax line would vary. In this case, with the two previous and following lines was enough, so the syntax lines

used were *<kwal +t*CHI +s"what*" -w2+w2 @>* for *what*; and *<kwal +t*CHI +s"when*" -w2+w2 @>* in the case of *when*.

Once all this was done, all the information and data extracted were sorted in an excel document which is attached in an electronic format to this dissertation. This document consists of two sheets, one compiles all the information of the spontaneous data whereas the other one that of the narrative data. Likewise, both sheets comprise the same information: the participant's number, the participant's type (i.e. TD or SLI), the file, the number of *wh*-words produced (differentiating at the same time between *what* and *when*), the MLU, the age, the example, and the sentence type (i.e. *wh*-relative or interrogative clauses).

Table 9. Properties of the study's corpus									
Participant		File	*Wh*- pronoun		MLU	Age	Example	Sentence type	
Number	Type		*What*	*When*				Inter.	Rel.
Child 1	SLI	444.cha	1	0	2.0	4;2.4	***What*** (is that sound?	1 (*what*)	0

Table 9 shows the different properties which are reflected in the excel document for each child and for each *wh*-word, either *what* or *when*, that they produce.

5.4. <u>Discarded cases</u>

In the data selection and classification processes, some instances which do not follow the criteria initially set arise and are, therefore, discarded. This subsection offers an overview of these cases and the reason why they are discarded.

Regarding the productions of *what,* there are some cases in which its compound form *whatever* appears. Nevertheless, in those cases, the *wh*-word *whatever* is not used as an interrogative clause marker or as a relativizer. It seems that it is used as an afterthought or even as an interjection, as shown in examples (54) and (55).

(54) *CHI: *and then the boy was looking down the hole like a mole hole or something like that (.) **whatever**.*

(55) *CHI: *then he ended in a river or pond **whatever**.*

(Target child,16;6.28, fssli519.cha)

Examples (54) and (55) are produced by a SLI child in narrative data. So, all the cases like these ones in which the *wh*-word *whatever* does not function as an interrogative clause marker or as a relativizer are discarded from the study.

Apart from these cases, there are other instances in which the child's productions of *what* and *when* are not clear. In these cases, the child doubts about what he is saying and, makes a pause and afterwards continues with the speech, as illustrated in examples (56) and (57). Hence, those cases cannot be considered proper elements of the final utterance, but as isolated words that the child produces before realizing the whole clause.

(56) *CHI: *and <**when**> [/] when he done making a castle a bunny rabbit (.) making it castle too.*

(57) *CHI: <**when**> [/] when rabbit said hi.* (Target Child, 5;2.25, 575.cha)

As in (56) and (57), angle brackets are used in the transcription of the data. This symbol indicates that the words which are between them overlap with another part of speech. Furthermore, right after the angle brackets another symbol appears, *[/]*, which signifies the previous idea mentioned: the child interrupts the speech in order to revise or replace what he has previously said (MacWhinney, 2018). So, in these cases, the children are reformulating their own speech and, hence, these instances are discarded from the study, too.

At this point, once all the methodology followed to obtain the results is explained, the results obtained are presented in the following section.

6. RESULTS OF THE STUDY

In this section of the dissertation, the results obtained after carrying out the analysis are shown. There are three main subsections, each of them related to one of the research questions sets previously explained in section 4.

Firstly, the findings of how much children use the *wh*-words *what* and *when* are presented. Afterwards, whether the TD and the SLI children have a preference for one *wh*-word over the other is depicted. Finally, if the children use *what* and *when* more in *wh*-relative clauses or in *wh*-questions is addressed. Additionally, each research question set has a double comparison in relation to the type of data, as previously explained.

In addition, in order to conclude this part, a developmental graph is included. This shows the rate at which these *wh*-words are used according to the participants' MLUs. In this way, it is possible to see whether their performance on *wh*-words is better as their MLUs increase.

6.1. Question set 1: General performance

6.1.1. Research question 1: production of *wh*-words

The first research question has to do with the number of times both the TD and the SLI children use the *wh*-words *what* and *when*. The principal aim here is to see if the production of these *wh*-words differs across the two child groups; and, if so, to what extent it varies.

Table 10. *Wh*-words per child group and data type			
	WH-WORDS	NARRATIVE DATA	SPONTANEOUS DATA
TD CHILDREN	97 (100%) 31%	61 (62.9%)	36 (37.1%)
SLI CHILDREN	216 (100%) 69%	119 (55.1%)	97 (44.9%)
TOTAL	313 (100%)	180 (57.5%)	133 (42.5%)

Table 10 shows a broad overview of the production of *wh*-words per child group and data type (results on data type are explained in the following subsection). In order to answer research question 1, *what* and *when* are initially grouped and analyzed together as *wh*-words.

As the total row from table 10 reveals, there are 313 (100%) instances in which children use either *what* or *when*. Out of the total, 31% of *wh*-words are produced by the TD children, whereas the other 69% belongs to the output of the SLI children. Overall, these results suggest that the SLI children use these *wh*-words at a higher rate than their MLU-matched TD participants. Besides, this difference is quite prominent as the production of *wh*-words by the SLI children is more than double the one by the TD children (216 cases *versus* 97 cases). Therefore, in order to answer the first research question, it can be said that these words are far more used by the SLI children (69%) than by the TD children (31%). Consequently, these results go in a different direction from what was expected: it was anticipated that the TD children would employ these *wh*-words more than their MLU-matched SLI participants but the opposite is found in the data analyzed.

6.1.2. <u>Research question 2: production of *wh*-words per data type</u>

In this case the focus of analysis is placed on what happens when the data are separated per type (i.e. narrative *versus* spontaneous), which can be seen in the last two columns from table 10. The main objective in this research question is to analyze whether the production of *what* and *when* changes or not in relation to the data type.

As table 10 shows, 313 *wh*-words are used by these children, among which 180 (57.5%) are used in narrative data, and the remaining cases, 133 (42.5%), in spontaneous data. This suggests that, although there is a predominance of *wh*-words in narrative data, the difference is not remarkable. So, it can be established that *wh*-words equally appear in spontaneous as well as in narrative data.

Besides, these results are analyzed separately per child group. The production of *wh*-words by the TD children amounts to 97 (100%) of the cases of which 61 (62.9%) belong to spontaneous data and the rest, 36 (37.1%), to narrative data. These results indicate that the use of *wh*-words in narrative data by the TD children is almost doubled if compared to the one in spontaneous data. Conversely, although the SLI children also produce more *wh*-words in narrative data, 55.1% (119 instances), than in spontaneous data, 44.95% (97 examples), the difference is not that relevant as it is in the TD children.

Thus, as it can be observed, there is a preference of using these *wh*-words in narrative data by both the TD and the SLI children. This could be a consequence of the study itself, as in those cases the investigator tries to guide the participants' production. Moreover, it also suggests that these children have already acquired these *wh*-words although they may produce them more often when they are pushed to do it, rather than when they are in a regular conversation. In short, and in order to answer the research question 2, it can be concluded that the production of *wh*-words differs per data type and child group.

6.2. <u>Question set 2: performance per *wh*-type</u>

The second research set deals with the production of the *wh*-words *what* and *when* separately. As in the previous question set, firstly the production of *what* and *when* by the TD and the SLI children is analyzed, without dividing per data type, an issue that is studied subsequently.

6.2.1. <u>Research question 3: production of *what* and *when*</u>

Children seem to follow and order in the acquisition of *wh*-words, as pointed out above in section 3. So, once the data have been analyzed without distinguishing between *what* and *when* instances, this differentiation is now done to see whether children produce *what* earlier than *when*, as expected.

Table 11. Production of *what* and *when*			
	Wh-words	*What*	*When*
TD children	97 (100%)	23 (23.7%)	74 (76.3%)
SLI children	216 (100%)	143 (66.2%)	73 (33.8%)

Table 11 shows the production of *what* and *when* by both the TD and the SLI children. As it can be seen, out of the total number of *wh*-words (97, 100%), the TD children produce 23 instances of *what* (23.7%) and 74 of *when* (76.3%). On the contrary, the results obtained from the SLI children are utterly opposed to the TD children's productions, since they employ the *wh*-word *what* in 143 occasions (66.2%) and *when* in 73 (33.8%).

In regard to the production of *what* and *when* by the TD children, it can be concluded that the results do not precisely follow what was expected. The production of *when* (76.3%) is more than three times higher than that of *what* (23.7%). Nevertheless, this does not mean that the acquisition of *when* in these participants occurs earlier than that of *what*. In order to establish whether this is a matter of acquisition or just a question of preference in their productions, a developmental study of the data is needed, as in graph 1 below.

By contrast, the results from the SLI children meet the expectations, as they produce the *wh*-word *what* (66.2%) twice as much as *when* (33.8%). Therefore, the results from the SLI participants confirm what was expected. Besides, by doing a comparison between the two groups, it is clear that the TD children make more use of *when*, whereas the SLI participants do so of *what*. This could mean that although they are MLU-matched, the production of *when* by the SLI children, at this stage, may still be problematic. Perhaps for that reason, they produce far more other *wh*-words such as *what*.

So, in order to answer research question 3, it can be concluded that there is a correlation between the use of one *wh*-word in each child group. In other words, the TD children produce more instances of *when*, whereas the SLI participants far more of *what*.

6.2.2. Research question 4: production of *what* and *when* per data type

Since the production of *when* is higher in TD children and the one *what* in SLI participants as previously established; what happens when the data are analyzed separately per data type? In this subsection, the data are examined in order to answer this question.

Table 12. Production of *what* and *when* per data type					
	NARRATIVE DATA			SPONTANEOUS DATA		
	Total	*What*	*When*	Total	*What*	*When*
TD Children	**61 (100%)**	11 (18.0%)	50 (82.0%)	**36 (100%)**	12 (33.3%)	24 (66.6%)
SLI Children	**119 (100%)**	68 (57.1%)	51 (42.9%)	**97 (100%)**	75 (77.3%)	22 (22.7%)

Table 12 shows the data classified per *wh*-word (i.e. *what* or *when*) and per data type (i.e. narrative or spontaneous) and, as previously explained, in both child groups the production is higher in narrative data (table 10 above).

As it can be appreciated from table 12, in narrative data the difference in the production of *what* and *when* is extremely noticeable. The percentage of *when* produced by the TD children amounts to 82.0% whereas 18.0% deals with the realization of *what.* This preference is noticed as well in spontaneous data since the production of *when* (66.6%) doubles that of *what* (33.3%). Therefore, the TD children clearly use *when* more than *what* in both types of data.

Conversely, this great disparity in the percentages is not found in the case of the SLI children in narrative data but it is in spontaneous data. In the former data type, *what* is used more but with a slight difference, since it is used 57.1% of the times whereas *when* 42.9% of the cases. However, in the production of these *wh*-words in spontaneous data, there is a clear preference for *what* (77.3%) over *when* (22.7%). These results could suggest that when the data are collected in a spontaneous way, the SLI children tend to use *what,* as it is believed to be easier syntactically speaking, more than *when*. Nevertheless, when they are being studied under some patterns or conditions, as in the case of the compilation of narrative data, they produce the *wh*-word *when* in more occasions than in a normal speech,

mainly because they are led to do it. So, it can be concluded that it does not seem to be a matter of acquisition, as it is proved that they produce both *wh*-words, but as an issue of preference or easiness linked to the extra-linguistic conditions in which the children produce language.

6.3. Question set 3: performance per sentence-type.

Research question set 3, the last one, deals with the production of these *wh*-words in relation to sentence type (i.e. *wh*-question or *wh*-relative clause). So, in this subsection, it is shown whether in the production of the TD and the SLI children there is any specific correlation between one *wh*-word and one sentence type. Lastly, it is explained if any difference, from the previous results, can be found when analyzing the data per data type.

6.3.1. Research question 5: distribution of cases across sentence-type.

This research question has as its first focus to examine the production of *wh*-questions and *wh*-relative clauses by TD and SLI children. Afterwards, which *wh*-word is used in each instance is studied in order to determine whether there is a possible correlation between word type and sentence type. Therefore, to observe these issues, the data are classified in the following tables.

Table 13. Distribution of cases across sentence type			
	#of cases	*Wh*-questions	*Wh*-relative clauses
TD CHILDREN	97 (100%)	29 (29.9%)	68 (70.1%)
SLI CHILDREN	216 (100%)	150 (69.5%)	66 (30.5%)
TOTAL	313 (100%)	179 (57.2%)	34 (42.8%)

Table 13 shows the number of instances in which *what* and *when* are used as interrogative markers, in *wh*-questions, or as relativizers, in *wh*-relative clauses. As it can be appreciated from the total row, in general terms, these children produce more *wh*-questions (57.2%) than *wh*-relative clauses (42.8%). Therefore, in this regard, what has been previously

mentioned in section 3, about the acquisition and production of *wh*-relative clauses, is confirmed since the production of this sentence type is the lowest one. Likewise, these results significantly vary when analyzing the data per child group.

Consequently, the data per child group are examined and several conclusions can be drawn. On the one hand, the data extracted from the TD children go against the previous generalization about the predominance of *wh*-questions over *wh*-relative clauses. These children produce 70.1% of *wh*-relative clauses and 29.9% of *wh*-questions. On the other hand, the SLI children do meet the previous premise, since their instances of *wh*-relative clauses amount to 30.5% whereas that of *wh*-questions raise to 69.5%. Thus, from these results, a clear correlation can be established. The TD children use either *what* or *when* more in *wh*-relative clauses than in *wh*-questions, unlike the SLI children who incorporate these *wh*-words more in *wh*-questions.

Besides, according to Schuele and Dykes (2005), although these children have a similar MLU range, the TDs produce generally complex syntax and the SLIs tend to use the less complex structures. However, the fact that the SLI participants generate more instances of *wh*-questions may be the consequence of three different factors combined: they would fail in the production of *wh*-relative clauses because these are problematic for them (as they are syntactically more complex than questions because they involve subordination); they do not have properly acquired this construction yet; or it is just as a matter of preference.

Table 14. Distribution of cases across sentence type and *wh*-word						
	Wh-questions			*Wh*-relative clauses		
	Total	*What*	*When*	Total	*What*	*When*
TD CHILDREN	29 (100%)	22 (75.9%)	7 (24.1%)	68 (100%)	0	68 (100%)
SLI CHILDREN	150 (100%)	143 (95.3%)	7 (4.6%)	66 (100%)	0	66 (100%)

Table 14 displays the distribution of cases per sentence type (i.e. *wh*-questions or *wh*-relative clauses) and *wh*-word (i.e. *what* or *when*). As is shown, a correlation between one *wh*-word with one sentence-type can be established, too. Generally, both the TD and the

SLI children tend to use *when* in order to introduce *wh*-relative clauses and *what* to formulate *wh*-questions. The preference of *what* over *when* in *wh*-questions proves that for both child groups it is easier to use *what* in order to ask something since *when* implies a more complex and abstract context. Likewise, in the case of *wh*-relative clauses, there is an absolute predominance of *when* as there are no instances in which these participants employ the *wh*-word *what* as a relativizer. This suggests that all the participants have acquired some knowledge of grammar, as the use of *what* in *wh*-relative clauses is not common, and in some cases could even be ungrammatical.

As can be appreciated from table 14, in the case of the TD children, out of the total number of *wh*-questions produced 75.9% are introduced by *what* and 24.1% by *when*. In regards to the SLI children, these percentages amount to 95.3% in the case of *wha*t and to 4.6% in the case of *when*. Besides, in relation to *wh*-relative clauses, 100% of the time they are introduced by *when*, both in the case of the TD and the SLI children. Thus, it can be concluded that there is a highly prominent and consistent correlation in the use of *what* in *wh*-questions and of *when* in *wh*-relative clauses. Although there are some instances in which *when* has been used as an interrogative marker, too.

6.3.2. Research question 6: distribution of cases across sentence and data type

The last research question deals with the data subclassified in terms of the type of data. This classification demonstrates whether the TD and the SLI children tend to use *wh*-questions and *wh*-relative clauses more in spontaneous or in narrative data.

Table 15. Distribution of cases across sentence and data type						
	Wh-questions			*Wh*-relative clauses		
	Total	Narrative data	Spontaneous data	**Total**	Narrative data	Spontaneous data
TD CHILDREN	**29 (100%)**	14 (48.3%)	15 (51.7%)	**68 (100%)**	47 (69.1%)	21 (30.9%)
SLI CHILDREN	**150 (100%)**	73 (48.7%)	77 (51.3%)	**66 (100%)**	46 (697%)	20 (30.3%)

Table 15 compiles the distribution of cases across sentence type (i.e. *wh*-questions or *wh*-relative clauses) and data type (i.e. narrative or spontaneous). It can be appreciated from this table that, for both child groups, the production of *wh*-questions is higher in spontaneous data as opposed to that of *wh*-relative clauses which is more predominant in narrative data. Moreover, the difference in the case of *wh*-questions is almost irrelevant whereas in *wh*-relative clauses, actually, it is fairly prominent. The TD children use *wh*-relative clauses 69.1% in narrative data and 30.9% in spontaneous data. Similarly, the SLI children's production of *wh*-relative clauses amounts to 69.7% in the former data type and to 30.3% in the later. Consequently, it can be concluded that whenever children are being guided to produce specific structures (as in the narrative data), they tend to use them. By contrast, when they are free to choose their utterances, they do not largely use complex syntax, as *wh*-relative clauses; they would rather use syntactically simpler ones.

Concerning *wh*-questions, the difference in the production per data type is minimum, both in the TD and the SLI participants. However, there is a slight preference in the use of this structure in spontaneous data. The TD children use *wh*-questions in narrative data 48.3% of the times and 51.7% in spontaneous data. Regarding the SLI participants, they produce *wh*-questions in narrative data 48.7% of the cases and 51.3% in spontaneous data. Hence, it can be concluded that in terms of *wh*-questions it does not really matter whether the data are narrative or spontaneous.

All in all, it can be concluded that from these results extracted from the TD and the SLI participants several resemblances in the production of these grammatical issues could be found across the two child groups; although there are still some relevant and prominent differences. These mainly affect the distribution of cases across *wh*-words since there is a clear preference for *when* by the TD participants and for *what* by the SLI children; and the distribution across sentence type as the TD children produce more *wh*-relative clauses, whereas the SLI children *wh*-questions.

6.4. <u>Performance by the TD and SLI children through MLU stages.</u>

The rates at which the target *wh*-words are used in relation to the children's MLUs appear in the following two graphs. These present whether their performance is better at higher or lower MLUs rates, and whether it is constant or it suffers variations.

Graph 1. TD children's data per MLU stages
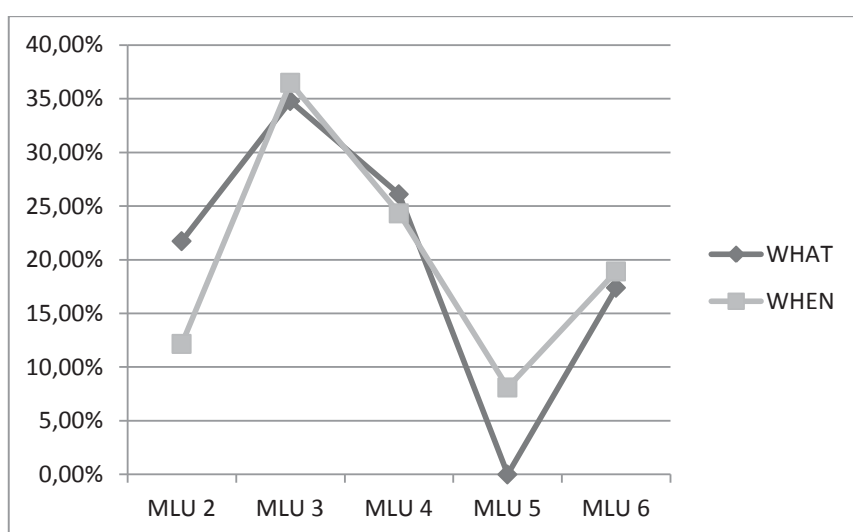
As it can be appreciated from graph 1, the production of *what* and *when* by the TD children follows a similar pattern. In both groups these *wh*-words *what* and *when* emerge at similar MLUs. The TD children experiment a high increase in their productions of the target *wh*-words at MLU 3, that is, between 3.0 and 3.9. At this point, they produce approximately 35% of the *wh*-words. However, after this rise their productions begin to decrease down to a 25% when their MLUs are between 4.0 and 4.9. Then, it continuously decreases around 8%, in the case of *when*, and to 0%, in the case of *what*, when their MLUs range between 5.0 and 5.9. Nevertheless, after this decrease, production once again begins to increase until it reaches almost 20%. These results could suggest that when their MLUs ranged between 3.00 and 3.9 they begin to produce them constantly, since they have acquired them recently.

However, they stopped producing those structures until they have higher MLUs, which may signify that they have properly acquired them and use them constantly and adequately.

Graph 2. SLI children's data per MLU stages
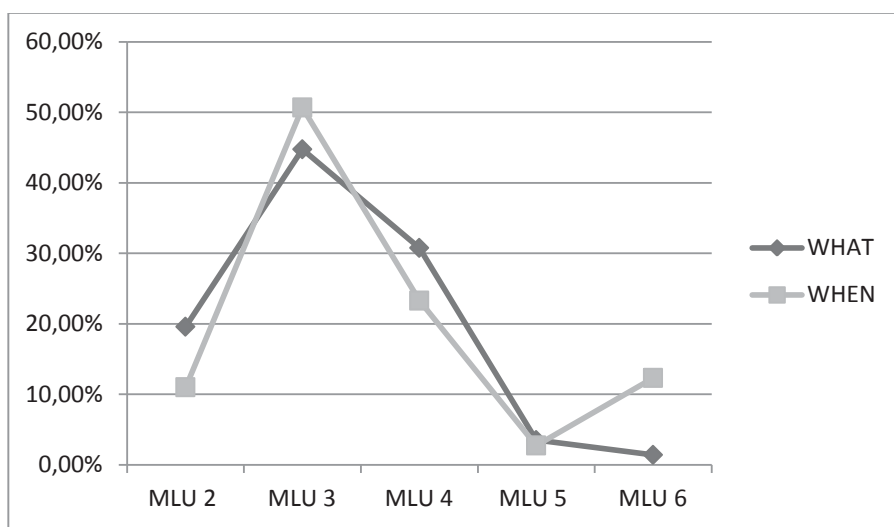
In regard to the SLI children, as graph 2 shows, the production of *what* and *when* in relation to their MLUs follows almost the same pattern as that of the TD children, since variations can be found when their MLUs are at the highest rates. As in the case of the TD children, they experiment a high increase when their MLUs ranged between 3.00 and 3.9. This increase is followed by a high decrease where the production of both *wh*-words is almost nonexistent: 3.50% in the case of *what* and 2.74% for that of *when*. Nevertheless, as opposed to the TD children who experience another increase in the production of both target *wh*-words, the SLI children continue reducing their production of *what* to 1.40% in MLU 5, whereas their production of *when* presents an increased to 12.32%.

In short, it can be concluded that the production of these target *wh*-words by both the TD and the SLI participants is higher when their MLUs ranged between 3.00 and 3.9 and that

their production patterns present the same process with just one variation. When they reach their highest MLUs, the TD children increase their production of both *wh*-words, whereas the SLI children only experience this rise in the production of *when*, as their realization of *what* continues to decline.

7. CONCLUSIONS

This dissertation presents a study on the production of two *wh*-words, *what* and *when*, regarding different grammatical variables (i.e. *wh*-type, sentence type, and data type), as they appear in the production of TD and SLI English children. To conduct this study, the necessary data are extracted from different corpora (i.e. the ENNI corpus, the Conti-Ramsden 3 corpus, and the Conti-Ramsden 4 corpus), all of them available in the online database CHILDES (MacWhinney 2000). After selecting the data, the results are classified following the different grammatical variables and several conclusions are reached in order to answer each of the research questions initially set.

Firstly, the results obtained present unexpected findings in relation to the research question set 1. Based on previous studies about the topic, it was expected that the TD children produce more *wh*-words than their SLI peers. However, the results prove just the opposite: the SLI children use considerately more the target *wh*-words than the TD participants. Likewise, their productions are consistent in relation to the data type: in both cases, they present more instances of *what* and *when* in narrative than in spontaneous data. This fact, as previously explained, could be linked to the different elicitation technique used to collect the data.

Regarding the second research question set, the results depict that there is a clear correlation between *wh*-word (i.e. *what* and *when*) and child group (i.e. TD and SLI children). These results point out that the TD children have a clear preference for the use of *when*, whereas the SLI for that of *what*. This last issue was expected, as the *wh*-word *what*

has been argued to be the easiest one and one of the first to be acquired by children. Moreover, the fact that TD children prefer the use of *when* over that of *what* reflects that, when they have their same MLUs, their productions are better, as the acquisition of *when* has been classified as one of the latest. Besides, in relation to the type of data, the results are consistent with the previous ones in that both the TD and the SLI children are more willing to produce these target *wh*-words in narrative than in spontaneous data.

Additionally, the last question set deals with the distribution of cases across sentence type. As previously explained, the acquisition and, consequently, the production of *wh*-relative clauses could be seen equally problematic to both TD and SLI children, as these structures are seen as complex ones. However, once again, the production by the TD children does not go in line with these previous ideas, since their results present a clear preference for the use of *wh*- relative clauses over that of *wh*-questions. Conversely, these former ideas are backed up by the SLI participants' results. Their production of *wh*-questions stands out on the *wh*-relative clauses ones.

Finally, the results have been analyzed following the different MLU stages established. Thus, in relation to their MLUs, it can be concluded that both the TD and the SLI participants produce the highest amount of the target *wh*-words when their MLUs ranged between 3.0 and 3.9. This could suggest that this is so precisely when they are acquiring all the syntactic properties that define these structures at first instance. Afterwards, again both participant groups experience a diminish in their productions which decrease to almost 0%. What it is even more important to bear in mind is that after this fall in their productions, the TD children show a rise, both of *what* and *when*; however, the production of *what* by the SLI children continues decreasing, whereas *when* begins to rise up to 12%. This last fact is surprising as it is the only point at which the SLI children show a higher production of *when*, resembling, in that way, to the production of TD children.

All in all, after reaching these conclusions, it can be added that these results have been extracted from a reduce number of children and that more grammatical and contextual

properties could be taken into account in order to amplify these results and conclusions. For instance, conclusions could vary in relation to the participants' ages. However, as in this dissertation the main focus has not been the acquisition but their productions, their ages have not been taken as a determining factor, which could be addressed in further studies. Moreover, more studies are needed in order to have a more comprehensive view of the use of these and other *wh*-words in *wh*-questions and *wh*-relative clauses in the production of TD and SLI children. Finally, it is important to keep in mind that sometimes the results go against the expectations. Therefore, further studies on the topic would clarify why this happens and whether this is a general tendency or rather a peculiarity of a specific group of participants.

8. BIBLIOGRAPHY

Biber, D. et *al.* 1999. *Longman Grammar of Spoken and Written English*. Longman. Print.

Bloom, L. et *al.* 1982. *Wh-Questions: Linguistic Factors That Contribute to the Sequence of Acquisition*. "Child Development", 53: 1084-1092. Print.

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

Crain, S and Thornton, R. 1947. *Investigations in Universal Grammar: A Guide to Experiments on the Acquisition of Syntax and Semantics*. Cambridge, MA: MIT Press. Print.

De Villiers, J. and Roeper, T. 1995. *Barriers, Binding, and Acquisition of the DP-NP Distinction*. "Language Acquisition", 4: 73-104. Print.

Flynn, S. and Lust, B. 1980. *Acquisition of Relative Clauses: Developmental Changes in their Heads*. "Cornell University Working Press Papers in Linguistics", 1: 33-45. Print.

Guasti, M. T. 2002. *Language Acquisition. The Growth of Grammar.* Massachusetts Institute of Technology. Print.

Hamburger, H. 1980. *A Deletion Ahead of its Time*. "Cognition", 8:389-416. Print.

Huddleston, R. 1984. *Introduction to the Grammar of English.* Cambridge University Press. Print.

Leonard, L. 1995. *Functional Categories in the Grammars of Children with Specific Language Impairment*. "Journal of Speech and Hearing Research", 38:1270-1282.Print.

---. 1998. *Children with Specific Language Impairment*. Boston, MA: MIT Press. Print.

Limber, J. 1973. "The Genesis of Complex Sentences". In *Cognitive Development and the Acquisition of Language*, ed. T. Moore, 169-185. New York: Academic Press. Print.

MacWhinney, B. 2000. *The CHILDES project* [Child Language Data Exchange System]. Web. https://childes.talkbank.org/. 01 Jun. 2018. Web.

---. 2018. *Tools for Analyzing Talk. Part 2: The CLAN Program*. Carnegie Mellon University. Print.

O'Grady, W. 1997. *Syntactic Development*. The University of Chicago Press. Print.

Quirk, R. et *al.* 1985. *A Comprehensive Grammar of the English Language*. Longman. Print.

Rizzi, L. 1990. *Relativized Minimality*. Cambridge, MA: MIT. Press. Print.

Rowland, C.F. et al. 2003. *Determinants of Acquisition Order in Wh-questions: Re-evaluating the Role of Caregiver Speech*. "Journal of Child Language", 30 (3): 609-635. Print.

Schuele, C.M, and Dykes J. 2005. *Complex Syntax Acquisition: A Longitudinal Case Study of a Child with Specific Language Impairment*. "Clinical Linguistics & Phonetics", 19(4):295-318. Print.

Sheldon, A. 1974. *The Role of Parallel Function in the Acquisition of Relative Clauses in English.*"Journal of Verbal Learning and Verbal Behavior", 13: 272-281. Print.

Universidad de Valladolid. 2009. *Grado Adaptación a Bolonia: Graduado/a en Estudios Ingleses*. Universidad de Valladolid. Print.

Van Der Lely, H.K.J. and Battell, J. 2003. *Wh-Movement in Children with Grammatical SLI: A Test of the RDDR Hypothesis*. "Language", 79 (1):153,181. Print.

Yule, G. 2006. *The Study of Language*. Cambridge University Press. Print.

AFTERWORD: OBJECTIVE REACHED IN THE UNDERGRADUATE DISSERTATION

Before doing this dissertation some objectives and expectations were set based on the official description of the degree (Universidad de Valladolid 2017-2018). Likewise, after having done the whole study, some of the most important competences that have been covered along the course of its elaboration are referred to. These are the following:

- Comprehension of the English grammar and its description.
- Consolidation of the general and specific competences related to the English language and its fields of study (i.e. grammar, linguistics, literature).

In terms of the first reached objective, through the elaboration of this dissertation I have been able to put into practice the knowledge acquired in different courses throughout these last four years. Thus, these courses are:

- Grammatical background. English grammar I, II and III (these courses were taken in the first and second years of the degree).
- Grammar and linguistic background: Syntax III, First Language Acquisition, and Second Language Acquisition (these courses were taken in the third year in an Erasmus exchange program in University College Dublin).
- English Language: Academic and Professional English (second year); and Applied Linguistics III and Information and Communication technologies applied to English Studies (fourth year).

Regarding the second objective, with this dissertation I have been able to extend and consolidate my knowledge on the field of grammatical analyses. This has been acquired not only in the elaboration of the dissertation but also in the whole process carried out: the topic selection, the research process, the data analysis and classification, and the study of the data themselves. Moreover, I have been able to see how these theoretical issues can be taken into account in practical English teaching strategies and methodologies.