The acquisition of complex wh- questions in the L2 English of Canadian French and Bulgarian speakers:
Medial wh- constructions, inversion phenomena, and avoidance strategies

Nikolay Slavkov

Thesis submitted to the
Faculty of Graduate and Postdoctoral Studies
In partial fulfillment of the requirements
For the PhD degree in Linguistics

Department of Linguistics
Faculty of Arts
University of Ottawa

ABSTRACT

THIS DISSERTATION EXAMINES THE DEVELOPMENT OF LONG-DISTANCE WH- MOVEMENT QUESTIONS IN THE L2 ENGLISH OF (CANADIAN) FRENCH AND BULGARIAN SPEAKERS. THE MAIN PHENOMENON UNDER INVESTIGATION IS *MEDIAL WH-* CONSTRUCTIONS (*WH- SCOPE MARKING* AND *WH- COPYING*). SUCH CONSTRUCTIONS ARE OF PARTICULAR INTEREST BECAUSE THEY ARE UNATTESTED IN BOTH THE L1 AND THE L2 OF THE TWO LEARNER POPULATIONS; AT THE SAME TIME, THEY ARE LICENSED OPTIONS IN A NUMBER OF OTHER TYPOLOGICALLY DISTINCT LANGUAGES, OF WHICH THE PARTICIPANTS REPORT NO KNOWLEDGE. AS SUCH, *MEDIAL WH-* CONSTRUCTIONS POSE A LEARNABILITY PROBLEM IN L2 ACQUISITION: HOW CAN A LEARNER "KNOW" SOMETHING THAT IS NOT SUPPORTED BY EITHER THE NATIVE LANGUAGE OR THE TARGET INPUT, BUT IS ATTESTED IN OTHER LANGUAGES?

TWO EXPERIMENTS, A WRITTEN GRAMMATICALITY JUDGMENT MULTIPLE-CHOICE TASK AND AN ORAL ELICITED PRODUCTION TASK, WERE CARRIED OUT WITH THE TWO DIFFERENT LEARNER POPULATIONS AND WITH ENGLISH NATIVE SPEAKER CONTROLS. THE WRITTEN EXPERIMENTS SHOWED THAT *MEDIAL WH-* CONSTRUCTIONS CO-EXIST AND COMPETE WITH THE TARGET ENGLISH LONG-DISTANCE STRUCTURE AT THE EARLY AND INTERMEDIATE STAGES OF ACQUISITION OF BOTH THE FRENCH AND THE BULGARIAN SPEAKING PARTICIPANTS; AT THE ADVANCED STAGES OF ACQUISITION, BOTH POPULATIONS SHOWED EVIDENCE THAT *MEDIAL WH-* REPRESENTATIONS HAD BEEN SUCCESSFULLY ELIMINATED FROM THE INTERLANGUAGE GRAMMAR, AND THE L2 DATA CONVERGED WITH THAT OF THE NATIVE SPEAKERS. IN THE ORAL ELICITATION EXPERIMENTS BOTH THE FRENCH AND THE BULGARIAN SPEAKING PARTICIPANTS RESORTED TO *MEDIAL WH-* AND A NUMBER OF OTHER STRATEGIES AIMED AT AVOIDING LONG-DISTANCE WH- MOVEMENT; I ARGUE THAT SUCH STRATEGIES ARE DUE TO BOTH THE DERIVATIONAL COMPLEXITY AND THE HIGH PROCESSING LOAD ASSOCIATED WITH LONG-DISTANCE WH- MOVEMENT.

THE ACCOUNT DEVELOPED TO ADDRESS THE FINDINGS OF THE DISSERTATION INCORPORATES INSIGHTS FROM BOTH NATIVIST AND DOMAIN-GENERAL VIEWS ON ACQUISITION. THE PROPOSAL IS THAT L2 GRAMMARS HAVE TO BE UG-CONSTRAINED IN ORDER FOR THE LEARNABILITY PROBLEM TO BE RESOLVED. IN ADDITION, THE ACQUISITION PROCESS HAS TO BE STRONGLY DRIVEN BY THE INPUT, ALLOWING LEARNERS TO MAKE EXTENSIVE USE OF A GENERAL PROBABILISTIC LEARNING MECHANISM; THIS MECHANISM HELPS THEM TO GRADUALLY ELIMINATE THE COMPETING REPRESENTATIONS UNSUPPORTED BY THE L2 INPUT AND TO CONVERGE WITH THE GRAMMATICAL TARGET. THIS APPROACH IS IN PRINCIPLE APPLICABLE TO BOTH L1 AND L2 ACQUISITION AND ACCOUNTS FOR SOME RELEVANT SIMILARITIES BETWEEN THE TWO.

THESIS ADVISORS: JUANA LICERAS   NINA KAZANINA
                      (UNIVERSITY OF OTTAWA)   (UNIVERSITY OF BRISTOL)
EXTERNAL EXAMINER: LYDIA WHITE (MCGILL UNIVERSITY)
INTERNAL EXAMINERS: MARÍA LUISA RIVERO (UNIVERSITY OF OTTAWA)
                      ASH ASUDEH (CARLETON UNIVERSITY)
                      ANDRÉS SALANOVA (UNIVERSITY OF OTTAWA)

Table of Contents

List of Tables

List of Figures

List of Appendices

List of Abbreviations

ACC	accusative
Agr	agreement
AH	accessibility hierarchy
BG	Bulgarian
C	complementizer
DCH	derivational complexity hypothesis
DCM	derivational complexity metric
DO	direct object
DP	determiner phrase
DTC	derivational theory of complexity
ECP	empty category principle
EF	edge feature
EPP	extended projection principle
FF	formal feature
FOC	focus
FR	French
GB	Government and Binding
Gen	genitive
IL	interlanguage
IM	internal merge
IO	indirect object
L1	first language
L2	second language
L2A	second language acquisition
LD	long distance
LF	logical form
MSIH	missing surface inflection hypothesis
NAD	non-adjacent dependencies
NL	native language
NOM	nominative
OT	optimality theory
PF	phonological form
prt	participant
PV	preverbal or verbal modifier
Q	Q morpheme (silent scope marker)
RI	root infinitive
SLA	second language acquisition (see also L2A)
Subj/ Su	subject
T	tense
TL	target language
WM	working memory
WS	Williams Syndrome

Acknowledgements

My deepest gratitude goes to my two advisors and friends Juana Liceras and Nina Kazanina. Their combined efforts have been an invaluable contribution to this dissertation and to my research profile in general.

I met Juana during the first year of my Ph.D. program at the University of Ottawa and was impressed by the detailed feedback that she would give even on minor assignments in the classes that she taught. Since then I have continued to receive such feedback on my papers, comprehensive exams, independent studies and, of course, this dissertation. Juana's generosity with her time and the fact that she never seems to get stressed – "just busy" as she puts it – is truly amazing and a great quality for an advisor. I have also benefited from Juana's ability to look at the big picture and from the dose of healthy scepticism which she applies to all research. Juana has been able to put things into perspective for me by bringing up older references and at the same time by being up-to-date with newer proposals. Her friendship and continual support, both administrative and financial, have contributed a lot to making my doctoral program a positive experience. Regardless of the circumstances, Juana has always been there for me, and I will always be grateful for this.

Nina came to the University of Ottawa in my third year and infected our department with her energy and enthusiasm about everything related to linguistics and beyond. I was not spared and my dissertation planning received a big boost as a result of just meeting her and working as a teaching assistant for her. Nina quickly became one of my best friends and volunteered her help with my experimental design a long time before she officially joined my dissertation committee as an advisor. Her knowledge about both experimental work and linguistic theory is impressive, but what really amazes me are her daring and sharp analytical abilities in any field or issue that she may come across. Nina's openness and breadth both as a researcher and as a person is something truly unique that I have been very lucky to be exposed to. Her comments and suggestions on this dissertation were insightful and timely, which are essentially the two things that you need from an advisor. However, I should also thank Nina for other support, such as coaching me through and giving me confidence in doing research, in preparing for conference talks, and in writing. Finally, I am ever grateful to her for staying on with me until the completion of my dissertation, even after she left the University of Ottawa.

Another person who has been a great influence on me is María Luisa Rivero. Marisa has always been available to read over drafts and to give invaluable theoretical advice on this dissertation; in addition to this, she has also offered useful guidance and support in my career planning. In the earlier stages of my studies, she has helped shape my interests in Slavic linguistics, encouraged me to publish in the field of Slavic syntax, and provided me with very generous financial support. Marisa is also one of the most engaging and organized lecturers I have seen and I must acknowledge that I have "stolen" some of her pedagogical techniques and used them in my own teaching. Finally, Marisa's serious academic attitude is complemented by a subtle sense of humour which somehow always comes as a surprise. This gives her a unique charm and makes interacting with her – which is always very useful professionally – a great pleasure.

Chapter I
Introduction

- ▪ SUMMARY: THIS CHAPTER PROVIDES AN OVERVIEW OF SOME OF THE GENERAL ACQUISITION ISSUES AS WELL AS THE SPECIFIC L2 PHENOMENA THAT THE DISSERTATION ADDRESSES.

Non-target patterns in L2 acquisition are an extremely interesting and, at the same time, challenging phenomenon.[1] Seeking consistent errors within an individual or a population of L2 learners and trying to account for such systematicity is a key goal in this dissertation and, indeed, in most other L2 research. The great challenge in this effort is that L2 errors can arise from a variety of sources: different feature configurations between the L1 and the L2, morphological mismatches, processing constraints, age effects, motivational factors, etc. How, then, are we to tease apart the different potential factors conspiring together to bring about a particular non-target L2 behaviour? Are there non-target phenomena that allow us a glimpse into the acquisition process without the confounds just described? Could such insights be generated from spontaneous L2 production or under controlled, experimental circumstances?

The phenomenon that this dissertation investigates is *medial wh-* constructions in the L2 English of (Canadian) French and Bulgarian speakers. Before I introduce *medial wh-* constructions in more detail, I will explain how they belong to a particular class of L2 phenomena that pose a special challenge in L2 acquisition theory, and are therefore of special interest. Generally, non-target utterances can be linked to some properties of the L1 or the L2. On rare occasions, however, such L2 utterances seem completely unrelated to any properties of the native or the target language, but surprisingly are attested in other natural languages of which the L2 learners do not have any knowledge. This constitutes an obvious learnability problem in second language acquisition: how can an L2 learner 'know' something that does not exist in the L1 or the L2, but is typologically attested in other languages? Several phenomena described in the L2 literature exemplify this problem. One such case is the existence of resumptive pronouns in relative clauses as in (1).

---

[1] In this chapter a *non-target* pattern is broadly defined as any pattern (morphological, syntactic, phonological, etc.) that deviates from what native speakers of the particular target language (variety) would produce and/or accept. However, later on in this dissertation the terms *target* and *non-target* are used focusing on purely syntactic phenomena and morphological, lexical, etc. errors are ignored (see chapter 4).

(1) a. *This is the woman who(m) I love **her**.
     b. 'This is the woman who(m) I love.'

Equivalent utterances were reported by Hyltenstam (1984) in situations where neither the L1s (Finnish; Spanish) nor the L2 (Swedish) allow this specific kind of resumption. However, such structures are attested in other languages such as Persian and Arabic.

Another example where non-target utterances do not seem to be licensed by either the L1 or the L2 is offered by MacLaughlin (1996; 1998; see also Finer and Broselow 1986), who investigates reflexives in the L2 English of Chinese and Japanese speakers. She argues that a subset of the participants in her study systematically accepted long-distance binding of reflexives out of non-finite clauses; this setting is available in Russian, for example, but not in either the learners' L1s or in the L2. Schwartz and Sprouse (1994) discuss yet another example that belongs to this class of phenomena. They claim that a Turkish learner of L2 German went through a stage of acquisition at which he produced inverted V2 structures with pronominal subjects and non-inverted structures when the subject was a full DP. In developing an account in terms of cross-linguistic variation of case feature checking mechanisms, the authors argue that the systematic differentiation between pronouns and full DPs with regards to inversion is not found in either the native or target language of the learner. However, such differentiation is found in French, a language of which the speaker reports no knowledge.

More details about these data and analyses will be provided in the next chapter. At this point, however, it is important to mention some of the questions that the above examples raise. Why is it that L2 learners sometimes show systematic non-target language patterns that cannot be explained in terms of either the L1 or the L2, but can be found as grammatical options in other natural languages? Could it be that L2 learners, just like child L1 learners, explore grammatical options made available to them by Universal Grammar and gradually determine which ones are licensed by the particular language they are acquiring? Alternatively, could there be processing reasons that can be invoked to account for such learnability issues? Finally, could the learnability problem result merely from a coincidental combination of, or a mismatch between abstract L1 and L2 features?[2] These are some of the fundamental questions this dissertation will address.

---

[2] See chapter 2, section 2.1.4, for a specific example illustrating this possibility.

As already mentioned, the phenomenon on which I focus is known as *medial wh-,* and what makes its investigation particularly worthwhile is that it constitutes another example of the L2 learnability problem. In essence, *medial wh-* constructions are long-distance wh- questions (i.e. wh- extraction out of an embedded clause), where an "extra" wh- word appears in the beginning of the embedded CP, as in (2)-(3). *Medial wh-* constructions are divided into two types. The first one is known as *wh- copying* because an exact copy of the matrix wh- word appears in the embedded clause, as in (2). The second type is called *wh-scope marking*[3], as in (3); in this construction the contentful wh- word appears only in the embedded clause, while an expletive *what* is inserted in the matrix to mark scope over the entire sentence[4]. Both *wh- copying* and *wh- scope marking* are ungrammatical in all varieties of English. The grammatical equivalent of these *medial wh-* constructions is given in (4) where a single (contentful) wh- phrase has undergone successive cyclic movement to the specifier of the matrix clause.

(2) ***Who**$_i$ do you think **who**$_i$ John kissed t_i?
(3) ***What** do you think **who**$_i$ John kissed t_i?
(4) **Who**$_i$ do you think t_i John kissed t_i?

Despite the fact that *medial wh-* as in (2) and (3) is ungrammatical in adult English, Thornton (1990) attested such sentences in elicited production experiments in child L1 English. Similar utterances have also been observed in the L1 acquisition of Dutch (van Kampen, 1997), French (Oiry, 2002; Oiry and Demirdache, 2006), Spanish (Gutierrez, 2004a; 2006) and Basque (Gutierrez, 2004b), even though *medial wh-* is not part of the adult grammar in these languages either.

More recently, *medial wh-* has also been reported in the L2 acquisition literature (mostly in elicited production experiments): Okawara (2000), Wakabayashi and Okawara (2003), Yamane (2003), and Schulz (2006) attested such structures in the L2 English of Japanese speakers; and Gutierrez (2005) observed *medial wh-* utterances in the L2 English of Spanish/Basque bilinguals. Interestingly, while *medial wh-* constructions are not a grammatical option in any of the languages mentioned so far, they are a well-

---

[3] Sometimes also referred to as *partial movement.*
[4] There is a debate in the literature on whether the wh- element in the matrix clause is an inserted expletive or a phonological realization of a feature that has moved to Spec matrix CP. See next chapter for more on this issue.

documented grammatical phenomenon in other natural languages such as (dialects of) German, Frisian, Hungarian, Romani, Afrikaans, Hindi, etc.[5]

In terms of the learnability problem described above, and working within a generative framework, it is tempting to evoke Universal Grammar in proposing an explanation for *medial wh-* constructions in language acquisition. It can be argued that since this phenomenon is attested in a number of natural languages, it represents a valid option in UG; this option is available and can be selected in the acquisition process regardless of whether it is licensed or not in the target language; eventually, *medial wh-* utterances disappear from the target language, if they are not part of its grammar, due to indirect negative input (i.e. *medial wh-* utterances are absent from the input). Such a scenario can in principle apply to both first and second language acquisition, and a version of it has been proposed by most of the researchers mentioned above (but see discussion of Schulz (2006) in Chapter 2 for an example of a different account). Naturally, this scenario would need to address questions such as why *medial wh-* seems to be more likely to occur under experimental circumstances rather than in spontaneous production, and why exactly *medial wh-* (as opposed to wh- *in situ*, for example) was attested in the experiments mentioned above.

Furthermore, while the assumption that UG plays a crucial role in L1 acquisition is in general well-established, it is subject to much more controversy in L2 acquisition. Some researchers argue that L2 learning is UG-constrained (e.g. White 1985, 1989a, 2003; Schwartz and Sprouse 1996; Flynn 1987, Epstein, Flynn and Martohardjono 1996; Gregg, 2003, inter alia), while others claim that the mechanisms involved in L1 and L2 acquisition are entirely different and UG has very limited or no role in the latter (e.g. Bley-Vroman, Felix and Ioup 1988, Bley-Vroman 1990, Bates and MacWhinney 1989, Clahsen and Muysken 1986, Schachter 1988, among others; see also O'Grady 2001b, 2008, among others). Proponents of this view usually explain L2 acquisition in terms of more general cognitive rather than strictly linguistic processes. While at first sight the class of L2 non-target phenomena that show no relation to the properties of the native or the target language seem to be most naturally explained in terms of a UG-constrained L2

---

[5] Note that Schulz (2006) also attested *medial wh-* constructions in the L2 English of Germans speakers. However, this fact does not pose a learnability problem as such constructions are licensed in (dialects of) German and can thus be straightforwardly accounted for in terms of L1 transfer.

model, they can also be accounted for assuming a non-access view. For example, it could be argued that such phenomena result from specific L2 processing constraints or memory load effects. In addition, although unlikely with the particular construction under investigation in this dissertation, it is possible that different L2 learning strategies, or explicit instruction might also be responsible for the emergence of such paradoxical phenomena.

The issue of whether or not L2 acquisition involves access to UG is further confounded by the high number of additional variables involved in L2 learning. Age of exposure, L1 transfer, length of studies, type of acquisition setting, and motivation are just a few of the many factors that distinguish L2 from L1 acquisition. Moreover, the different degrees of success in language learning as well as the greater variability generally attested in L2 studies compared to L1 studies, make the task of providing empirical support for or against access to UG in L2 even more challenging.

One of the general goals of this dissertation is to provide new empirical data and a discussion relevant to the issues outlined above. An important task in this respect is to address the question of what type of data can support the claim that an L2 grammar is constrained by UG and what kind of evidence would point to the null hypothesis (i.e. that UG is not involved in L2 acquisition). A more specific goal, on the other hand, is to demonstrate that *medial wh-* is attested in the interlanguage of (Canadian) French and Bulgarian speakers acquiring L2 English. A crucial question that I investigate in this respect is whether *medial wh-* in the L2 English of these two populations represents a *direct* or an *indirect dependency*. The *direct dependency* $[_{CP}$ wh_i-… $[_{CP}$ wh_i-…$t_i]]$ constitutes a complex (biclausal) question with long-distance (LD) wh- movement. In this case, pronouncing an intermediate wh- word in the embedded clause is indeed ungrammatical in the two L1s (French; Bulgarian) and the L2 (English), and is of great interest with regards to the learnability problem discussed earlier. On the other hand, the *indirect dependency* $[_{CP}$ wh-$_i$… $t_i]$ $[_{CP}$ wh-$_j$… $t_j]$ constitutes two independent local movements corresponding to two simple (monoclausal) sequential questions with separate wh- phrases, as in (5)-(6).

(5) **Who** do you think… **who** did John kiss?
(6) **What** do you think? **Who** did John kiss?

Such questions can occur in both the L1 and the L2 of the two populations, and therefore do not bear relevance to the L2 learnability problem. However, they are still of interest as it could be argued that L2 learners go through a stage where they favour sequential questions and avoid long-distance movement because the latter is a more complex syntactic operation and is more difficult to process.

In addition to the question of the nature of the dependency, this dissertation also focuses on the inter-relation between *medial wh-* and the complementizer *that*. In particular, I investigate whether *medial wh-* utterances with an overt complementizer, as in (7)-(8), are more or less likely to occur than utterances without such complementizer, as in (2)-(3).

> (7)   *__Who__ do you think **__that__ who** John kissed?
> (8)   *__What__ do you think **__that__ who** John kissed?[6]

This comparison arises from Thornton's 1990 pioneering research on this topic where she proposed that *medial wh-* words in child L1 English occupy complementizer position. Thus, if a *medial wh-* word and an overt complementizer *that* compete for the same position in L2 English, utterances like (7) and (8) should be unlikely to occur or even impossible.

Another important issue that I investigate is whether negation in the matrix clause poses an intervention effect on *medial wh-* utterances in the L2. It has been noted in the literature on German that negation in the matrix clause renders *medial wh-* ungrammatical, as in (9), even though *medial wh-* utterances without negation are licensed in (dialects) of this language, as in (10).

> (9)   *_Was_ _glaubst_ _du_   __nicht__   _mit_   _wem_ _Maria_   _gesprochen_   _hat?_
>   **what** believe you   **not**   **with who** Maria   spoken   has
>   Who do**n't** you believe Maria spoke to?'
> (10) _Was_ _glaubst_ _du_   _mit_   _wem_ _Maria_   _gesprochen_   _hat?_
>   **what** believe you   **with who** Maria   spoken   has
>   'Who do you believe Maria spoke to?'

Based on this observation for German, I test whether negation might have similar effects on *medial wh-* in the L2 English of French and Bulgarian speakers.

---

[6] For a detailed explanation of why the complementizer *that* appears to the left of the *medial wh-* word in these examples, see section 3.2.4.2.

To summarize, this chapter provides an overview of both the general and the more specific questions that the dissertation addresses. These will be re-introduced and elaborated upon in the following chapters.

Chapter II
Background

- SUMMARY: THIS CHAPTER BEGINS BY DISCUSSING CASES IN THE ACQUISITION LITERATURE WHERE AN L2 PHENOMENON CANNOT BE ACCOUNTED FOR IN TERMS OF THE PROPERTIES OF EITHER THE NATIVE OR THE TARGET LANGUAGE; THIS IS FOLLOWED BY AN OVERVIEW OF THE TYPOLOGY OF WH- QUESTIONS, PAYING PARTICULAR ATTENTION TO *MEDIAL WH-* CONSTRUCTIONS; FINALLY, THE CHAPTER LOOKS AT THE *MEDIAL WH-* LITERATURE IN L1 AND L2 ACQUISITION.

2.1. Cases of L2 patterns showing properties of neither the L1 nor the L2

As Eckman (2004) points out, cases of consistent L2 error patterns that cannot be accounted for in terms of either the native language (NL) or the target language (TL) of the learners are of particular interest in L2 research:

> One of the most interesting phenomena in second language acquisition (SLA) is a pattern of learner utterances—often, though not always, errors—in which the regularity is not explainable either in terms of the native language (NL) or the target language (TL), but is nevertheless an attested pattern found in other human languages. (p. 89)

Typically, such phenomena are interpreted as evidence for access (full or partial) to UG in L2 acquisition. However, it is also possible to explain them in terms of L2 processing constraints or other mechanisms in combination with or independent of UG. Below is a (non-exhaustive) survey, of some of these phenomena and the issues they raise.[1]

2.1.1. Resumptive pronouns

Hyltenstam (1984) discovered that second language learners systematically produce resumptive pronouns in relative clauses, even when such pronouns are not licensed in either the L1 or the L2. The target language in his study was Swedish, and the four L1s were Finnish, Spanish, Greek and Persian. Note that resumptive pronouns in a relative clause can in principle occur in a number of different positions along what is known as the *Accessibility Hierarchy* (Keenan & Comrie 1977; later revised by Lehmann 1986; Croft 1990). The original version of the AH is represented in (1), and the idea behind it is that the degree of difficulty with which relative clauses are formed *increases* from left to right; on the other hand, the likelihood of availability of a particular type of relative clause in a given language *decreases* from left to right.

---

[1] Some of the discussion in the following sections is based on White's (2003) review of L2 patterns that reflect parameter settings of neither the L1 nor the L2.

(1) Accessibility Hierarchy (AH)
    Subj > DO > IO > Oblique > Gen > Object of a comparative

A resumptive pronoun in any position of the AH is ungrammatical in English, as illustrated in (2)-(7).

(2) *There is the woman *who she is my sister. (subject)*
(3) *There is the woman *who(m) I registered her. (direct object)*
(4) *There is the woman *to whom I sent her an application. (indirect object)*
(5) *There is the woman *whom I read about her in the newspaper. (oblique)*
(6) *There is the woman *who her sister graduated last year. (genitive)*
(7) *There is the woman *who I am older than her. (object of a comparative)*

(examples adapted from Eckman 2004)

Resumptive pronouns in any of the AH positions are also ungrammatical in Swedish, the L2 in Hyltenstam's study, as well as in two of the L1s: Finnish and Spanish. However, the other two L1s, Greek and Persian, allow or require resumptive pronouns in several different positions of the AH. The total number of speakers of the four languages was 45, and the author used a picture-naming task to elicit relative clauses in L2 Swedish. The results showed that participants from all four speaker populations produced resumptive pronouns in Swedish relative clauses. Of particular interest are the results of the Spanish and the Finnish participants because in their case resumptive pronouns in relative clauses are not licensed in either the NL or the TL.[2] The results were especially systematic among the Spanish group, where eleven out of 12 subjects consistently produced resumptive pronouns in at least one kind of relative clause listed in (2)-(7), and the majority of the subjects produced such pronouns in several different kinds of relative clauses. As predicted, the Greek and the Persian groups also produced resumptives widely and systematically, presumably due to L1 transfer; but what is especially interesting is that some participants produced resumptive pronouns in contexts where their L1 does not require or does not allow them. Overall, these data seem to suggest that L2 learners have access to and make use of grammatical options provided by the languages of the world, regardless of whether these options are instantiated in either the NL or the TL. Such data are often considered to be evidence that adult L2 learners resort to the same mechanisms of language acquisition as child L1 learners. An important assumption in such claims is that both children acquiring their L1 and children or adults

---

[2] It has been shown that Spanish allows some resumptive pronouns but they appear in different contexts.

acquiring a second language would not produce structures that are outside of the realm of UG; this assumption finds support in Hyltenstam's data (see also chapter 4).

In a recent interpretation of Hyltenstam's study, Eckman (2004) proposes a constraint-based account of resumptives using the *Optimality Theory* (OT) framework. Focusing on the question why second language learners produce constructions which cannot be accounted for in terms of either the L1 or the L2, but are attested in other natural languages, the author assumes the following: a) the grammars of all languages are subject to the same inventory of universal constraints but differences between languages occur as a result of different ranking of the constraints; b) if a given constraint contains a hierarchy of sub-constraints (e.g. the AH described above), the ordering of these sub-constraints remains invariable across languages (i.e. a version of the AH would be universal); c) the initial state of interlanguage grammars is dominated by a set of markedness constraints, which ultimately have stronger determining power over the nature of L2 utterances than the effects of L1 transfer. This last assumption is crucial in explaining the puzzling resumption phenomena. Essentially, the idea, shared by both Hyltenstam and Eckman, is that IL grammars are initially characterized by unmarked structures, even though both the NL and the TL may contain their marked counterparts. Based on this assumption, Eckman proposes a family of constraints, as in (8), penalizing constructions that contain non-adjacent dependencies caused by relative clause formation, reordering of constituents, co-referential relationships in binding, etc.

   (8) *Non-Adjacent Dependencies (*NAD): penalize structures with non-
        adjacent dependency relations.

The proposal is that in the initial state of the interlanguage grammar, the non-adjacent dependencies constraint outranks all other constraints, including the ones operative in the NL grammar. In other words, L1 transfer is outranked by *NAD, as illustrated in (9).

   (9) *NAD >>   NL ranking

This constraint ordering is what accounts for structures not attested in the L1 or the L2. Since non-adjacent dependencies create gaps that are difficult for the learners, structures containing such gaps are expected to be avoided in the initial stages of acquisition. In the specific case under consideration, the formation of relative clauses creates such gaps, and

depending on the kind of relative clause listed in the AH, the degree of difficulty varies. Since non-adjacent dependencies are penalized by the system, the learners find a way of filling or "bridging" them by using resumptive pronouns, despite the fact that such pronouns are not licensed in either the NL or the TL. The more difficult the relative clause along the AH, the more likely it is that the non-adjacent dependency will be penalized (i.e. AH dictates a particular order of sub-constraints) and that a resumptive pronoun would be used. This explains the general pattern observed in the data, whereby resumptive pronouns occurred more often in the lower positions of the AH (i.e. the more difficult and more marked types of relative clauses).

Note that this account is essentially a processing one. It assumes that during the initial stages of acquisition learners try to shorten long-distance dependencies by filling the gaps they create. Thus, even though phenomena which cannot be accounted for by the properties of either the L1 or the L2 are typically taken for evidence that the learners are accessing UG in the acquisition process, Eckman makes a convincing case for a processing explanation. He argues that the proposed constraint types and ordering fits well within the literature that accords a significant role to processing in explaining L2 phenomena (e.g. Clahsen 1984; Pienemann 1989; O'Grady 2001a&b, 2008, among others). Interestingly, while some researchers argue that processing effects show important differences in the acquisition mechanisms used by child L1 learners and adult L2 learners, Eckman suggests that within the markedness framework incorporated in OT, interlanguage grammars are constrained in the same way as L1 grammars. He further argues that IL grammars and L1 grammars should be viewed as the same kind of systems, obeying the same kind of general principles and differing systematically from each other in the same way natural languages differ from each other.

Eckman's (2004) proposal is relevant to this dissertation in several ways. Apart from providing an example of an L2 phenomenon unavailable in both the NL and the TL, it also evokes a processing account of shortening long-distance dependencies by inserting or overtly pronouncing material in syntactic gaps. This idea is applicable to the empirical data that this dissertation provides and will be revisited and evaluated in chapters 3 and 4.

2.1.2. Reflexives

Another case of systematic L2 errors which cannot be explained on the basis of either the L1 or the L2 is discussed by MacLaughlin (1996, 1998). The author investigated the acquisition of reflexives in L2 English by Chinese and Japanese speakers, and argued that a subset of the participants restructured their grammar to a setting unavailable in either the L1 or the L2, but attested in another natural language, Russian. The key facts about MacLaughlin's claim are as follows. Crosslinguistically, reflexives can be of two types with regards to their morphological properties: simplex (monomorphemic, i.e. a head, X^0) or complex (polymorphemic, i.e. a phrase, XP). English has only polymorphemic reflexives (e.g. *myself*, *himself*, etc.). Chinese and Japanese, on the other hand, have both polymorphemic and monomorphemic reflexives. Complex reflexives must always be bound locally. Simplex reflexives, however, can engage in either local or long-distance binding, depending on whether the particular language expresses overt morphological agreement or not (see also Bennet 1994; Progovac 1992, 1993). In cases where agreement is never overt, like in Chinese and Japanese, long-distance binding of simplex reflexives is always possible. In languages like Russian, however, morphological agreement is expressed overtly in finite clauses, but not in non-finite clauses. Therefore, Russian allows long-distance antecedents for monomorphemic reflexives, but only out of non-finite clauses.

MacLaughlin's experiment involved a co-reference judgement task in which the Chinese and Japanese participants were asked whether or not a reflexive can refer to different antecedents in a sentence in English. The results showed the following: 6 out of the 15 participants were treating the reflexives in accordance with the TL settings (i.e. local antecedents); two of the fifteen participants applied the settings for monomorphemic reflexives from their L1 (i.e. long-distance antecedents); and seven participants showed a pattern characteristic of Russian (i.e. long-distance binding of reflexives only in the case of non-finite clauses). Since both the type of reflexive (simplex or complex) and the presence or absence of overt morphological agreement play a role in the distribution of reflexives across languages, MacLaughlin suggests that the subset of participants who showed the Russian pattern mistakenly treated English reflexives as monomorphemic, and at the same time correctly identified English as a language with (some)

morphological agreement in finite clauses and null agreement in non-finite clauses. Thus, these participants arrived at a setting which is unavailable in either their L1 or L2, but allowed by the principles of UG and attested crosslinguistically.

White (2003), while considering MacLaughlin's data to be evidence of access to Universal Grammar in second language acquisition, raises the legitimate question of why certain learners arrive at this new setting, which is not instantiated either in the native language or the target input. Essentially, what MacLaughlin's participants seem to have done is correctly re-set the agreement parameter to the English setting but at the same time they have failed to restructure the reflexive parameter to the polymorphemic-only setting. According to Progovac and Connell (1991), speakers of languages that do not realize morphological agreement overtly can discover that English does have overt agreement from input containing third-person-singular agreement. However, by the same logic, the learners should be able to discover from the input that English reflexives are polymorphemic, which, as White points out, seems to be a fairly transparent fact.

Despite these unresolved questions, MacLaughlin's data are still seen as a strong case for access to UG in L2 acquisition. Further evidence in favour of this argument comes from McDaniel, Cairns and Hsu (1900) and Thomas (1994), who claim that children acquiring English as their first language also initially fail to recognize reflexives as polymorphemic and to restrict them to local antecedents only.

The main point that this research demonstrates is that L2 learners can be sensitive to complex abstract linguistic properties whose source is difficult to trace. The *medial wh-* data that will be presented in chapter 3 of this dissertation will show evidence of similar abstract grammatical knowledge in the L2 English of (Canadian) French and Bulgarian speakers and an appropriate explanation of its source will be pursued at that point.

2.1.3. Case checking

Another example of an L2 pattern that is not attested either in the NL or the TL is offered by Schwartz and Sprouse (1994) who examine production data from an adult Turkish speaking learner of German as a second language. The authors divide the speaker's L2 development into several distinct stages, claiming that during each of them

the interlanguage obeys a number of different constraints or makes use of different syntactic mechanisms. In stage one, the speaker produces SOV order, which is typical of both Turkish and German. In stage two, the speaker begins to acquire the V2 order (a main clause structure in which a constituent other than the subject is fronted, and the verb is obligatorily placed in second position). This order is characteristic of German and some other Germanic languages, but not of Turkish. Thus, it appears that at this stage the Turkish speaker of L2 German is in the process of restructuring his grammar in a way that meets the requirements imposed by the L2 word order. Interestingly, however, the speaker consistently makes an extra distinction with regard to V2, which German does not have: in the cases where the target V2 order is produced, the subject position is almost always filled by a personal pronoun (10); on the other hand, when the subject position is filled by a full DP, the learner consistently fails to place the verb in second position (11), which is ungrammatical in German.

(10) *Dann trinken* **wir** *bis neun Uhr.*   → *pronominal subject, Target V2*
     then drink   we   until nine o'clock
     'Then we will drink until nine o'clock.'

(11) **In der Türkei* **der Lehrer** *kann den Schüler schlagen.*   →   *full DP, V2 fails*
     in the Turkey the teacher can the pupils beat
     'In Turkey, teachers can hit pupils.' (adapted from Schwartz and Sprouse 1994)

The account that Schwartz and Sprouse propose in explaining the puzzling distinction that the learner applied with regards to pronouns and full DPs is in terms of crosslinguistic mechanisms of case feature checking. Since DPs have case features that need to be checked, different languages provide various options of meeting this requirement. Assuming that case feature checking in German is achieved through agreement between a specifier and a head, a subject DP should raise to AgrP to check its case feature against a corresponding Agr feature. Schwartz and Sprouse further assume that the same mechanism must also be used in Turkish. However, a different mechanism, namely incorporation, is available in French (Rizzi and Roberts 1989). In this language, subject clitics are incorporated into the verb in order to check case features. As Schwartz and Sprouse suggest, the Turkish learner of L2 German, seems to have arrived at the French option of subject clitic incorporation. This would explain his production of inverted structures with a pronominal subject as in (10), and uninverted ones with a full

DP as in (11), where the incorporation option is in principle impossible. The authors claim that in the next stage of L2 development, the learner acquires the mechanism that enables him to achieve the TL setting, regardless of whether the subject is pronominal or a full DP.

Schwartz and Sprouse's proposal assumes that such data are best explained in terms of a UG-constrained view of adult L2 acquisition. White (2003) also supports such an account, while once again pointing out that the reasons for the Turkish learner to go through an incorporation analysis of German pronominal subjects in target V2 contexts remain unclear. Questions such as whether the L2 input might contain a trigger for the incorporation mechanism, or whether some unknown properties of the L1 can somehow lead the learner to this unexpected analysis are still to be resolved.

Once again, a parallel can be drawn between the above data and the results on *medial wh-* in the L2 English of French and Bulgarian speakers that will be presented in the following chapters. Crucially, the learners are able to "unlearn" non-target constructions in the later stages of acquisition of the L2, despite the lack of any obvious triggers or direct evidence against such constructions in the input. The details of a proposal accounting for such facts will be presented in chapter 3 and then reinforced in chapter 4.

2.1.4. Null subjects

The last example in this section comes from null subjects and is indicative of an interesting controversy in the literature. It is based on a study by Clahsen and Hong (1995) meant to show evidence for failure of parameter re-setting and lack of real access to UG[3] in L2 acquisition. The study is subsequently re-interpreted by White (2003), however, who claims that the results actually provide support for fully operative UG in L2, and represent another example of an L2 phenomenon that cannot be explained in terms of the properties of either the NL or the TL.

---

[3] Clahsen and Hong (1995) grant that limited access to UG is available to L2 learners, but only through their first language. In other words, principles of UG remain accessible to L2 learners to the extent provided to them by their native grammar; however, learners cannot access UG to reset any of their L1 parameters to meet the requirements of the TL.

The view that Clahsen and Hong take is that parameters in L2 acquisition do not apply in the same way as they do in L1 acquisition. They evoke the idea, advanced under the GB framework, that parameters "co-vary". In other words, a parameter does not apply in isolation, but consists of a set of different properties closely related to one another; alternatively, separate parameters may be very closely related and interdependent in the grammar. In L1 acquisition, it is commonly assumed that once a child acquires one property of a particular parameter, she will simultaneously, or soon thereafter, acquire the rest of the properties of this parameter or the interdependent parameter(s). The authors propose, however, that in L2 acquisition this is not the case because full access to UG has been lost and parameters no longer "co-vary".

To test this hypothesis, Clahsen and Hong focus on the null subject parameter in the L2 acquisition of German by adult Korean learners. It should be noted that the L1, Korean, is a null subject language, whereas the L2, German, requires overt subjects. It is commonly assumed in the GB literature that null subjects are a grammatical option in a given language if they can be a) licensed and b) identified (Jaeggli and Safir 1989; Rizzi 1886). The licensing condition refers to a particular property in the grammar that allows null subjects in principle. The identification condition, on the other hand, refers to the availability of a mechanism to recover the missing referent in the empty syntactic position. It is claimed that null subjects can be identified in certain languages via sufficiently rich overt verbal agreement morphology (e.g. Spanish, Italian, Bulgarian, etc.). Languages with poorer agreement (e.g. English, French, etc.) typically do not allow null subjects because the identification requirement cannot be met. However, there are languages lacking overtly expressed verbal agreement altogether, such as Korean and Chinese, which allow null subjects to be identified through preceding topics in the discourse (Huang 1984). Finally, there are also languages, such as German (see below), that do not allow null subjects despite having a rich agreement paradigm.[4]

---

[4] Note that in Minimalism certain assumptions about null subjects have changed. In a recent L2 account of null subjects Park (2004) assumes that their licensing is conditioned by the interpretability of agreement features. Assuming Alexiadou and Anagnostopoulou's (1998) proposal, when a language has [+ interpretable] agreement features, the verb can raise to T and satisfy an EPP requirement. Thus, a null subject would be allowed. However, if a language has [– interpretable] agreement, a subject must be merged in the specifier of TP in order to check the EPP feature (i.e. a subject is obligatory). The discussion in this section follows the GB framework in keeping with the analysis provided by Clahsen and Hong (1995), as well as the subsequent reanalysis by White (2003). Recasting these analyses in minimalist terms

To return to Clahsen and Hong's study, both the L1 and the L2 are assumed to license null subjects. However, while Korean can identify null subjects, using the topic mechanism, German cannot identify null subjects. Thus, if licensing and identification are considered to be two interdependent parameters, Korean and German have the exact same value with regards to licensing, but differ with regards to identification. One might legitimately argue that German has relatively rich morphology and, just like in Spanish and Italian, the null subject can be recovered based on the *phi-features* (person and number) expressed overtly in verbal agreement. To address this, it is claimed that agreement in a language must have the feature [+pronominal] in order to identify null subjects; German, unlike Spanish and Italian, is [–pronominal] which makes it impossible to identify null subjects.

In order to acquire the relevant target grammatical representations in L2 German, the Korean learners can transfer certain properties of their NL, but must restructure others according to the TL settings. On the one hand, they can transfer their implicit knowledge of null subject licensing, as this setting is the same in both the NL and the TL. On the other hand, the learners must discover that German, unlike Korean, is a language with overt agreement, which is a property that in principle allows null subject identification. They must also discover that *Agr* in German is non-pronominal, which is a property that would lead them to the target grammar requirement for subjects to be expressed overtly. In Clahsen and Hong's terminology, the latter two properties described above "co-vary" (i.e. are inter-related). Their claim finds support in L1 studies (Clahsen 1990/91; Clahsen and Penke 1992) where children were found to allow null subjects in German, but only until correct verbal agreement morphology was acquired. Thus, it is argued that acquiring *Agr* made the children realize that its value is [–pronominal], which in turn caused them to stop allowing null subjects. Based on this finding, Clahsen and Hong propose that if parameter re-setting in L2 acquisition mirrors parameter setting in L1 acquisition, L2 learners should also disallow null-subjects as soon as they acquire the German agreement paradigm.

---

is beyond the scope of this literature review whose main purpose is to show cases in which an L2 pattern is not attested in either the NL or the TL.

The authors use a sentence-matching procedure where the participants are asked to identify pairs of sentences as same or different. Since it takes native speakers longer to recognize ungrammatical sentences as same (Freedman and Forster 1985), Clahsen and Hong predict that if learners have correctly acquired the value of the identification parameter in the L2, when they are faced with pairs of identical sentences containing incorrect agreement they should have a longer response time than when they are faced with pairs of sentences containing correct agreement. Furthermore, incorrect sentence pairs with null subjects should take longer that correct pairs with overt subjects. The results showed that almost all of the 20 German native controls behaved as expected: their response latencies were shorter in identifying as identical grammatical sentence pairs with correct agreement as well as grammatical sentence pairs with overt subjects (compared with sentence pairs with incorrect agreement and sentence pairs with null subjects, respectively). The 33 Korean speakers of L2 German responded as follows: a) the responses of 2 participants seemed to suggest that they treated German as a null-subject language without agreement (i.e. just as Korean); b) 13 participants were able to identify grammatical sentence pairs with agreement and grammatical sentence pairs with overt subjects faster than their ungrammatical counterparts with incorrect agreement and null subjects (i.e. a pattern equivalent to the German native controls); c) 5 participants did not differentiate in their response times between sentence pairs with null subjects and sentence pairs with overt subjects, but did have slower responses to sentence pairs with incorrect agreement versus pairs with correct agreement; d) 13 participants showed longer latencies on null-subject sentence pairs versus overt subject sentence pairs, and shorter latencies on sentence pairs with correct agreement versus pairs with incorrect agreement.

The last pattern is of particular interest, as it suggests that the L2 learners treat German as a language which has an overt subject requirement, but at the same time no morphological agreement. Clahsen and Hong propose that this indicates that the two properties do not co-vary in the L2 German interlanguage of the Korean speakers the way they do in child L1 acquisition of German. In other words, the L2 acquisition pattern is different from the one found in L1 children, who acquire agreement and then disallow overt subjects because they have realized that Agr is non-pronominal.

White (2003), however, re-interprets these results and advances a proposal that runs counter to Clahsen and Hong's. She argues that the 13 participants described in (b) above may well have acquired the target responses to agreement and null subjects in L2 German through UG-led parameter resetting. Furthermore, she claims that the 5 participants in (c) above have arrived at the Italian setting of overt agreement and null subjects (i.e. [+pronominal] Agr). Thus, she affirms that these 5 learners have arrived at a setting unattested in the L1 or the L2, but attested in other natural languages, of which the learners have no knowledge. Presumably, these learners have accessed UG in order to achieve this setting. As White (2003) concludes, 20 out of 33 participants showed behaviour consistent with some correct value of the null subject identification parameter: 2 participants showed the L1 setting; 13 participants showed the correct L2 setting; and 5 participants showed a third setting licensed by UG. With regards to the remaining 13 participants who show a pattern not supported by UG (considered by Clahsen and Hong to be evidence of failed parameter clustering, i.e. lack of co-variation), White suggests that abstract agreement can in fact be present in the grammars of these speakers, but it may not always be realized accurately in the surface morphology. Adopting this view, which has become known as the *Missing Surface Inflection Hypothesis* (see also Haznedar 2001; Haznedar and Schwartz 1997; Herschensohn 2001; Ionin and Wexler 2002; Lardiere 1998a, 1998b, 2000; Prévost and White 2000a, 2000b, among others), White (2003) proposes that these learners' responses may not be as problematic as suggested by Clahsen and Hong (1995).[5]

To give a specific example of the MSIH proposal, White (2007) explains that failure to produce verb inflection for past tense, for instance, does not necessarily mean that the L2 learner is lacking the [±past] feature altogether; rather, the abstract tense feature may exist in the learner's grammar but the link between this feature and its PF representation /-ed/ (in English) may not always be established. White further explains that this is due to a breakdown between the syntax and the lexicon, or a problem in processing or computational terms, rather than a permanent representational deficit in the

---

[5] An alternative explanation would be that these learners, similarly to the ones who seem to show the Italian setting, have arrived at a representation licensed by other natural languages. Cases of obligatory subjects and lack of overt agreement have been attested in Creole languages (thanks to Juana Liceras for drawing my attention to this fact).

learners' interlanguage. This breakdown may be seen as a mapping problem (Lardiere, 2000) or as lexical access difficulty (Prévost and White, 2000a, 2000b). Evidence for the MSIH comes from different sources: missing inflection often alternates with accurate inflection; copula and auxiliary verbs are often present and inflected appropriately (in contrast to inflection on lexical verbs), which suggests that abstract features and checking mechanisms are in place; inflectional substitutions are sometimes reported instead of omissions, which suggests presence of abstract features and mechanisms but failure in their overt morphological realization. Additional evidence and discussion of the predictions of the MSIH compared with some competing proposals is provided by White (2007, and references therein). For the purposes of this literature review, it is sufficient to say that the hypothesis can be used to argue against Clahsen and Hong's proposal described above. More importantly, the MSIH will be invoked again in this dissertation in relation to some of the data and the definition of target versus non-target productions in chapter 4.

2.1.5. Discussion

The purpose of this section was to review cases reported in the literature where a particular L2 language behaviour cannot be explained in terms of either the NL or the TL, but is attested in other natural languages. While such cases are relatively few, the issues that they raise are of great interest. The main question to be asked is how do L2 learners arrive at these unexpected language representations? Why would they produce or accept as grammatical a linguistic structure or property that is not attested either in their native grammar or in the target input, and is presumably not part of L2 instruction? One of the possible ways of answering this question is by drawing a parallel between child L1 and adult L2 acquisition. It is well known that children acquiring their native language are able to correctly derive certain grammatical properties despite insufficient target input; children are also known to produce language behaviour unlicensed in their target grammar (i.e. absent from the input), yet attested in other natural languages (e.g. Thornton 1990). Thus, it seems plausible that UG is implicated in both L1 and L2 acquisition and provides a common framework or a pool of language properties, from which the target can be derived. In the process of acquisition, the child or the second

language learner may not always draw the appropriate properties for the TL from this general pool, and thus certain phenomena may be characteristic of other, unrelated languages. Following this reasoning, a view assuming that UG is operative in both L1 and L2 acquisition can tackle such learnability issues.

It should be acknowledged, however, that a UG account of L2 acquisition can be problematic. For example, as pointed out earlier, it is almost never clear what exactly triggers the paradoxical phenomena described above. No property of the target input, the native grammar, explicit instruction, etc., has been identified as a potential cause for the learners' departure from the L1 and obvious misanalysis of the L2 (see White 2003 for further discussion). Another problem with a UG-constrained view is the variation that applies across study participants and results. Not all learners produce and/or accept such unexpected patterns in the L2: some are able to arrive at the target representation immediately, while others seem to retain their L1 settings. Somewhat similar variability is also reported in L1 acquisition but it generally occurs at a lower rate, and the ultimate convergence to the target is essentially guaranteed. Of course, the greater variability in L2 acquisition could be attributed to factors such as the presence of an L1 grammar that can never be 'erased' from the speaker's mind, the greater burden on the memory, the interference of learning strategies and explicit instruction, the imperfect proficiency testing when recruiting participants for a particular study, the higher level of inhibitions which could affect performance, etc. Overall, while a UG-based account of the above phenomena often seems appealing, it will also often face criticisms, some of which remain difficult to address.

An alternative account to some of the above phenomena can be sought in processing. As suggested by Eckman (2004), such an explanation is especially plausible in the case of gaps created by long-distance dependencies. This idea is particularly important as it is applicable to *medial wh-* constructions, which will be introduced later in this chapter. Essentially, a processing explanation would assume that due to a high memory load, especially during the earlier stages of acquisition, L2 learners may overtly pronounce or insert syntactic elements in certain positions, in order to shorten or 'bridge' a long-distance dependency. Thus, accounts along these lines differ from strictly UG-based views in the sense that the phenomena are explained in terms of some specific

processing pressure, as opposed to an available universal option that can be selected by the language learner. Crucially, it is often pointed out in the processing literature that L1 and L2 acquisition seem to involve fundamentally different mechanisms, and studies comparing child L1 and adult L2 participants generally yield distinct results (see Clahsen and Felser 2006 for a recent overview).

The message that this intermediary discussion aims to convey is that L2 phenomena can be accounted for by using different and often competing mechanisms, which makes understanding SLA processes an extremely challenging task. A further discussion of these issues will be provided in the later chapters of the dissertation, and, ultimately, a stronger and more specific stand will be taken after the results of the present research are presented.

2.2. Types of wh- questions crosslinguistically

It is well-known that the mechanisms of wh- question formation vary across languages. The main distinction is drawn between wh- movement and wh- *in situ*; however, there are some other interesting wh- properties that show variation in the languages of the world. Before introducing the phenomenon of *medial wh-*, below is a brief overview of some general facts about wh- questions cross-linguistically.

2.2.1. *Wh- movement*

One of the two basic strategies of wh- question formation is wh- movement. This is common in many languages of the world and is illustrated in English in (12); in this simple (monoclausal) question, the wh- word raises from the site of its base-generation, to a new clause-initial position.

(12) Who$_i$ did John kiss t_i?[6]

In cases where more than one wh- phrase is involved, wh-movement languages fall into two distinct categories. On the one hand, languages like English front only one wh- word and the rest remain in their basic positions (i.e. *in situ*), as in (13). On the other hand, there are Slavic languages (see Rudin 1988) which typically front all wh- words, as illustrated by Bulgarian in (14).

---

[6] Trace notation is used in most examples for convenience.

(13) Who$_i$ t_i gave what to John?
(14) *Koj$_i$ kakvo$_k$ t_i dade t_k na Ivan?*
    who what   gave   to Ivan
    '*Who gave what to John?*'

In complex (biclausal) questions, movement languages raise the wh- phrase from the site of its base-generation in the embedded clause to an initial position in the matrix clause (usually intermediate steps in movement are assumed), as illustrated in (15). The distinction between languages with and without multiple fronting applies in the same way as in monoclausal questions, as shown in English (16) and Bulgarian (17), respectively.

(15) Who$_i$ do you think t_i gave the book to John?
(16) Who$_i$ do you think t_i gave what to John?
(17) *Koj$_i$ kakvo$_k$ misliš   (če) t_i dade t_k na Ivan?*
    who what   think-2sg (that)   gave   to Ivan
    '*Who do you think gave what to Ivan?*'

To summarize, movement languages front one or more wh- element to the left periphery of the matrix clause, regardless of whether the question is simple or complex. Interestingly, some languages that have wh- movement, may also employ other mechanisms of forming wh- questions. This option will be illustrated in the following section and again in the next chapter.

2.2.2. *Wh- in situ*

The second main mechanism for forming wh- questions is to leave the wh- phrase *in situ*. This is the case in Chinese, for example. As illustrated in (18), the object position of a declarative sentence is the same as the position in which the object wh- word of the equivalent interrogative appears in (19).

(18) *hufei mai-le   **yi-ben-shu**.*
    Hufei buy-asp.   one-cl.-book
    'Hufei bought a book.'
(19) *hufei mai-le   **sheme***
    Hufei buy-asp.   what
    '*What did Hufei buy?*'   *(Cheng 1990)*

Similarly, in complex questions, the wh- word also remains in its base position. This is illustrated by Japanese in (20).

(20) *John-wa   [Mary-ga   dare-o   syotaisi-ta to]   omotteiru-no?*
       John-Top   Mary-Nom who-Acc invite-Past Comp think-Q
       'Who does John think Mary should invite?'   (Schulz 2006)

Note that Japanese is an SOV language and the embedded clause containing the wh- object *in situ* is located between the matrix subject (*John*) and the matrix verb (*think*).

Despite the fact that English was already classified as a wh- movement language, it also has a wh- *in situ* option in the so-called echo-questions, as illustrated in (21).

(21) John thinks he should buy what?

Although such structures are more marked and uttered only under specific discourse circumstances, their availability is useful in demonstrating the base-generation sites of wh- words in wh- movement languages.

2.2.3. Medial wh-

Medial wh- constructions, which constitute the primary focus of this dissertation, are a typological alternative in complex wh- question formation in certain languages. Note that I use the term *medial wh-* to refer to both the *wh- copying* and the *wh- scope marking* constructions described below (sections 2.2.3.1 and 2.2.3.2, respectively)[7]. This reflects the basic insight that in both constructions an "extra" wh- word appears in an intermediate position of a complex question (i.e. in the embedded clause), regardless of whether this "extra" wh- word is an exact copy of the one in the matrix clause or not. Before introducing examples of the two types of *medial wh-* and discussing some of their properties, it should be pointed out that whenever attested in a particular language, these constructions usually co-exist with one of the other two mechanisms for forming complex wh- questions described above (i.e. movement or *in situ*). Thus, it seems that *medial wh-* can only be an additional mechanism of complex wh- question formation in a given language (i.e. it is a marked or a secondary strategy within the movement or the *in situ* mechanism). This raises a number of issues with regards to the precise sociolinguistic and semantic status of *medial wh-*, but these remain beyond the scope of the dissertation.

---

[7] Thornton (1990) uses the term *medial wh-* to refer to the *wh- copying* construction only, and uses the term *partial movement* (following McDaniel 1986) to refer to the *wh- scope marking* construction. The terminology in later research also varies to some degree, especially in the acquisition literature.

Typically, *medial wh-* structures are reported in languages that employ wh-movement in question formation (e.g. German, Afrikaans, etc.); however, they have also been attested in *in-situ* languages such as Hindi (*wh- scope marking* only). A further observation with regards to the distribution of the two constructions is that if a given language licenses *wh- copying*, *wh- scope marking* is likely to be available as well (e.g. German). On the other hand, the presence of *wh- scope marking* does not necessarily entail the presence of *wh- copying* (e.g. Hindi). Examples of the two constructions from several different languages are given below.

2.2.3.1.   Wh- copying

As Felser (2004) points out, *wh- copying* constructions are attested in the adult speech of a number of languages, including German (22), Frisian (23), Afrikaans (24), and Romani (25) (for details see Du Plessis, 1977; Hiemstra, 1986; Höhle, 2000; McDaniel, 1986; Reis, 2000, among others).

(22) *Wen glaubst Du, wen sie getroffen hat?*   German
who think   you who she met   has
'Who do you think she has met?'

(23) *Wêr   tinke jo   wêr't   Jan   wennet?*   Frisian
where think you   where that-CL Jan   resides
'Where do you think that John lives?'   (Hiemstra)

(24) *Waarvoor dink   julle waarvoor werk   ons?*   Afrikaans
wherefore think   you   wherefore work   we
'What do you think we are working for?'   (Du Plessis)

(25) *Kas   o Demìri mislenola kas   i Arìfa   dikhla?*   Romani
whom   Demir   think   whom   Arifa   saw
'Who does Demir think Arifa saw?'   (McDaniel, 1989, adapted by Felser)

As these examples illustrate, both wh- arguments and adjuncts can participate in *wh-copying*. While there are various accounts of these constructions in the literature, typically, they have been analysed as an example of successive cyclicity, where the wh-phrase originates in a lower position of the embedded clause and raises through intermediate steps to the left periphery of the matrix clause. The wh- word in the embedded clause represents a copy, which has not been deleted in the process of the derivation, and has been sent to PF spell-out.

The *wh- copying* construction has been studied extensively in German, and has been found to have a number of interesting properties. For example, it is subject to negative islands, as in (26), and is also blocked in the cases of multiple wh- word constructions, as in (27), (examples by Reis 2000 and Fanselow & Mahajan 2000, adapted from Felser 2004).

(26) a. *Wen   glaubst du   nicht, dass sie   liebt?*
      whom believe you   not   that she   loves
      'Who don't you think (that) she loves?'

    b. **Wen   glaubst du   nicht, wen   sie   liebt?*   *(Negative Island)*
       whom believe you   not   whom   she   loves
      'Who don't you think (that) she loves?'

(27) a. *Wen hat Peter wann gesagt, dass er besuchen   wird?*
      who   has Peter when said   that he visit   will
      'Who did Peter say when (that) he is going to visit?'

    b. * *Wen hat Peter wann gesagt, wen er besuchen wird?* *(Multiple Wh-Phrases)*
       who has Peter when said   who he visit   will
      'Who did Peter say when (that) he is going to visit?'

Another interesting property of the *wh- copying* construction is that the *medial wh-* word can co-occur with an overt complementizer introducing the embedded clause. That is, the construction is available in certain varieties of German that accept a doubly-filled complementizer, as in (28).

(28) *Wer glaubst   du,   wer dass du   bist?* *(Doubly-Filled Complementizer)*
    who think   you   who that you   are
    'Who do you think you are?'   (Fanselow and Mahajan 2000)

It is important to note that typically *wh- copying* constructions have a LD counterpart where the intermediate wh- copy is deleted, as illustrated in the (a) variants of the examples in (26)-(27) above. Thus, *wh- copying* is indeed a sub-option of the more standard type of LD movement. Nonetheless, *wh- copying* constructions are widely attested in many varieties of German[8].

While most of the different properties of *wh- copying* will not be discussed any further, the two that are relevant to the L2 experiments of this dissertation are the negation asymmetry illustrated in (26), and the doubly-filled complementizer in (28).

---

[8] As Höhle (2000) notes, *medial wh-* constructions have not been reliably correlated with a particular dialect of German; thus, it seems that they are linked to the colloquial register in many varieties, rather than to a particular dialectal area.

The importance of these properties with respect to the present research will be explained in more detail in the next chapter.

2.2.3.2.   Wh- scope marking

Wh- scope marking[9] seems to be more productive crosslinguistically than *wh-copying*. It has been attested in a number of typologically distinct languages such as German (see collection of articles in Lutz, Müller & von Stechow 2000, among others), Frisian (Hiemstra 1986; McDaniel 1989), Hungarian (Marácz 1990, 2000; Mycock 2006), Russian and Polish (Stepanov 2000), Hindi (Dayal 1994, 1996, 2000), Kikuyu (Sabel 2000b), Malay (Cole and Hermon 2000; Mycock 2006), Warlpiri (Legate 2002), Passamaquoddy (Bruening 2006), etc. In this construction, just as in *wh- copying*, a wh-word appears in an intermediate position of the complex question; this wh- word, however, is not necessarily an exact copy of the one appearing in the matrix. The *wh-scope* construction is illustrated in German in (29)*a*[10]; note that, similarly to *wh- copying*, *wh- scope marking* also has a counterpart with full LD movement, as illustrated in (29)*b*.

(29) a. *Was   glaubst   du   wen   sie   liebt?*
      what   think   you   who   she   loves
      'Who do you think she loves?'
  b. *Wen   glaubst   du   daß   sie   liebt?*
      whom   think   you   that   she   loves
      'Who do you think she loves ?'

Under most accounts *wh- scope marking* is analysed as follows: the contentful wh- word appears in the embedded clause; the wh- element in the matrix clause is an inserted expletive whose purpose is to mark scope. There are analyses, however, where the matrix wh- word is proposed to be argumental, not just an expletive (Dayal 1994, 1996, 2000), and analyses where the wh- element in the matrix is considered to be just a

---

[9] This phenomenon is also sometimes referred to as *partial movement*. I use the term *wh- scope marking*, as it also encompasses *in situ* languages where it would be inaccurate to use the term *movement*.

[10] Note that apart from the German type of *wh- scope marking* where two wh- elements appear overtly (i.e. one in the matrix and one in the embedded clause), there are also languages in which *wh- scope marking* occurs with only one overt wh- element, which is usually located in the embedded clause (Bahasa Indonseia, Malay, Kitharaka, and Quechua). Such constructions are sometimes called *silent scope marking* as the scope marker in the matrix clause (i.e. the equivalent of the matrix *was* in (29)*a* above) is claimed to be covert. For the purposes of this dissertation *silent scope marking* constructions are treated on a par with overt *wh- scope marking* (e.g. German). Relevant examples of *silent scope marking* are provided in (53)-(55), and section 2.3.3 provides further discussion (see also chapter 4).

wh- feature that has separated from the contentful wh- word in the embedded clause and moved to the specifier of the matrix CP (Cheng 2000).

It is important to point out that some of the key properties of *wh- copying* discussed above also apply to *wh- scope marking*. In particular, negation blocks the *wh- scope marking* structure, while the standard LD movement equivalent is grammatical, as illustrated in German by the contrast in (30)*a&b*. Furthermore, as shown in (31), *wh- scope marking* is compatible with an overt complementizer in the embedded clause (in dialects where the doubly-filled complementizer is grammatical; see Müller 1997).

(30) * a. *Was   glaubst   du   nicht mit wem   Maria   gesprochen   hat*?
         what   believe   you   not   with who   Maria   spoken   has
         'Who don't you believe that Maria talked to?'
     b. *Mit wem   glaubst   du   nicht dass   Maria gesprochen hat?*
         with whom   believe   you not   that   Maria spoken   has
         'Who don't you believe that Maria talked to?' (Rizzi 1991, in Beck 1996)

(31) *Was glaubst du   warum (dass) sie das gemacht hat? (Doubly-Filled Comp.)*
      what think you why   that she that done   has
      "Why do you think (that) she has done that?"   (Müller, p.c.)

Considering that *wh- scope marking* and *wh- copying* have certain properties in common, it can be argued that the two are variants of the same construction (see Bayer 1996; Höhle 2000); at the same time, the two constructions have some characteristics that do not coincide, which suggests that they can also be treated as separate (see Dayal 2000; Höhle 2000; Felser 2004, among others). I abstract from this debate, as it does not play a significant role with regards to the L2 acquisition experiments I present in the following chapters of this dissertation. As already mentioned, I use the term *medial wh-* referring to both constructions for convenience, and will draw a distinction between them only when the L2 acquisition data requires it. However, another distinction that has introduced a great deal of controversy in the literature is of crucial importance to this dissertation and will be described in some detail in the following section.

2.2.3.3. Direct versus indirect dependency approaches

The nature of the wh- dependency is one of the key issues in the literature on *medial wh-*.[11] Under the *direct dependency approach*, there is a direct relationship between the wh- phrase in the embedded clause and the one in the matrix CP. Proposals following this general line of reasoning have been made for German by van Riemsdjik (1982), McDaniel (1989), Müller (1997), Cheng, (2000), Sabel (2000a, 2000b); for Hungarian by Marácz (1990) and Brody (1995); and for Hindi by Mahajan (1990). In all these analyses, the wh- element in the matrix clause is a vacuous element that serves the purpose of marking scope and typing the clause as interrogative. Depending on the specific account, this vacuous scope marker may be: a) inserted directly in the specifier of the matrix CP and forming a chain with the contentful wh- phrase at some level, b) replaced by the contentful wh- phrase at LF, or c) represent a feature separated from the wh- phrase and moved up to the matrix Spec CP.

On the other hand, the *indirect dependency approach*, represented by Dayal (1994, 1996, 2000), among others, assumes that the wh- element in the matrix CP is not a vacuous scope marker but a contentful argumental wh- phrase. The analysis is based on data from Hindi, as in (32), but Dayal claims that it applies cross-linguistically (see also Bruening 2006).

(32) *Jaun **kyaa** soctaa hai ki   merii **kis-se**   baat karegii?*
     Juan what think-PR that   Mary who-INS   talk do-FUT
     'Who does John think Mary will talk to?'

Note that Hindi is a wh- *in situ* SOV language and both wh- words in the example above appear in object position in their respective clauses. The wh- phrase *kyaa* 'what' in the first clause quantifies over propositions to which John is in a *think* relationship; the wh-phrase *kisse* 'who' is a question over individuals. Essentially, under the *indirect dependency* approach, there is a contentful wh- phrase in both clauses, which are adjoined to each other and interpreted as two independent questions in their own right. In English, this would be roughly equivalent to two sequential questions, as in (33) or (34).

---

[11] This has been debated mainly with regards to *wh- scope marking*, and the literature referred to in this section focuses on this construction. However, the same debate can in principle apply to both types of *medial wh-* constructions (i.e. *wh- scope marking* and *wh- copying*). This is the case particularly in language acquisition, as will be shown in the following chapter.

(33) What does John think? Who will Mary talk to?

(34) What does John think? What will Mary buy at the store?

The complex debates between the proponents of the *indirect* and the *direct dependency* approaches in the literature have not been resolved and the question still stands as to which approach better captures the various properties of *wh- scope marking* cross-linguistically. It is beyond the scope of this dissertation to relate the debates in any more detail; however, the crucial point that will become relevant in the next chapters is that the *direct dependency* roughly corresponds to long-distance (successive cyclic) movement, illustrated in (35), while the *indirect dependency* roughly corresponds to two independent local movements as in sequential questions, illustrated in (36).

(35) $[_{CP}\, wh\text{-}_i \ldots\, [_{CP}\, wh\text{-}_i \ldots t_i]]$

(36) $[_{CP}\, wh\text{-}_i \ldots\, t_i]\, [_{CP}\, wh\text{-}_j \ldots\, t_j]$

This (simplified) distinction between the two approaches will be used as a key diagnostic of whether *medial wh-* does indeed represent a learnability problem in L2 acquisition (i.e. a phenomenon unrelated to either the native or the target language) in the following chapters. That is, the *direct dependency medial wh-* representation, as in (35), is unattested in both the L1 and the L2 of the two groups of participants (French speaking and Bulgarian speaking learners of English) and as such would constitute a learnability problem if attested in their interlanguage. The *indirect medial wh- dependency*, on the other hand, is equivalent to two sequential questions with local movement of independent wh- phrases and as such poses no learnability problem.

This concludes the discussion of the properties of *medial wh-* constructions in natural languages (i.e. adult L1 grammars). The rest of the chapter focuses on this phenomenon in developing L1 and L2 grammars.

2.3. *Medial wh-* in L1 acquisition

2.3.1. L1 English

Thornton (1990) was the first to report *medial wh-* productions in child L1 acquisition. In an elicited production experiment originally designed to test for *that-trace* phenomena, she found that some English speaking children produced both *wh- copying*

constructions, as in (37), and *wh- scope marking* constructions, as in (38); such utterances were attested in subject, object and adjunct extractions in L1 English. As shown in (39), Thornton also attested *that–trace* effects (i.e. utterances with overt complementizer *that* in subject extractions which are ungrammatical in the adult grammar of English).

(37) ***Who** do you think **who's** under there?
(38) ***What** do you think **who** ate the cheese?
(39) ***Who** do you think **that's** in the box?   (Thornton 1990)

The children Thornton tested were between 2;11 and 5;5 years old (M = 4;3) and some of them produced the ungrammatical *medial wh-* and/or *that*–trace structures consistently, while others only occasionally. The elicitation task was a guessing game in which the children were prompted to talk to a puppet and ask him questions about the location or activity of different toys and props. Interestingly, Thornton found that the ungrammatical structures with *medial wh-*, as in (37)-(38), and the ungrammatical structures with overt complementizer *that* (i.e. *that*–trace effect), as in (39), were in complimentary distribution. In other words, utterances with both a *medial wh-* word and an overt complementizer, as in (40)-(41), were never attested.

(40) a. ***Who** do you think **that who's** under there?   → NEVER ATTESTED
     b. ***Who** do you think **who that's** under there?   → NEVER ATTESTED
(41) a. ***What** do you think **that who** ate the cheese? → NEVER ATTESTED
     b. ***What** do you think **who** that ate the cheese? → NEVER ATTESTED

Since medial wh- and overt complementizers were observed in identical contexts and the two phenomena never co-occurred, Thornton proposed a unified account for them. She claimed that the intermediate wh- word is misanalysed by the children as a complementizer and that it is located in the head of the embedded CP position (just as the complementizer *that*). This proposal is grounded in Rizzi's (1990) idea that in different languages complementizers may agree overtly or covertly with their specifiers. An example of overt agreement (limited only to subject extractions) is the que/qui alternation in French, as (42).

(42) a. *Qui crois-tu   **qu'**est (= **que** est) parti?*
        who believe you   **that** is   left
        'Who do you think left?'
     b. *Qui crois-tu   **qui** est   parti?*
        who believe you   **that** is   left
        'Who do you think left?'

As illustrated above, the complementizer *que* changes to *qui* when the wh- phrase *qui* 'who' is extracted from subject position. Rizzi (1990) proposes that this is an example of specifier-head agreement; that is, the complementizer in C^0 agrees with the trace of the wh- phrase which has moved through the Spec CP position of the embedded clause, as in (43).

(43) [$_{CP}$ qui $_i$ … [$_{CP}$ t_i [$_C$ qui   [$_{IP}$ t_i …]]]].

<div style="text-align:center">spec-head
agreement</div>
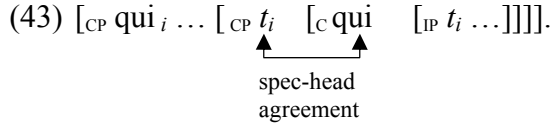
According to Rizzi (1990), an equivalent agreement process applies to English subject extractions, but is not expressed overtly because the complementizer is phonologically null in such contexts, as illustrated in (44).

(44) [$_{CP}$ Who$_i$ do you think [$_{CP}$ t_i [$_C$ ø   [$_{IP}$ t_i kissed John]]]]?

<div style="text-align:center">spec-head
agreement</div>
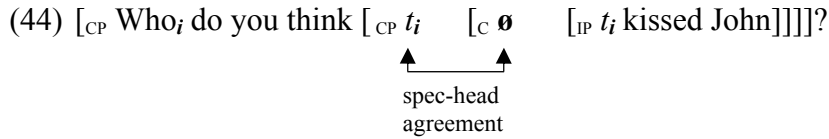
To return to the *medial wh-* utterances observed by Thornton, the proposal is that the children who produced them had not yet realized that English complementizer agreement is covert. Thus, their *medial wh-* words are claimed to be overt complementizers agreeing with the traces of wh- phrases in the specifier of CP.[12] As illustrated in (45), under this analysis a *medial wh-* utterance consists of one 'real' wh- word that has undergone successive cyclic movement and one overt complementizer located in *C;* in the children's grammar, this complementizer undergoes overt agreement with the wh- word in the specifier of the embedded CP and is thus pronounced as homophonous with the wh- element. A similar account applies to the *that*-trace violations that the children produced.

(45) [$_{CP}$ Who$_i$ do you think [$_{CP}$ t_i [$_C$ **who**   [$_{IP}$ t_i kissed John]]]]?

<div style="text-align:center">spec-head
agreement</div>
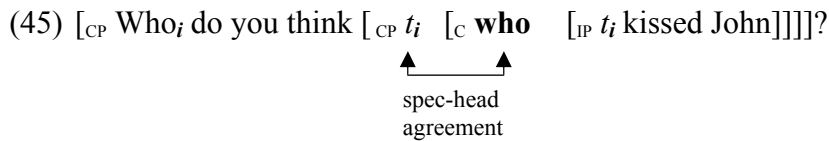
One of the key proposals that Thornton advances is that children's productions may be ungrammatical in the target language, but they are still grammatical options made

---

[12] Note that this analysis applies to the *wh- copying* construction only, which Thornton refers to as *medial wh-*.

available by Universal Grammar. That is, even if children incorrectly postulate extra wh-words as agreeing complementizers, their conjectures about the L1 that they are acquiring are consistent with other natural languages and, more broadly, with the principles of UG. Further evidence for this claim comes from the observation that children never produced *medial wh-* in infinitival clauses, as in (46).

(46) * **Who** do you want **who** to eat the pizza?   → NEVER ATTESTED

Medial wh- in infinitival clauses is unattested cross-linguistically, and the fact that children did not show evidence of such representations demonstrates that the acquisition process obeys UG constraints. As Thornton (1990) points out, the reason for *medial wh-* to be incompatible with infinitival clauses is that specifier-head agreement, as in (45), is by definition restricted to finite clauses (Rizzi 1990) and cannot apply to the example in (46).

Overall, in producing *medial wh-* utterances in finite clauses, children show a grammar that is (temporarily) divergent from the target (English) but consistent with UG; the fact that children did not produce *medial wh-* utterances in infinitival contexts further demonstrates their implicit awareness of abstract UG principles. Hypothesising that L2 acquisition is a process similar to L1 acquisition, one would expect the same results to obtain with a population of L2 English learners. That is, the L2 learners would be expected, similarly to children acquiring their L1, to show evidence of a grammar temporarily divergent from the target but consistent with the principles of UG. *Medial wh-* in the interlanguage of learners whose L1 and L2 do not support such constructions would constitute evidence of such a grammar and of a fundamental similarity between L1 and L2 acquisition. A further prediction with regards to the L2 learners is that the non-target *medial wh-* constructions would disappear at the advanced stages of development, as with L1 learners, and that infinitival clauses would be unattested at all stages of acquisition. These predictions will be tested in the following chapters of the dissertation. In what follows immediately, however, the discussion returns to other cases of *medial wh-* in child L1 acquisition.

2.3.2. L1 Dutch

Following child L1 English, *medial wh-* was attested in the acquisition of Dutch by van Kampen (1997). The author reports on two children acquiring Dutch as a first language who produced complex questions like (47)-(48) both spontaneously and in elicited production.

(47) **Wat** *denk je* **wat** *ik zie?*
 what think you   what I   see
 'What do you think I see?'

(48) **Waarom** *denk je* **waarom** ik op zwemles   *zit*?
 why   think you   why   I   on swimming lessons   am
 'Why do you think I take swimming lessons?'   (van Kampen 1997)

Even though such sentences are ungrammatical in the target language, the two children produced them up to the age of 7 and 8, respectively.[13] In parallel to the English speaking children reported on by Thornton (1990), van Kampen's two Dutch participants produced *medial wh-* only in contexts licensed by Universal Grammar. In other words, they never produced *medial wh-* in infinitival clauses.

Other studies have recently replicated van Kampen's findings of *medial wh-* utterances in Dutch child L1 acquisition. Strik and Jakubowicz (2005) and Jakubowicz and Strik (2008) report results from an elicited production task in which groups of children (ages 3, 4 and 6) produce *medial wh-* constructions, as in (49)-(50).

(49) **Wat** zei   Billy **waarom** Kikker   weggaat?
 what said   Billy why   Frog   leaves?
 "Why did Billy say that Frog leaves?"

(50) **Waar**   zei Lala **waar** de   vis   zwemt?
 where   said Lala where the   fish   swims?
 "Where did Lala say that the fish swims?"   (Strik and Jakubowicz 2005)

The authors reiterate the claim that UG plays an important role in the acquisition of the native language. They further propose that the *Derivational Complexity Hypothesis* (Jakubowicz 2004, 2005) can account for the *medial wh-* data obtained. In essence, the hypothesis predicts that since *wh- copying* and *wh- scope marking* involve a lower degree of derivational complexity, they emerge in child language before the target long-distance questions.

---

[13] Recent research suggests that despite being ungrammatical in standard Dutch, *medial wh-* constructions are attested in certain dialects or registers of this language (Schippers 2006a&b)

One problem of proposals based on derivational complexity is that they predict strong hierarchies, which are rarely fully supported by the data (see Jakubowicz and Strik 2008). A further complication arises from the fact that derivational complexity is difficult to measure in the sense that it is not clear what exactly it constitutes. That is, the number and the nature of computational operations, the distance of the dependency, and the intervening elements between gaps and antecedents are all factors that can be taken as indicative of derivational complexity. However, the way these factors interact and the way they should be ordered when calculating derivational complexity remains unclear. Despite these problems, derivational complexity can provide general insights into why certain structures are preferred to others and why certain structures appear before others in both L1 and L2 acquisition. Such insights will be sought in the discussion of the production data in chapter 4 of this dissertation.

2.3.3. L1 French

Another case of *medial wh-* in L1 acquisition is reported by Oiry (2002) and Oiry and Demirdache (2006). Using an elicited production experiment (see Jakubowicz 2004; Crain & Thornton 1998), these authors attested *medial wh-* in L1 French. The children tested were between 3:02 and 5:11, and an example of a *medial wh-* utterance is given below.

(51) **Tu penses **quoi** que je lis?*
      you think   **what** C° I   read
      'What do you think I'm reading?'   (Oiry, 2002)

Note that (51) differs from the English and Dutch examples above in that the wh- phrase is not located in the left periphery of the matrix clause. Oiry and Demirdache (2006), however, propose that despite this seeming difference, (51) is indeed an instantiation of *medial wh-*. They argue that this is a *scope marking* structure in which a silent non-lexical *Q morpheme* is present in the beginning of the matrix clause (see Mathieu 1999; Cheng and Rooryck 2000). This silent *Q morpheme,* illustrated in (52), marks scope over the matrix clause and is essentially equivalent to the overt *wh- scope* marker *was* 'what' in German, which was discussed in section 2.2.3.2 above, and to overt scope markers attested in other languages.

(52) [**Q_i**]Tu   penses [_{CP}   **quoi**   [que [je lis **t_i**]]]?
       you think   what   C°   I   read   (Oiry, 2002)
       'What do you think I'm reading?'

The proposal of a silent scope marker in these non-target utterances in child L1 French is

supported by cross-linguistic data. As Oiry and Demirdache (2006) point out, silent scope

markers are attested in Bahasa Indonesia (53), dialects of Quechua (54), and Kitharaka

(55).

(53) ∅ *Bill tahu*   **siapa** *yang Tom cintai?*
       Bill knows who   FOC   Tom loves
       'Who does Bill know that Tom loves?'   (Saddy 1991)
(54) ∅ *Jose munan* **may-man**   *Maria away-na-n-ta?*
       Jose wants   where-to   Maria go-NOM-3-ACC
       'Where does Jose want Maria to go?'   (Cole and Hermon 1994)
(55) ∅ *U-ri-thugania*   *ati*   **n-uu**   *John a-ring-ir-e-t?*
       2ndSG-T°-think   that   FOC-who   John SUBJ-beat-T°-FINAL VOWEL
       'Who do you think that John beat?'   (Muriungi 2004)

Thus, Oiry and Demirdache (2006) make the claim that even though the utterances

produced by their child participants are ungrammatical in adult French, they are still

constrained by UG (i.e. *silent scope marking* structures are attested cross-linguistically).

    Further evidence for this claim is offered by Strik and Jakubowicz (2005) and

Jakubowicz and Strik (2008). The authors use the same elicitation task used to collect the

Dutch data mentioned in the previous section, and attest several different types of *medial*

wh- in child L1 French, as illustrated in (56)-(59). Note that because of the use of the

same elicitation task, the *wh- copying* and the *wh- scope marking* constructions in (56)-

(57), respectively, are close counterparts to the Dutch data in (49)-(50). The examples in

(58)-(59), on the other hand, do not have Dutch equivalents, but are similar to the data

reported by Oiry (2002) in (51), and thus the analysis proposed in (52) can be applied to

them.

(56) *Qu'est-ce que Billy a   dit   pourquoi   Grenouille part?*
       what   Billy has   said   why   Frog   leaves
       "Why did Billy say that Frog leaves?"
(57) *Où   Lala a dit   où   le   poisson nage?*
       where   Lala has said where   the   fish   swims
       "Where did Lala say that the fish swims?"
(58) *Billy a   dit   quoi que   Grenouille mange?*
       Billy has   said   what that   Frog   eats
       "What did Billy say that Frog eats ?"

(59) *Lala a   dit   que   c'est qui   qui   boit   de l'eau?*
Lala has said   that   it is   who who drinks   water
"Who did Lala say drinks water?"   (Strik and Jakubowicz 2005)

As with their Dutch data, Strik and Jakubowicz (2005) and Jakubowicz and Strik (2008) propose that the *Derivational Complexity Hypothesis* (DCH) provides an insight into the reasons underlying these utterances. They claim that the non-target structures in (56)-(59) are predicted to pre-date developmentally their target LD counterparts of the adult grammar. The authors further propose that there is a hierarchy of acquisition even within the different kinds of *medial wh-* questions: due to the differences in derivational complexity, *in situ* utterances, such as (58)-(59), are predicted to emerge earlier (i.e. are less complex) than wh- movement structures, such as (56)-(57).

2.3.4. L1 Basque and L1 Spanish

Medial wh- utterances (both *wh- copying* and *wh- scope marking*) were also attested by Gutierrez (2004a, b), who conducted a longitudinal study with a Basque and a Spanish speaking child (aged 5;01 and 4;09, respectively, at the beginning of the study). Examples of *wh- scope marking* and *wh- copying* are given in (60)-(61) for Basque, and in (62)-(63) for Spanish.

(60) *__Zer__ uste duzu **zergatikan** bildurtu   dan   gizona*?
what think aux   why   get scared aux-comp   the man
'Why do you think the man got scared?'

(61) *__Nor__ uste duzu **nor**   bizi dela   etxe   horretan*?
who think aux   who   lives aux-comp house that-in
'Who do you think lives in that house?'

(62) *__Tú__ **qué**   crees **cómo** ha   hecho   el   castillo*?
you what   think how   has   made   the castle
'How do you think he made the castle?'

(63) ***Dónde** crees   **dónde**   ha   ido   el niño*?
where   think-2sg   where   has   gone the child
'Where do you think the child has gone?'

In agreement with the already mentioned literature, Gutierrez makes the claim that since such utterances are ungrammatical in the adult grammars of Spanish and Basque, they show evidence for children's access to UG in L1 acquisition.

Overall, *medial wh-* utterances have become a well-documented phenomenon in child language acquisition. Either *wh- copying* or *wh- scope marking*, or both, have been

attested in the L1 acquisition of English, Dutch, French, Spanish and Basque. Although the data are subject to some variation, in all cases the researchers argue that *medial wh-* is an instantiation of children's access to UG in the process of acquisition. In the next section, I will provide an overview of similar data from L2 acquisition.

2.4.   *Medial wh-* **in L2 acquisition**

The literature on *medial wh-* phenomena in L2 acquisition is much more recent and limited in comparison with the L1 studies described above. All relevant L2 research to date focuses on English as a second language, although the L1 backgrounds of the participants differ.

2.4.1 *L2 English; L1 Japanese.*

Yamane (2003) attested *medial wh-* questions in the L2 English interlanguage of Japanese speakers. She administered an oral translation task to 30 low-level learners of English. The results showed that 9 of the learners consistently produced *medial wh-* constructions and 18 produced such constructions occasionally. Yamane also reported on an acceptability judgment task with 60 learners which supported the findings of the oral translation task. Furthermore, she observed that her subjects accepted *medial wh-* in infinitival clauses at a very low rate and almost never produced such structures; this supports Thornton's original results as well as the cross-linguistic properties of the construction in natural languages (i.e. *medial wh-* does not occur in infinitival contexts). However, Yamane did not find *medial wh-* to block the appearance of an overt complementizer *that* and *vice versa*. Thus, these L2 data do not fit the analysis of *medial wh-* words as agreeing complementizers proposed by Thornton (1990).

Okawara (2000) and Wakabayashi and Okawara (2003) also investigated the presence of *medial wh-* phenomena in the L2 English of Japanese speakers. They used an adaptation of Thornton's elicitation task and tested 16 Japanese university students. Four different kinds of non-target productions were observed: *wh- scope marking* with and without inversion (T-to-C movement) in the embedded clause (64)-(65); lack of a wh-word in the embedded clause, but presence of inversion (66); and lack of a *wh- scope* marker in the matrix clause (67).

(64) **What** do you think **who did** he love?
(65) **What** do you think **who** loved Mr. Yellow?
(66) **Who** do you think **did** Mr. Yellow kiss?
(67) Do you think **what** is in the bag?   (Wakabayashi and Okawara 2003)

These results differ from the data originally reported by Thornton in several respects. On the one hand, no structures with *wh- copying* nor overt complementizer *that* in subject extraction questions were produced by the Japanese participants. On the other hand, questions like (64), (66) and (67) were not attested in Thornton's child L1 English data. Note also that the example in (67) is equivalent to what Oiry and Demirdache analysed as a *silent wh- scope marking* construction in non-adult L1 French (52). Thus, (67) also parallels the data from Bahasa Indonesia (53), Quechua (54), and Kitharaka (55) discussed earlier. The key point in this study is that L2 learners seem to produce a variety of structures not supported by the target grammar, but attested in a number of typologically distinct languages, or else attributable to their L1. No *medial wh-* in infinitival context was reported.

2.4.2. L2 English; L1 Spanish/Basque (bilingual)

Another study that attested *medial wh-* constructions in L2 English was carried out by Gutierrez (2005; see also Gutierrez and García Mayo 2006). The participants in this research were 260 Spanish/Basque bilinguals acquiring English in a school context. The learners were at the beginner and low-intermediate levels. Similarly to Wakabayashi and Okawara (2003), Gutierrez used an adaptation of Thornton's guessing game to elicit long-distance questions. She attested productions of both the *wh- copying* and the *scope marking* constructions in the interlanguage of 32 participants, which is a relatively small subset (12.3%) of the 260 learners. However, the rate of occurrence of *medial wh-* constructions was about 30%, which makes these utterances a robust phenomenon in the L2 English of some Spanish/Basque bilinguals. Gutierrez also found a very high overall occurrence rate of utterances with overt complementizer *that*, similar to the non-adult questions Thornton (1990) reported for English L1 children in (39) above. However, this result is not surprising considering that overt complementizers are obligatory in both Spanish and Basque in subject, object and adjunct extractions. In other words, the

Spanish/Basque bilinguals' errors of this type in L2 English can be straightforwardly explained in terms of transfer from the two L1s.

What Gutierrez's research clearly shows is that *medial wh-* constructions are only one possible way of interlanguage development, and that different learners may take a different route in acquiring target complex questions in L2 English. Even though *medial wh-* constructions were attested only in a small subgroup of learners, however, they could not be discounted as accidental. The challenge of these data is to explain why only some learners produced *medial wh-* while others did not. Another important question that arises is how to account for the fact that *medial wh-* utterances usually co-exist with other constructions in the interlanguage grammar. A more detailed discussion of these issues will be provided in the following chapters.

2.4.3. L2 English; L1 German, L1 Japanese

Finally, Schulz (2006) also reported on *medial wh-* constructions in L2 English. She focused exclusively on *wh- scope marking* and tested two different populations: one with German L1 (59 subjects) and one with Japanese L1 (54 subjects). Schulz administered three different tasks in order to show that *medial wh-* is a robust phenomenon in the interlanguage of her subjects, and not merely an effect produced or exaggerated by a specific experimental technique. The tasks were an oral elicitation procedure adapted from Thornton (1990), a grammaticality judgment test, and an on-line self-paced reading task of the 'stop-making-sense' type (i.e. the participants would see the words of a sentence appear on a screen one after another and press a button at the point where the sentence 'stops making sense' or becomes ungrammatical). Using these three instruments, Schulz was able to attest *wh- scope marking* utterances in both the Japanese and the German populations.

Unlike most other researchers who have reported *medial wh-* in the L1 or L2 acquisition literature, Schulz explored a processing explanation for the *wh- scope marking* constructions in the interlanguage of her participants. More specifically, she hypothesised that *wh- scope marking* may result from processing limitations in the early stages of L2 acquisition and that it can be overcome during the later stages when sufficient processing resources become available. That is, the intermediate wh- word may

serve to shorten the wh- dependency, a strategy that is assumed to pose a lower processing load than standard LD wh- movement involving deletion of all intermediate copies. Schulz predicted that if *medial wh-* is not a valid option in the interlanguage grammar, learners who have insufficient processing resources would find it difficult to reject the *medial wh-* word representation in on-line processing and would experience a delay at the site where the embedded clause is introduced. The results from the on-line self-paced reading experiment, however, showed no evidence of a slow down in the critical region (the *medial wh-* word) of the *wh- scope marking* construction in comparison with the target LD wh- movement. The lack of significant differences in the reading times of the two constructions led Schulz to conclude that acceptance of *wh- scope marking* is not a processing-induced effect.

Looking for an alternative explanation, Schulz proposed that *wh- scope marking* constructions can be attributed to L1 transfer (see also Yamane 2003). This account is not surprising with regards to the German population: since *medial wh-* is licensed in the L1, or at least in many colloquial varieties of the language, it is expected that the construction would be transferred into L2 English. However, the explanation for the speakers with L1 Japanese background is more complex (recall that *medial wh-* is unattested in Japanese). Essentially, Schulz developed an account in terms of abstract transfer of a feature separation operation from Japanese into English. She argued that in the L2 *medial wh-* utterances, a contentful wh- word moves into the embedded CP where its wh- feature separates and moves up to the matrix CP, appearing as a scope marker *what.* This explanation draws on Cheng (2000), who makes such a proposal for *wh- scope marking* in German (see next chapter for details). Schulz proposed that even though *medial wh-* constructions are not part of the Japanese grammar, the feature separation operation is attested in other cases in Japanese, and thus transfer of this operations causes the learners to postulate *medial wh-* in the L2.

It is important to point out that since Schulz used an explanation based on L1 transfer, she advanced the proposal that *medial wh-* as a general English L2 phenomenon may be best explained not by a unified account but as a structure that is constrained or conditioned by language-specific (i.e. L1-specific) factors. One of the implications of this proposal is that *medial wh-* constructions may not be attested in the English interlanguage

grammars of speakers with other typologically different L1 backgrounds. Due to the limited data available on *medial wh-* in L2 acquisition, it is difficult to say whether this prediction can receive support at this point. The data that will be presented in the following chapters, however, adds two new typological backgrounds (L1 French and L1 Bulgarian) to the L2 English *medial wh-* literature. As such, this dissertation will contribute to the body of research necessary to evaluate Schulz's proposal.

2.5. Conclusion

This chapter began by reviewing the literature on different L2 phenomena that cannot be accounted for in terms of either the properties of the NL or the TL, but are attested in other natural languages. Such phenomena pose a learnability problem in L2 acquisition: How can a speaker arrive at a language representation that is not evidenced by the target input, yet licensed by the abstract grammatical principles that apply to the languages of the world? While the existence of such phenomena has typically been interpreted as evidence for UG in L2 acquisition, some of the research proposals described in section 2.1 argued for processing explanations ranging from being consistent with a UG account to strongly restricting learners' recourse to UG principles. Assuming a UG account, one of the challenges is to provide evidence that learners' hypotheses about the L2 never go beyond the UG hypothesis space. In other words, patterns not licensed by UG (i.e. not attested cross-linguistically) are generally not expected to occur in the interlanguage of L2 learners. While acquiring such evidence constitutes a more ambitious and longer research program than this dissertation can undertake, the data in chapter 4 will offer some relevant facts in this respect. Assuming a processing account, on the other hand, raises, among other things, the question whether and how processing interacts with the underlying grammatical competence of the learners. While the data which will be presented in the following chapter are best explained in terms of the learners' implicit grammatical competence, the results in chapter 4 are best explained in terms of processing effects. As such, in this dissertation I will argue that grammatical and processing phenomena work hand in hand and that L2 patterns (target and non-target) may be attributed to both of these factors. This idea is consistent with Phillips' (1996)

claim that processing and grammar are inseparable (for more details on this claim see chapters 3, 4 and 5).

Another important issue that will be raised in the following chapters is individual variability. The question why particular non-target L2 patterns may be attested with certain speakers but not with others constitutes a central theme in L2 research. Assuming a UG-constrained L2 acquisition model, there is no specific reason requiring learners to undergo the same route in acquiring the target grammar. That is, individual learners may use a variety of mechanisms and go through entirely different interlanguage representations, as long as they are consistent with UG and convergence with the target is eventually achieved. Under a processing-based proposal, on the other hand, individual differences can be explained by the different memory and other cognitive capacities of the learners. These ideas will be further elaborated on and evaluated in the following chapters of the dissertation.

Returning to the current chapter, apart from providing an overview of L2 phenomena that cannot be explained on the basis of the properties of either the NL or the TL, it also offered a summary of the typology of wh- question formation mechanisms across languages; in addition, the chapter provided examples and discussion of the properties of *wh- copying* and *wh- scope marking* constructions in languages where such representations are valid options in the adult grammar. The negation asymmetry and the possibility of co-occurrence of a *medial wh-* word and an overt complementizer were identified as two crucial properties that will become relevant in the next chapter.

The last part of this chapter provided a review of the L1 and L2 acquisition literature showing that *medial wh-* is a well-documented phenomenon in the former, and an emerging topic in the latter. A variety of proposals involving access to UG, processing, and abstract L1 transfer were mentioned with regards to the existence of *medial wh-* in L2 English. Apart from explaining *medial wh-* structures in the L2, however, all these proposals raise the issue of whether and how such non-target structures disappear from the interlanguage at the later stages of acquisition. The role of L2 input is particularly interesting in this respect: in essence, there does not seem to be a particular input trigger indicating to the learners that such constructions are not licensed in the L2, and yet the data seem to show that advanced participants with enough L2 experience are able to

"unlearn" *medial wh-* structures. This indicates that indirect negative input (i.e. lack of *medial wh-* in the target input) has an effect on the L2 acquisition process. A more precise formulation of this idea in terms of a Bayesian statistical framework (see Tenenbaum and Griffiths 2001) will be offered in the next chapter of the dissertation.

Finally, before proceeding to the written L2 experiments in the next chapter, it is important to point out that *medial wh-* representations have been attested in both L1 and L2 acquisition mainly under experimental conditions (i.e. oral elicitation tasks, grammaticality judgements, processing measures, etc.). While it could be argued that *medial wh-* also needs to be attested in spontaneous production in order to be considered an important property of developing (L1 or L2) grammars, complex wh- questions in general are very rare in spontaneous circumstances. Thus, the studies described in the next two chapters involve controlled experiments in line with most of the previous research.

Chapter III
Written Experiments

- ▪ SUMMARY: THIS CHAPTER RERPORTS ON THE RESEARCH QUESTIONS, DESIGN, AND RESULTS OF A WRITTEN STUDY ON *MEDIAL WH-* CONSTRUCTIONS CARRIED OUT WITH TWO DIFFERENT POPULATIONS: FRENCH SPEAKERS OF L2 ENGLISH, AND BULGARIAN SPEAKERS OF L2 ENGLISH. THE CHAPTER INCLUDES A DISCUSSION OF THE FINDINGS AND OFFERS SOME CONCLUSIONS; THESE ARE FURTHER EXAMINED AND MODIFIED IN THE FOLLOWING CHAPTER, IN WHICH I REPORT ON THE ORAL EXPERIMENTS.

3.1 Introduction and research questions

As indicated in the previous chapter, *medial wh-* is extensively studied as a phenomenon licensed in the adult grammar of certain languages, such as German, Hindi, Hungarian, Romani, etc. It is also well documented in the child L1 acquisition of some languages in which it is not a valid option in the adult grammar (e.g. English, Dutch, French, Spanish, and Basque). However, while the phenomenon undoubtedly raises many interesting and potentially challenging questions in second language research as well, only a handful of L2 studies have been dedicated to it. So far, *wh- scope marking* or *wh- copying* (or both) have been attested in the L2 English of Japanese, German and Spanish/Basque L1 speakers. These results suggest that *medial wh-* is a phenomenon that occurs in the interlanguage of English L2 learners with several distinct typological backgrounds. At the same time, in order to confirm that *medial wh-* is indeed a robust phenomenon in L2 English, and perhaps in the L2 acquisition of other languages, further studies attesting its existence are needed. As such, one of the main contributions of this dissertation is adding two new typologically distinct L1 backgrounds to the L2 English *medial wh-* literature: L1 (Canadian) French and L1 Bulgarian.

One interesting issue raised by Schulz (2006) with regards to *medial wh-* in L2 acquisition is whether it should receive a unified explanation or whether it is a phenomenon determined by the particular L1 background of the L2 speakers. Another interesting question is whether *medial wh-* in L2 English is best explained in terms of recourse to UG, in terms of processing constraints, in terms of domain-general learning mechanisms, or in terms of a combination of some of these factors. These are among the general acquisition issues that this dissertation ultimately aims to address. The more specific research questions of the L2 experiments on which I report in this chapter are listed below.

The main question this dissertation addresses is whether *medial wh-* constructions (*wh- scope marking* and/or *wh- copying*), exist in the interlanguage of either (or both) of the two learner populations. As already mentioned, the L2 literature on *medial wh-*, although limited, indicates that such constructions may not be restricted to a particular L1 background. Thus, the prediction is that they should also be attested in the English interlanguage of French and Bulgarian speakers. However, since *medial wh-* in both L1 and L2 acquisition studies does not seem to be a phenomenon that can be attested with all participants, and usually occurs at steady but relatively low rates (e.g. between 10% and 30%), it was expected that similar distribution and occurrence rates would be attested with the French and Bulgarian learners of L2 English.

The potential implications of the above research question can be far-reaching. On the one hand, if *medial wh-* is attested in the L2 English of French and Bulgarian speakers, then these data can be added to the pool of Spanish/Basque and Japanese speakers who have already been found to produce and/or accept such constructions in L2 English.[1] Enlarging this pool of L2 English learners with typologically different L1 backgrounds would be a strong indication that *medial wh-* is a universal phenomenon in (English) L2 acquisition, and as such deserves a unified explanation. On the other hand, if *medial wh-* cannot be attested in the English interlanguage of French and/or Bulgarian speakers, then such data would indicate that the phenomenon is largely determined by the specific L1 background of the learners, and thus a language-specific (i.e. L1-specific) account would be more appropriate.

To recapitulate, the main research question of this dissertation is formulated as follows.

R1. Do *medial wh-* representations exist in the interlanguage of French and/or Bulgarian learners of L2 English?

As already mentioned, it was expected that *medial wh-* would be attested in the two speaker populations, and thus the rest of the research questions were based on this prediction.

---

[1] Recall that *medial wh-* was also attested in the L2 English of German speakers (Schulz 2006). However, this group is not mentioned here because German, unlike Japanese, Basque and Spanish, does have *medial wh-* as a licensed grammatical option in colloquial speech. As such, *medial wh-* in a German-English interlanguage receives a straightforward explanation in terms of L1 transfer.

The second research question concerns the distribution of the particular type of *medial wh-* (i.e. *copying* or *scope marking*) in the English interlanguage of the two populations.

R2. If *medial wh-* is attested, is it of the *wh- copying* or the *wh- scope marking* kind?

As discussed in chapter 2, for the purposes of L2 acquisition, the *wh- copying* and *wh- scope marking* constructions are treated on a par in this dissertation because they share some common properties, and because they are both ungrammatical in the target and the native languages of the participants. Nonetheless, it is important to determine if one of the two constructions occurs more often than the other. While the literature lacks specific studies on the crosslinguistic distribution of *wh- copying* and *wh- scope marking*, it seems that the latter is more likely to occur as a licensed option in the adult grammar of a given language than the former (see Lutz et al. 2000). Furthermore, the rates with which *wh- scope marking* has been attested in previous L2 studies are higher than the rates of occurrence of its *copying* counterpart. Thus, the prediction would be that the *wh- copying* construction would be attested to a lower extent than the *wh- scope marking* one.

The third research question is with regards to the nature of the wh- dependency: *direct* or *indirect*.

R3. If *medial wh-* is attested, does it constitute a *direct* or an *indirect dependency*?

Recall from the previous chapter that the analyses of *medial wh-* in the syntax literature are split between a *direct* and an *indirect dependency* approach. I use these terms (slightly modified) to correspond to a long-distance successive cyclical wh- dependency versus two independent local wh- movements (i.e. sequential questions).[2] The schematic representation of the *direct* and the *indirect dependency* given in chapter 2 is repeated here in (1) and (2), respectively.

(1) Direct dependency
     $[_{CP} wh\text{-}_i \ldots [_{CP} wh\text{-}_i \ldots t_i]]$

(2) Indirect dependency
     $[_{CP} wh\text{-}_i \ldots t_i] [_{CP} wh\text{-}_j \ldots t_j]$

---

[2] As mentioned in chapter 2, these terms are used in the syntax literature mainly with regards to *wh- scope marking* and do not exactly correspond to LD wh- movement versus sequential questions.

An important criticism that has been voiced with regards to some of the previous research of *medial wh-* in L2 is that these constructions can be analysed as either (1) or (2), and most of the studies do not offer a principled way of distinguishing between the two (see Schulz 2006 for an overview). As such, the status of a *medial wh-* L2 utterance, as in (3), is ambiguous. At first sight, it looks like a *direct dependency* where the wh- word has originated in the embedded clause and moved to initial position in the matrix, leaving an undeleted copy along the way. At the same time, however, it could also be a sequence of two semantically but not syntactically related questions with independent wh- movements as in (4). Of course, the second question in this sequence is ungrammatical in English, as it lacks *do* support and T-to-C movement (i.e. subject-auxiliary inversion). Its grammatical equivalent is given in (5). However, it is well known that L2 learners are often inaccurate with do support and T-to-C movement, and thus it is very plausible that many research participants could produce and/or accept an utterance as in (3), assigning to it an *indirect* representation as exemplified (4) and schematized in (2).

(3) What do you think *what* they like?
(4) What do you think? What they like?
(5) What do you think? What *do* they like?

The nature of the dependency has serious implications in investigating *medial wh-* utterances in the L2 English of French and Bulgarian speakers. On the one hand, the *direct dependency* configuration is unavailable in both the NLs and the TL. Thus, if the attested *medial wh-* utterances are assigned a *direct dependency* analysis, they would constitute a learnability problem: What makes learners resort to a representation unattested in the native grammar and in the target input, but licensed in a number of typologically distinct languages, of which they report no knowledge? On the other hand, if the *medial wh-* utterances can be accounted for in terms of an *indirect dependency* (i.e. sequential questions), no learnability issues arise; that is, sequential questions are attested in both the L1s and the L2 of the participants. Nonetheless, such utterances would still be of interest, as they could indicate a stage of L2 development in which the speakers avoid long-distance movement of wh- phrases and favour local short movement. The reasons for such utterances can be sought in a processing strategy arising from the complexity of LD wh- movement, for example.

While both the *direct* and the *indirect dependency* representations bring up very interesting acquisition issues, the question that needs to be addressed is how to distinguish between them. The appropriate measure in this respect is to control for the participants' knowledge of inversion. That is, if a particular participant has acquired T-to-C movement in simple (monoclausal) questions, then the *indirect dependency* representation in this speaker's interlanguage would be as in (5) above (i.e. featuring two separate T-to-C movements). On the other hand, if the same speaker, who has acquired inversion, produces or accepts an utterance as in (3) as grammatical, then this would constitute evidence of the *direct dependency* representation. In other words, if the speaker has acquired inversion, inversion in both clauses would be equivalent to sequential questions; inversion in the matrix clause and lack of inversion in the second clause, on the other hand, would point to one single question with LD movement.[3] As already illustrated in (5), T-to-C movement in English may involve *do support*. In such cases, in order to invert the subject and the auxiliary, the participant would first need to insert *do* in the representation. Knowledge of *do* insertion as such, however, is not relevant to the issue of the nature of the dependency, and no further reference to it will be made. For the purposes of the experiments described below, controlling for T-to-C movement means controlling for inversion of the subject and the auxiliary, regardless of whether the latter is *be, do* or a modal.

To recapitulate, in order to be able to distinguish between the *indirect* and the *direct dependency*, a study must control for the participants' knowledge of inversion, and subsequently exclude the participants who have not acquired this property from the analysis. The current study controls for the participants' knowledge of T-to-C movement, and as such offers a principled way of approaching the question of the nature of the *medial wh-* dependency. More details on the specific procedure that I adopt will be given in the following sections of this chapter.

The fourth research question in this study is based on Thornton's (1990) results from child L1 English. Recall Thornton's observation (chapter 2, section 2.3.1) that

---

[3] Note that the particular formulation of the terms *direct* and *indirect dependency* adopted in this dissertation allows for T-to-C movement to be used as a diagnostic for the nature of the dependency in L2 English. This does not necessarily apply to natural languages where *wh- scope marking* is attested (e.g. German, Hungarian, etc.) as the *direct* versus *indirect dependency* distinction is not defined in terms of T-to-C movement (see Lutz et al. 2000, among others, for details).

medial wh- words occurred in complementary distribution with the overt complementizer *that*. Thus, she proposed that children mistakenly categorized *medial wh-* words as complementizers. If this account of L1 child English is correct, it can also be evoked in L2 acquisition. That is, L2 learners might also be misanalysing *medial wh-* words and using them interchangeably with complementizers. If this were true, the prediction would be that *medial wh-* and overt complementizer *that* would not co-occur, just as in Thornton's L1 data; thus utterances as in (6) would not be attested. Conversely, if *medial wh-* has a status different from that of a complementizer, there should be no restriction on the co-occurrence of *medial wh-* and overt complementizer *that;* thus utterances as in (6) should be possible.[4]

(6)   Who do you think **that who** John kissed?

Based on the above reasoning, the formulation of the fourth research question is as follows.

R4. If *medial wh-* is attested, can it co-occur with an overtly expressed complementizer *that*?

The last research question pertains to a negation-induced asymmetry mentioned in the previous chapter (sections 2.2.3.1 and 2.2.3.2). Recall that one of the widely-studied properties of *medial wh-* in German is an intervention effect whereby negation in the matrix clause renders *medial wh-* ungrammatical, as in (7), while the long-distance equivalent remains unaffected as in (8).[5]

(7) **Wen   glaubst du   nicht, wen   sie   liebt?*
      whom believe you   not   whom   she   loves
     'Who don't you think (that) she loves?'
(8) *Wen   glaubst du   nicht,   dass sie   liebt?*
      whom   believe you   not   that she   loves
     'Who don't you think (that) she loves?'

                              (repeated from chapter 2, section 2.2.3.1)

Based on this observation, the current study tests whether negation might have a similar blocking effect on *medial wh-* in the English interlanguage of (Canadian) French and

---

[4] It could be argued against this idea if we assume that learners who accept or produce an utterance as in (6) assume a doubly-filled complementizer. This option will be investigated further in the discussion section of this chapter.
[5] The examples here are of the *wh- copying* construction, but the same negation asymmetry applies in *wh-scope marking* as well (see chapter 2, section 2.2.3.2).

Bulgarian speakers.

   R5. If *medial wh-* is attested, is it subject to a negation asymmetry?

Note that negation asymmetries in wh- questions are found in various languages. For example, French which has both a movement and an *in situ* option in wh- questions disallows negation in the former case, as in (9), but not in the latter, as in (10).

 Wh- in situ:
  (9) **Il ne mange pas quoi?*
    He NEG eat   NEG what
    'What doesn't he eat?'
 *Overt wh-*movement:
  (10)*Qu'est-ce qu'il ne mange pas?*
    What is-it that-he NEG eat NEG
    'What doesn't he eat?'   (adapted from Oiry and Demirdache 2006)

With regards to the negation intervention effect in (9), Mathieu (1999) argues the following: the wh- phrase consists of an operator and a variable; the operator is assumed to be phonologically null and to move to the specifier of CP; in this position it marks scope, provides a binder (antecedent) for the wh- variable and checks the strong (assumed in the earlier version of the MP) feature of C. Mathieu further proposes that the variable left *in situ* is non-referential and thus requires a local antecedent. Since antecedent government is strictly local, island effects (i.e. negation intervention) are expected. Relating the example in (9) to second language acquisition, it is important to recognize the possibility of negation effects arising in the L2 English of the French speaking participants due to both transfer from the L1 and certain cross-linguistic properties of the *medial wh-* construction itself. Such transfer effects cannot be predicted for the Bulgarian group as the L1 does not have a negation asymmetry; however, if negation does intervene with *medial wh-* in the L2 English of the Bulgarian speakers, then it could be argued that they are sensitive to some universal properties of *medial wh-* and negation.

  Overall, the main goal of the five research questions described above is to establish if *medial wh-* is a phenomenon that occurs in the English interlanguage of the two different L1 populations, and if so, to describe its properties and propose an appropriate analysis for it. I will return to the research questions later in this chapter when I lay out the specifics of the experimental design.

3.2   Experiment 1: L1 French; L2 English.

Before proceeding with the study of the L1 French population, it is important to review the different mechanisms of forming long-distance wh- questions in French. French has been described as a 'mixed' language in this respect, as it has both the movement and the *in-situ* options. This complexity is further augmented by the existence of different strategies within these options. In describing the basic typology of complex wh- questions in French, I follow Jakubowicz and Strik (2008). The movement option illustrated in (11)*a&b* with a subject and an adjunct wh- extraction, respectively, is considered the most "standard" or normative way of forming a long-distance question. Note that apart from wh- movement, these variants also have subject-verb inversion (V- to-C movement). However, the same questions are also possible without subject-verb inversion, as illustrated in (12)*a&b*.

(11) a. ***Qui*** *penses-tu qui lit des histoires*?
      who think you that reads stories
      'Who do you think is reading stories?'
   b. ***Où*** *dis-tu que Marie va?*
      where say you that Marie goes
      'Where do you say that Marie is going?'

(12) a. ***Qui*** *tu penses qui lit des histoires*?
      who you think that reads stories
      'Who do you think is reading stories?'
   b. ***Où*** *tu dis que Marie va?*
      where you say that Marie goes
      'Where do you say that Marie is going?'

                               (adapted from Jakubowicz and Strik, 2008)

Fronted LD wh- questions are also possible with the wh- + *ESK* construction, as shown in (13)*a&b* with an argument and an adjunct extraction, respectively.

(13) a. ***Qui*** *est-ce que tu penses qui lit des histoires?*
      who is it that you think that reads stories
      'Who do you think is reading stories?'
   b. ***Où*** *est-ce que tu dis que Marie va?*
      where is it that you say that Marie goes
      'Where do you say that Marie is going?'

                               (adapted from Jakubowicz and Strik, 2008)

As mentioned above, complex questions in French can also be formed by leaving the wh- phrase *in situ* as in (14).

(14) a. *Tu   penses   que **qui**   lit   des histoires?*
    you   think   that who   reads stories
    'Who do you think is reading stories?'

   b. *Tu   dis   que Marie   va   **où**?*
    you   say   that Marie   goes   where?
    'Where do you say that Marie is going?'

(adapted from Jakubowicz and Strik, 2008)

As a fourth strategy, complex questions can also be formed using cleft structures as in (15).

(15) a. *Tu   dis   que c'est   **où**   que Marie va?*
    you   say   that it is   where   that Marie goes?
    'Where is it that you say Mary is going?'

   b. *Tu   penses   que c'est   **qui**   qui   lit   des histoires?*
    you   think   that it is   who   that   reads stories
    'Who is it that you think is reading stories?'

(adapted from Jakubowicz and Strik, 2008)

In essence, this review of the various strategies of forming complex questions shows that French can employ a wh- long-distance movement mechanism, which is similar to English, and an in-situ mechanism, which is unavailable in English.[6] In addition, it is important to point out that in English complementizers are optional in object and adjunct extractions, as illustrated in (16) and (17), respectively, and illicit in subject extractions, as in (18). In French, on the other hand, complementizers are obligatory in all contexts. Thus, if the complementizers were deleted in any of the examples in (11)-(15) above, the sentences would be ungrammatical.

(16) What do you think (that) John bought?
(17) Where do you think (that) John went?
(18) Who do you think (*that) purchased the book?

Finally, note that the French declarative complementizer *que/qui* 'that' is homophonous with the wh- words *que* 'what' and *qui* 'who', as shown in examples (19) and (20), respectively.

---

[6] Except for the *echo* questions shown in section 2.2.2 of the previous chapter.

(19) *Que₁* penses-tu   *que₂* j'ai   acheté?   *que₁ ≠ que₂*
    What think-you   that   I have   bought?
    'What do you think I bought?'

(20) *Qui₁* penses-tu   *qui₂* est   arrivé?   *qui₁ ≠ qui₂*
    who think you   that   has   arrived
    'Who do you think arrived?'

Depending on the particular wh- word used and on agreement processes applying to the complementizer, homophony between the two does not always apply, as illustrated earlier in (11) (12)*b*, (13)*b*, (14)*a*, (15)*a*. However, when homophony as in (19) and (20) arises, its possible transfer effects with regards to the English interlanguage of the French participants should be taken into consideration (see discussion section of this chapter).

3.2.1. Participants

The total number of experimental participants in this study was 161 (prior to exclusion procedures). All of them were enrolled in a five-week immersion ESL program at the Centre Linguistique du Collège de Jonquière in Ottawa, Ontario. The participants received about six hours of ESL instruction per day, and were also engaged in additional social activities (sports, movies, dances, etc.) where the medium of communication was English. They were all placed with English-speaking home-stay families for the duration of the program. The program had an "English only" policy whereby the participants were asked to sign a formal pledge to communicate in English amongst themselves for the duration of their stay in Ottawa, and were subject to various discipline measures (warnings, detentions and ultimately expulsion) if they were heard using their L1 at any time.

The majority of the participants were high-school students from different towns across the French-speaking province of Quebec; however, there were also some students who came from Spain. All participants were between 15 and 18 years old and had had between 5 and 12 years of exposure to English (mainly in school setting) before enrolling in the program. For the purposes of this study all Spanish speakers were excluded, as were a few other participants who reported that their L1 was not French (see section 3.2.2.2 below for description of the language background questionnaire administered). The remaining experimental participants on which the study focuses is 130, as indicated

in table 1.

Table 1. Number of participants by L1 (experiment 1)

	Number	Excluded	Retained
L1 French	130	0	130
L1 Spanish	29	29	0
Other L1	2	2	0
Total Experimental	161	31	**130**

Apart from the experimental participants, the study also tested a control group of 21 native speakers of North American English. Since *medial wh-* utterances are considered ungrammatical in all varieties and registers of English[7], no specific criteria with regards to variety, age, etc., were adopted when recruiting the native speakers.

3.2.2. Design and procedures

3.2.2.1. Placement test

All experimental participants were given a placement test by the ESL program in order to determine their proficiency level at the beginning of classes. The placement test involved reading and listening comprehension, vocabulary questions, and other typical ESL activities. Due to the intensive nature of the program, it was not possible to administer a separate proficiency test specifically designed for the purposes of this study, and thus the proficiency levels used are based on the program's levels. Note, however, that the level of performance on the experimental task increases with the level of proficiency determined by the ESL program (see results sections below). This serves as an independent measure of the validity of the placement test.

The six proficiency levels of the L1 French participants are illustrated in table 2.

Table 2. L2 proficiency levels (experiment 1)

Group	Level	No. of participants
F1[8]	High Beginner	28
F2	Low Intermediate	25
F3	Intermediate	22
F4	High Intermediate	24
F5	Low Advanced	18
F6	Advanced	13
	Total:	**130**

---

[7] To date, *medial wh-* structures have never been reported for English.
[8] F1=French speakers, group 1, F2=French speakers, group 2, etc.

Neither the placement test nor any of the teaching materials or instruction during the ESL program involved English complex questions (i.e. long-distance wh- movement). Thus, the grammatical equivalent of the structures in which this dissertation is interested was not addressed through explicit instruction in the ESL program. Of course, this cannot be guaranteed for ESL classes prior to enrolment in the program. However, even if the participants had previously received classroom instruction on complex questions, this would not pose a problem for the experiment; on the contrary, if the participants had been taught how to form grammatical long-distance wh- questions in English and *medial wh-* was still to be attested in their interlanguage, then the import of such results would be even higher.

3.2.2.2.   *Language background questionnaire*

In order to determine the extent and context of exposure to the L2 prior to participation in the study, a language background questionnaire was administered following the placement test. The questionnaire was originally designed by the Language Acquisition Laboratory at the University of Ottawa and included questions about the participants' native language, knowledge of other languages, years of ESL instruction, frequency of L2 use, previous L2 immersions, etc. (see appendix 1). The results of the questionnaire were used mostly in deciding which participants to exclude from the study (i.e. due a native language other than French). The native control speakers were not given a questionnaire, but were asked to confirm verbally that English was their first language.

3.2.2.3.   *Experimental task*

The experimental task was a multiple-choice grammaticality judgement test (sometimes referred to as "multiple-choice fill in the blanks task" in ESL instruction).[9] The test contained 40 items in total: 24 experimental items, 8 control items, and 8 fillers.

---

[9] Opting for a multiple-choice test instead of a standard grammaticality judgement task was based on the reasoning that in the former the participants' judgements are focused on a specific part of a sentence (i.e. the part testing for *[+/–] medial wh-* and/or *[+/–] inversion*, as illustrated below). In a standard grammaticality judgement task, where the participants judge the whole sentence, it is possible that some of the L2 learners' answers are based on some other aspect, unrelated to the specific phenomenon under investigation (see Schulz 2006 for discussion of such concerns with a standard grammaticality judgment task in an L2 context).

The experimental items were divided into 3 conditions (8 items each). The number and types of items are summarized in Table 3.[10]

Table 3. Experimental task: types and number of items

Type of Item	No. of items
Condition 1: *Basic* (complex questions)	8 (4 wh- copy + 4 wh- scope marking)
Condition 2: *Complementizer* (complex quest.)	8 (4 wh- copy + 4 wh- scope marking)
Condition 3: *Negation* (complex questions)	8 (4 wh- copy + 4 wh- scope marking)
Control (simple questions)	8
Filler (declaratives)	8
Total:	**40**

All 40 items were introduced by a brief story of a few sentences which served to provide context and to prevent the task from becoming mechanical. The stories used simple vocabulary and sentence structure so as not to pose a comprehension challenge to less advanced learners. Each story was accompanied by a picture to provide further context and to facilitate comprehension. The multiple-choice items contained four choices: *A*, *B*, *C* and *D*. The participants were instructed to mark one choice as correct for each item. The four choices in the three experimental conditions represented four different types of embedded clause. *Choice A* was the target utterance (i.e. the grammatical LD option in English) with no *medial wh-* and no subject auxiliary inversion (T-to-C movement) in the embedded clause; choice *B* contained inversion in the embedded clause[11] but no *medial wh-*; *choice C* contained *medial wh-* in the embedded clause but no inversion; and *choice D* contained both *medial wh-* and inversion in the embedded clause. All items as well as the four choices contained within them were randomized. The four choices are summarized in Table 4.

Table 4. Experimental items: types of embedded clause

Choice	Embedded clause	Grammaticality in Standard English
A	[–] medial wh-, [–] embedded inversion	grammatical
B	[–] medial wh-, [+] embedded inversion	ungrammatical
C	[+] medial wh-, [–] embedded inversion	ungrammatical
D	[+] medial wh-, [+] embedded inversion	ungrammatical

---

[10] Note that some groups were given a shorter version of the test (see end of this section).

[11] Note that subject auxiliary inversion in the embedded clause is grammatical in some varieties of English (e.g. Irish English, etc.). See below for discussion of this issue and its possible impact on the results of the study. See also McCloskey (2006) for further discussion and syntactic analysis of such utterances in Irish English.

As mentioned above, the test contained three experimental conditions. The type of lead-in story and prompt, as well as the type of item (i.e. multiple-choice) were similar in all three conditions. However, the type of structure included in the multiple-choice items was different. The structures in *condition 1* contained no complementizer and no matrix negation, and thus this condition was called *basic,* (i.e. least marked). The structures in *Condition 2* contained an overt complementizer but no matrix negation; this condition was called *complementizer*. The third condition contained matrix negation but no complementizer; thus, this condition was called *negation*. Having these three different conditions allowed for the explicit testing of the research questions regarding overt complementizer and negation mentioned above (more details on how the experimental task addresses each of the research questions are provided later in this section). The differences between the three conditions are summarized in table 5.

Table 5. Experimental conditions

	Matrix negation	Complementizer
Condition 1: *Basic*	[−]	[−]
Condition 2: *Complementizer*	[−]	[+]
Condition 3: *Negation*	[+]	[−]

Sample test items from the *basic*, the *complementizer* and the *negation* conditions are given in (21), (22) and (23), respectively (see appendix 2 for a full list of the items).

Condition 1: *Basic*
(21) Sam likes to eat sweet things. Ten minutes ago he was eating a chocolate bar. Five minutes ago he was eating an ice-cream. Two minutes ago he was eating a piece of cake.

What do you think _____ right now?
    A. √ he is eating
    B. is he eating
    C. what he is eating
    D. what is he eating

Condition 2: *Complementizer*

(22) Anna and Martin want to buy a house in Ottawa. But they are not sure when to buy it. They are asking a real estate agent for advice.

When do you recommend _____ a house in Ottawa?

    A. √ that we should buy
    B. that should we buy
    C. that when we should buy
    D. that when should we buy

Condition 3: *Negation*

(23) Richard is buying a lot of video games these days. But he is spending too much money on video games so he's had to stop buying some other things.

What don't you think _____ these days?

    A. √ he is buying
    B. is he buying
    C. what he is buying
    D. what is he buying

Note that half of the 8 items in each experimental condition contained extraction of an object wh- phrase (i.e. *what,* or *who*) and the other half contained adjunct extraction (i.e. *when,* or *where*). [12] Furthermore, half of the items in each experimental condition contained *wh- copying,* and the other half contained *wh- scope marking.* Finally, to avoid repetition, the verb *think* in the matrix clause *What do you think…* was varied with the verbs *say, claim, recommend,* and *suggest.* Not all of these variations are shown in the sample items above (see appendix 2 for a full list of the items).

As pointed out in Table 3, the test also contained 8 control items. These were simple (mono-clausal) questions designed to control for knowledge of inversion (T-to-C movement) in matrix clauses, which is of crucial importance with regards to answering **R3** (see below for further details). [13] The control items were matched with the

---

[12] Subject extractions were not included as they do not allow for testing of inversion phenomena (T-to-C movement) in the embedded clause; that is, the key diagnostic to differentiate between a *direct* and an *indirect dependency* would not be available on such items. Furthermore, subject extractions would have made the task too long.

[13] Throughout the dissertation these items are called simple questions to reflect the fact they control for inversion in the matrix clause as opposed to the experimental items which test for different kinds of embedded clause. Nonetheless, the simple control items contain a second adjoined clause in order to approximate the complex questions in relative length and complexity.

experimental items for relative length and complexity. A sample control item is given in
(24).

Control Item
  (24) Barb is a teacher. She is not happy with her students. She thinks they are
  learning very slowly and always complaining. She is asking her colleagues:

Why _____so slowly and then always complaining about their teacher?
  A. √ are students learning
  B. students learns
  C. students learn
  D. students are learning

As illustrated in (24), the control items contained one grammatical choice with subject
auxiliary inversion (*A*) and three ungrammatical options without inversion (*B, C,* and *D*).

        Finally, the task contained 8 filler items. Unlike the rest of the items, these were
declaratives in order to introduce variety in the task. The fillers were always biclausal in
order to match the experimental items in relative length and complexity. A sample filler
is given in (25).

Filler Item
  (25) Ten years ago Jessica Johannes was killed in a nearby park. The police have
  finally captured the criminal.

The police captured the criminal _____Jessica Johannes 10 years ago.
                  A. √ that murdered
                  B. that murders
                  C. that murder
                  D. that he murdered

As illustrated by (25), the fillers contained a target utterance (*A*) and three ungrammatical
utterances (*B, C* and *D*) involving tense and/or agreement errors and resumptive pronouns.

        The experimental task was presented visually and auditorily to the participants.
First, the lead-in story/context would appear on a video screen accompanied by a
recording of a native speaker reading the story at normal speed. Then the story would
disappear from the screen and the multiple-choice item would appear. The item was not
accompanied by a voice recording because reading the different choices (*A, B, C,* and *D*)
would be tedious and unnatural, and would not allow the participants to review them at

their own pace. The response time was limited to 35 seconds for all items except for the *wh- scope marking* subset of the experimental items. The latter were slightly longer than the rest (see appendix 2), so the time allotted to each of them was 40 seconds. To warn the participants that an item was about to disappear from the screen, an animated *NEXT* sign appeared and was accompanied by a beep. The sign remained underneath the item for 4 seconds before the next screen appeared. The participants were provided with an answer sheet to mark the answer for each question (appendix 3). The entire task took about 35 minutes.

Before proceeding to the results, it is important to demonstrate that the experimental design adequately addresses the research questions. For convenience, these are repeated below.

R1. Do *medial wh-* representations exist in the interlanguage of French and/or Bulgarian learners of L2 English?

R2. If *medial wh-* is attested, is it of the *wh- copying* or the *wh- scope marking* kind?

R3. If *medial wh-* is attested, does it constitute a *direct* or an *indirect dependency*?

R4. If *medial wh-* is attested, can it co-occur with an overtly expressed complementizer *that*?

R5. If *medial wh-* is attested, is it subject to a negation asymmetry?

R1 is addressed by all three experimental conditions as two of the choices (*A* and *B*) are *[–] medial wh-*, while the other two (*C* and *D*) are *[+] medial wh-*. Thus, participants who choose *C* or *D* would provide evidence for *medial wh-* in the L2.

R2 is also addressed by all three experimental conditions as half of the items in each condition contain *wh- copying* in choices *C* and *D*, and the other half contain *wh- scope marking* in choices *C* and *D* (see appendix 2 for *wh- copying* versus *wh- scope marking* sample items).

R3 is also addressed in all the experimental conditions. The answer to this question is based on a the proposed link between presence of T-to-C movement in the embedded clause and an *indirect dependency* on the one hand, and lack of T-to-C movement in the embedded clause and a *direct dependency* on the other. In other words, if the participants postulate an *indirect medial wh- dependency* (i.e. two independent sequential questions), then T-to-C movement would apply to both clauses as they would

be treated as two separate matrix clauses (*choice D*). Conversely, if the participants postulate a *direct medial wh- dependency*, then T-to-C movement would not apply in the embedded clause (*choice C*).

It is important to recall that the above reasoning applies only if several crucial factors are taken into consideration. First, as it was mentioned in section 3.1, subject auxiliary inversion in the embedded clause can be used as a valid measure of an *indirect* versus *direct dependency* representation only if we know that the participants have acquired the English inversion rule. That is, participants who chose the *[+] medial wh-, [-] embedded inversion* option on the experimental items may simply be unaware of English T-to-C movement, and thus never invert in either matrix or embedded clauses. This is a distinct possibility, especially considering that T-to-C movement does not necessarily apply in the embedded clause of complex questions in the L1 (see example (11)*b* above). Second, one of the ways of forming simple questions in French is by using declarative (uninverted) word order and applying special intonation. Thus, it is plausible that this pattern may be transferred into the L2, and that the participants would not be likely to invert in any context (i.e. both matrix and embedded clauses).

Another issue to keep in mind is that participants who have acquired the subject-auxiliary inversion rule in matrix clauses may over-generalize it and thus apply it to embedded clauses as well. In this case if they choose the *[+] medial wh-, [+] embedded inversion* option on the experimental items, it would not necessarily mean that they posit an *indirect dependency*. This problem is further complicated by the fact that even though in standard English inversion in the embedded clause is ungrammatical, some varieties (e.g. Irish English) do use it at least in informal speech, as indicated in (26).

(26) I wonder what should we do.   (McCloskey 2006)

The above challenges, however, can be overcome with the help of the control items included in the study. As already mentioned, the control items are simple (monoclausal) questions where *choice A* (the target option) involves T-to-C movement and the three non-target choices (*B, C,* and *D*) lack T-to-C movement. Thus the experimental task provides a way of comparing the participants' performance on simple versus complex questions. In terms of **R3**, evidence for a *direct medial wh- dependency* would come only from participants who demonstrate high performance on the control

questions (i.e. identify the inverted *choice A* as correct) and at the same time choose the *[+] medial wh-, [-] embedded inversion* options (i.e. *choice C*) on the experimental items. Evidence for an *indirect dependency*, on the other hand, would come from participants who also have high performance on the control items, and at the same time choose the *[+] medial wh-, [+] embedded inversion* options (i.e. *choice D*) on the experimental items. This latter case could still be considered problematic if we assume that these participants overgeneralize the English inversion rule. However, participants who overgeneralize the inversion rule should also choose the *[–] medial wh-, [+] embedded inversion* option (choice B) on the experimental items, and thus this choice provides a way of controlling for overgeneralization (see the result section below for details on the specific control procedures).

So far, it was demonstrated that **R1**, **R2** and **R3** are addressed in all three experimental conditions. **R4** and **R5**, on the other hand, are only addressed in conditions 2 (*complementizer*) and 3 (*negation*), respectively. As already indicated, the 8 items in *condition 2* contain overt complementizer *that*, as in (22), and thus allow us to address **R4**. In other words, if participants choose *medial wh-* utterances in this condition, there will be evidence that the complementizer and the *medial wh-* word co-occur in their interlanguage. The fifth research question is addressed in *condition 3* where all 8 items contain negation in the matrix clause, as in (23). Thus, if a negation intervention effect applies to *medial wh-* in the L2, *medial wh-* would be less likely to occur in this condition than in the *basic* condition.

Note that **R4** and **R5** cannot be answered with regards to each of the 6 groups of participants listed in table 2 above. In the course of administering the experimental procedures it was decided that the experimental task may be too long and challenging for the beginner group, and at the same time too long and tedious for the upper-intermediate and advanced learners.[14] Thus, each of these groups received a shorter version of the test with 26 questions, where either the negation or the *complementizer* condition was omitted, the control items were reduced to 6, and the filler items were reduced to 4. The low-intermediate and intermediate groups, on the other hand, were given the full version of

---

[14] This was based on a suggestion by the ESL program staff which had to be accommodated for as per the terms of the research ethics approval for this project.

the test (40 questions), as it was deemed a meaningful ESL exercise suitable for their level.[15] The proficiency levels and the test version administered with each group are summarized in table 6.

Table 6. L2 proficiency levels and test version (experiment 1)

Group	Level	No. of participants	Test version
F1	high beginner	28	short (no *negation* condition)
F2	low intermediate	25	full version
F3	intermediate	22	full version
F4	high intermediate	24	short (no *negation* condition)
F5	low advanced	18	short (no *complementizer* condition)
F6	advanced	13	short (no *complementizer* condition)
Total full version		**47**	
Total short version (no *complementizer*)	**37**		
Total short version (no *negation*)		**46**	
Total all versions		130	

While the fact that not all participants received the same version of the test makes it difficult to draw comparisons for conditions 2 and 3 across certain groups, research questions 1 through 3, which are independent of these two conditions, remain unaffected. Furthermore, as will be indicated in the next section, the two groups that were given the full version of the test (group F2-3) had the highest and most systematic rates of occurrence of *medial wh-*, and thus the results of the other groups will provide mostly additive evidence.

3.2.3.   Results

3.2.3.1.   Statistical analysis

The statistical measures used in this dissertation are as follows: 1) logistic regression analysis of means (chi square) with factors *group, condition, copy* (*wh- scope* versus *wh- copy*) and *type* (arguments versus adjuncts). 2) Pair-wise two-tailed t-tests. The logistic regression analysis was used to determine the effect of the different factors on each of the four choices in the multiple-choice test. Logistic regression was also used to draw comparisons between groups.[16] The t-tests were used to determine the

---

[15] As per the approved research ethics procedures for this project, the participants were promised a meaningful ESL experience (i.e. an exercise that could potentially improve their English skills).

[16] Although some of these comparisons required less sophisticated statistical measures (i.e. when fewer factors were involved), logistic regression was used throughout for the sake of consistency.

significance of the differences in the rates of the four response types on the multiple choice test (i.e. *choice A* versus choice B, choice B versus *choice C*, *choice C* versus *choice D*, etc.). The software used was R$^©$ (version 2.5.0), and the codes and selected outputs from the logistic regression analysis are given in appendix 5.[17]

3.2.3.2.   *Group N (native controls, n=21)*

Although *medial wh-* constructions are starkly ungrammatical in all varieties of English, the possibility that native speakers might accept them needed to be excluded with absolute certainty. In addition, the native control group was necessary in order to test whether some of the other non-target options in each item of the grammaticality judgement multiple choice task could potentially be accepted. As, figure 1 indicates, the native speakers performed as predicted in rejecting *medial wh-* utterances in the three conditions (the 1% *choice C* responses in conditions 1 and 3 can be discounted as noise).

Figure 1. Complex questions: Total responses for choices A, B, C & D by condition, native control group, n=21
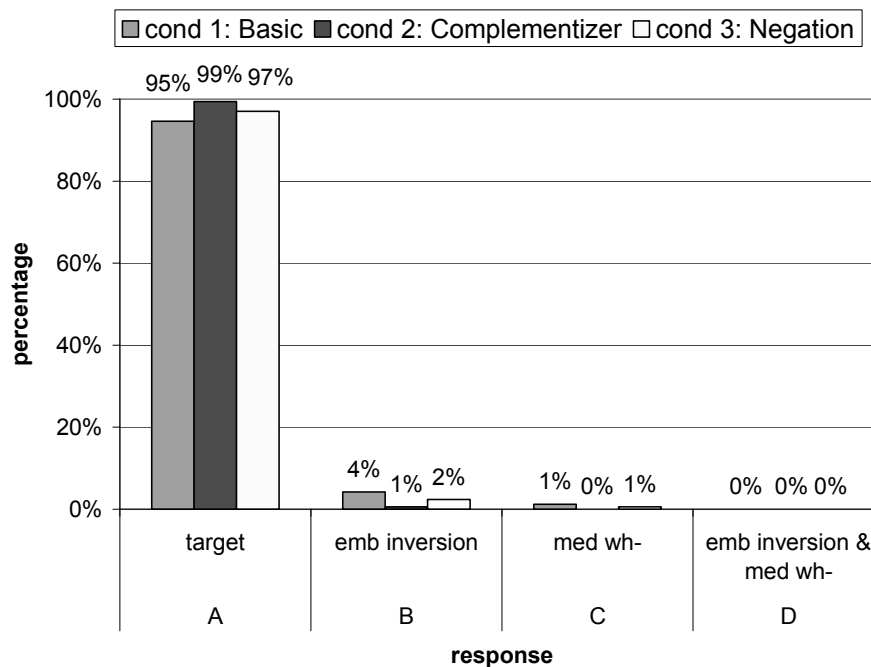
Note that the native speakers occasionally selected choice *B* (embedded inversion) which is an option that is potentially compatible with Irish English. However, none of these

---

[17] T-tests codes and outputs are not given in the appendix. They were performed using standard commands in a Microsoft$^©$ Excel 2002 spreadsheet.

participants reported to be a speaker of Irish English. As such, the choice *B* responses should be considered either noise (an average of 3% across all three conditions), or it could be that some North American English varieties may also feature T-to-C movement in the embedded clause. I leave this issue open to future research.

On the whole, the native controls performed at an accuracy level close to 100%, and thus provided a steady baseline against which the L2 learners' scores could be measured.

3.2.3.3.   *Group F2-3 (low)-intermediate (n=47)*

The presentation of the results for the non-native participants begins with groups F2 (low-intermediate) and F3 (intermediate) because they were given the full version of the experimental task, and the findings from these groups are most informative. Although the differences between these two groups were generally significant, their results are collapsed because the overall trends were identical. That is, group F2 had lower accuracy than group F3, but the distribution of the responses and the ensuing patterns were the same. From now on I will refer to this combined group of learners as *group F2-3*.

I begin by presenting the results for complex questions with the three experimental conditions combined. These are shown in figure 2.

Figure 2. Complex questions: Total responses by choices A, B, C, and D
Group F2-3, n=47
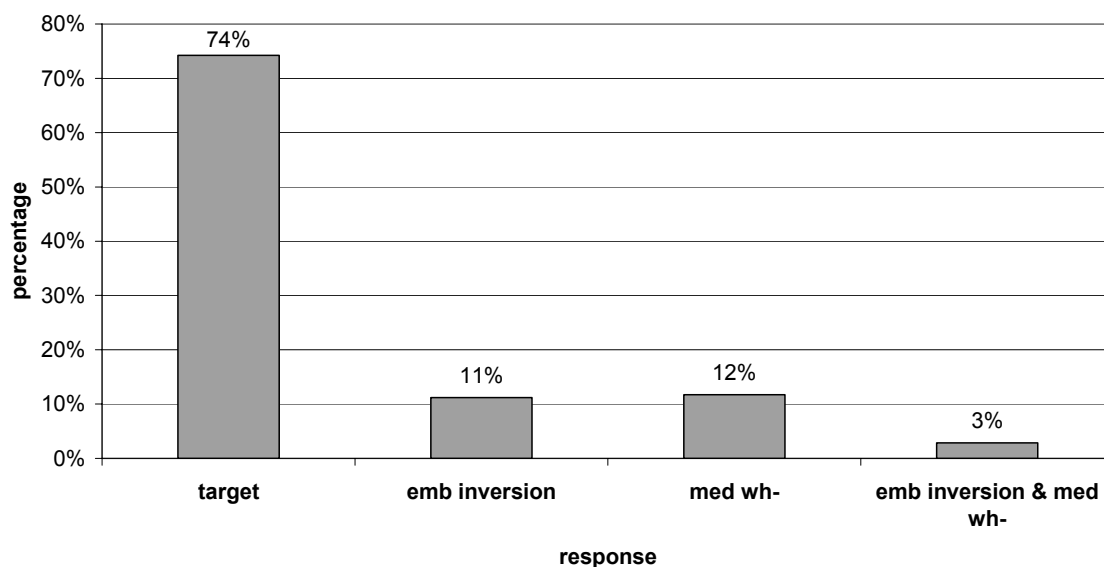
As these results indicate, the majority of the responses involve the target utterance (*choice A*); however, both inversion in the embedded clause (choice *B*) and *medial wh-* (*choice C*) are attested. The combination of embedded inversion and *medial wh-* (*choice D*), on the other hand, yields only 3% of the total responses. These results provide an answer to **R1**: *medial wh-* is attested at a rate of 15% (*choice C + D*) in the L2 English interlanguage of French L1 speakers. Overall, these responses indicate that the target representation (*choice A*) co-exists with the non-target representations (choices *B, C,* and *D*). Thus, it can be argued that the former competes with the latter but convincingly outranks them at this level of acquisition. Further development of this idea and discussion of its implications will be provided in section 3.2.4.

As for **R2**, a juxtaposition between the acceptance rates of *wh- copying* and *wh-scope marking* from all three conditions is shown in figure 3.

Figure 3. Wh- copy vs. wh- scope marking: Group F2-3, n=47
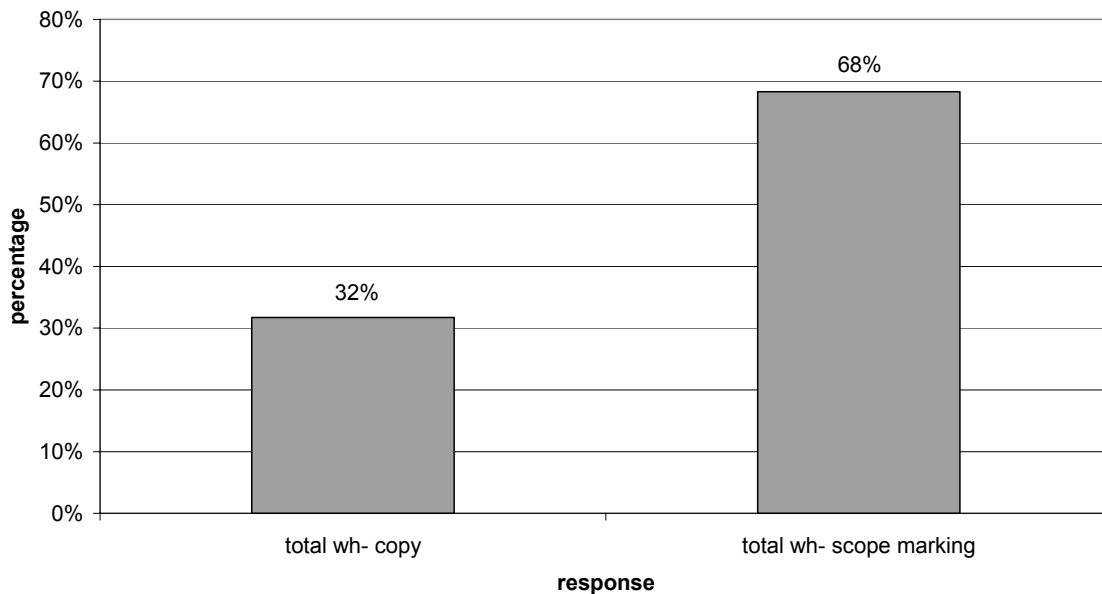
As these data indicate, 32% of all *medial wh-* utterances accepted as grammatical were *wh- copying* and the remaining 68% were *wh- scope marking*. A logistic regression revealed that this difference is significant ($p<0.0001$). This replicates previous L2 studies where *wh- scope marking* also occurred at a higher rate than *wh- copying*.

Note that in the coding process data items involving *what* in both the matrix and the embedded clause were considered to be *wh- copying*. This was done to preserve the

general insight that different wh- words in the matrix and the embedded clause constitute *wh- scope marking*, while utterances with identical wh- words in both clauses constitute *wh- copying*. However, utterances involving *what* in both the matrix and the embedded clause could also be analysed as *wh- scope marking*, in which the matrix *what* is an expletive or a phonological spell out of a wh- feature, and the embedded *what* is a contentful wh- word. If this is the case, the asymmetry in the distribution of *wh- scope marking* versus *wh- copying* in figure 3 would be even larger. This issue, however, is not pursued further. The key point is that *wh- scope marking* occurs at a much higher rate than *wh- copying* in the English interlanguage of French L1 speakers and this trend mirrors other L2 studies (for discussion see section 3.2.4).

Turning to **R3**, the question regarding the nature of the dependency, figure 2 shows a strong contrast between the *[+] medial wh-, [–] embedded inversion* option (12%) versus the *[+] medial wh-, [+] embedded inversion* option (3%). In fact, the latter percentage can be discounted as noise; thus, using presence or absence of embedded inversion in combination with *medial wh-* as a diagnostic for the nature of the dependency, it seems that there is evidence for a *direct dependency*. However, as explained in the previous section, this result can be misleading as the *[+] medial wh-, [–] embedded inversion* responses may be coming from participants who have not acquired the English inversion rule. It is therefore necessary to examine the results for the control (monoclausal) questions and identify the participants who have acquired T-to-C movement in matrix clauses. The relevant data are shown in figures 4 and 5.

Figure 4. Simple questions: Group F2-3, n=47
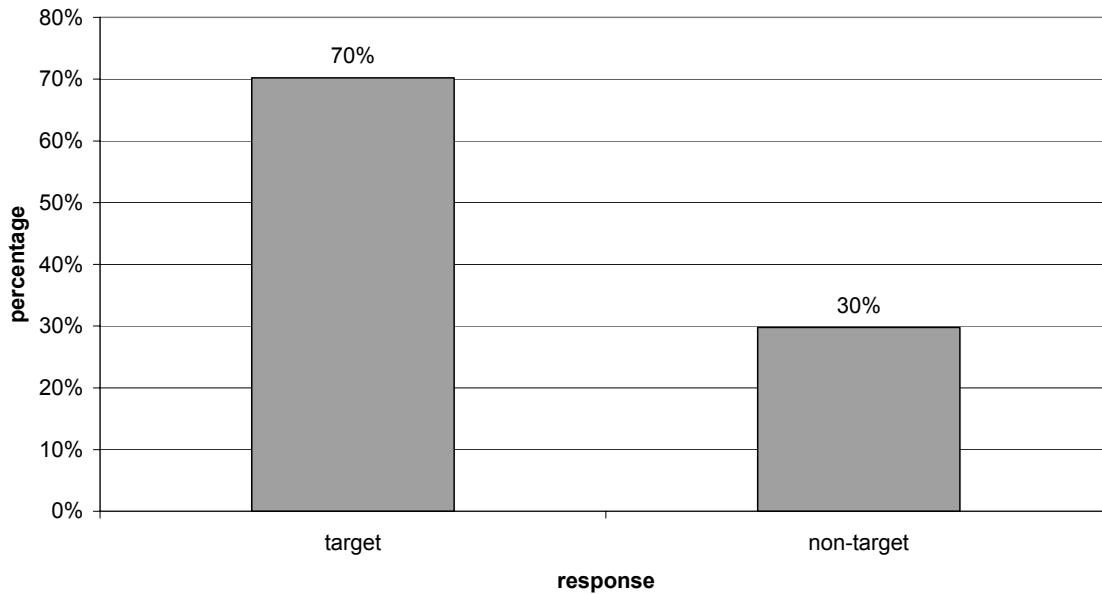
Figure 4 shows that 70% of the total responses on the simple (monoclausal) control questions involved correct T-to-C movement while the remaining 30% did not. The distribution of the 70% target responses is provided in figure 5.

Figure 5. Simple questions: Target responses (T-to-C movement)
Group F2-3, n=47
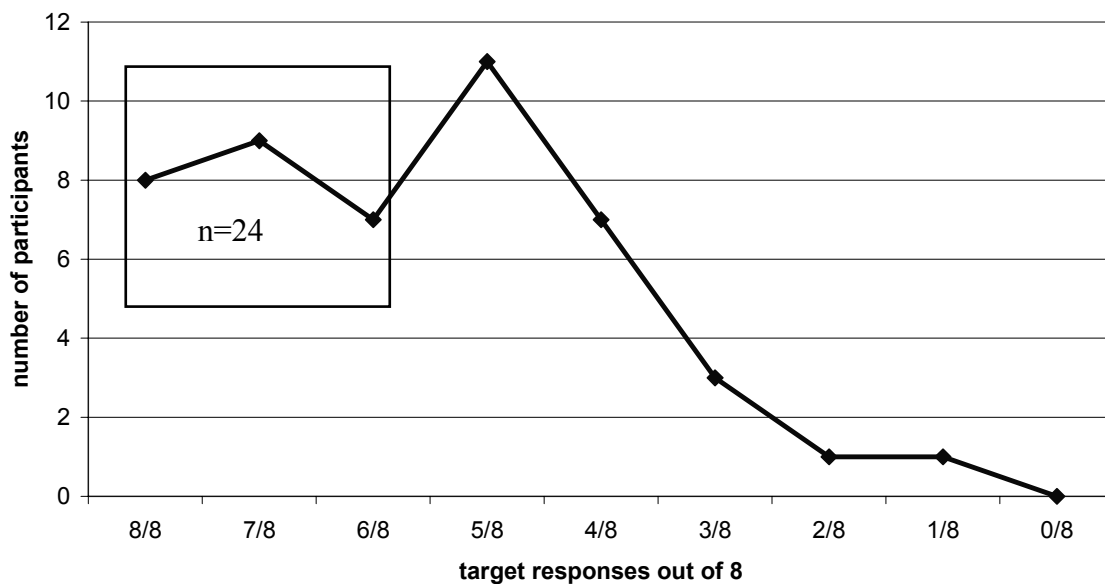
Recall that the task contained 8 control items. In order to consider a participant to have reliably acquired the target inversion rule in simple questions, a score of at least 6 out of 8 (i.e. 75% or higher) was set as a criterion.[18] Twenty-four of the 47 participants met this requirement, as shown in figure 5, while the rest were excluded for the purposes of **R3**.[19]

The next step was to re-run the analysis of complex questions for these 24 participants. The results are given in figure 6.

Figure 6. Complex questions: Total responses by choice A, B, C and D
Group F2-3, n=24 (participants who have acquired T-to-C movement)
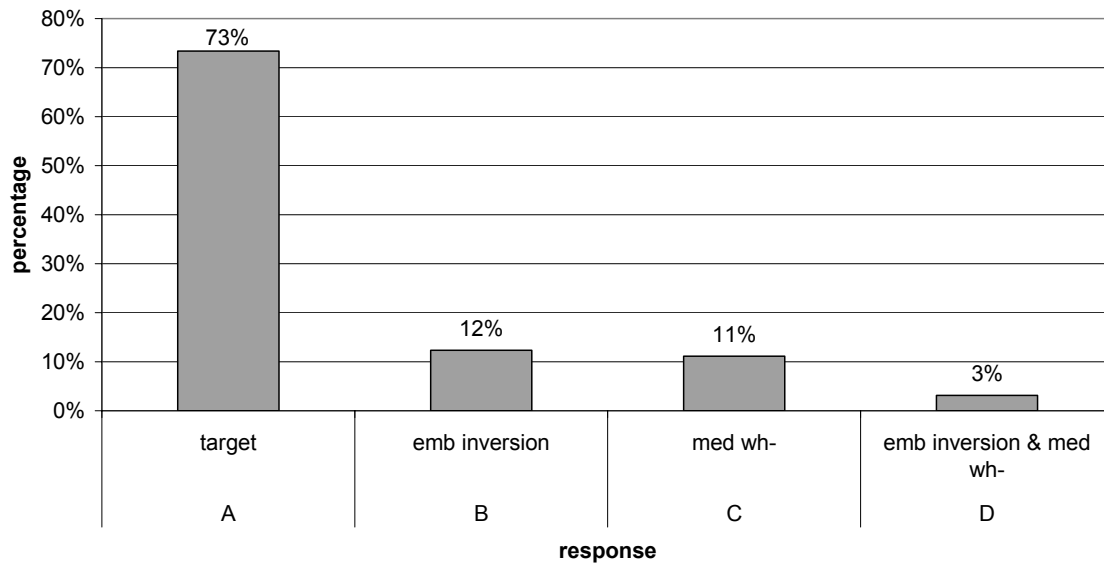
Interestingly, the distribution of responses on complex questions shown by the subgroup of participants who have acquired subject-auxiliary inversion in figure 6 essentially mirrors the overall pattern shown by all the participants in figure 2. However, it is only in figure 6 that a definitive answer to **R3** can be given. At this point the difference between the *[+] medial wh-, [–] embedded inversion choice C* (11%) and the *[+] medial wh-, [+] embedded inversion choice D* (3%) indicates that the *medial wh-* dependency attested is a *direct* one (a planned pair-wise two-tailed t-test comparison revealed that this difference is highly significant ($t(23)$=(4.6); p<0.001)). That is, the vast majority of *medial wh-* utterances that these 24 participants incorrectly identified as grammatical in L2 English

---

[18] In general, the criteria used for a particular construction to be considered successfully acquired are subject to substantial variation in the literature (see Zobl and Liceras 1994 for an overview with regards to L2 morpheme acquisition orders). The 75% criterion used here is quite high (see also footnote 20).

[19] For the purposes of the other research questions, the entire group of participants is used. This procedure is followed consistently with the rest of the groups as well, and also in *experiment 2*.

cannot be sequential questions. This finding indicates that such utterances constitute ungrammatical representations in both the L1 and the L2, and are thus an example of the learnability problem that is a central theme in this research. Potential explanations for this problem will be discussed later in this chapter. At this point, however, it is important to address the results of the remaining 23 participants who were not included in the subgroup that has acquired T-to-C movement. While the same strict criteria cannot be applied to their performance in order to determine if they posit a *direct* or an *indirect* *dependency*, it is still most logical to assume that they opt for the *direct* one: 18 of them had relatively high scores on the simple questions (between 50% and 63% inversion rate[20]); at the same time, their rate of inversion in the second clause of *choice D* responses (i.e. the sequential question representation) was 3% or lower. This difference strongly suggests that these learners do not analyse *medial wh-* constructions as sequential questions (for further evidence see individual results below).

Having answered **R3** and determined that the participants who accept *medial wh-* as grammatical in L2 English postulate a *direct dependency*, we can now turn to **R4** and **R5**. Since these were addressed in conditions 2 and 3 respectively, a break down of the results by condition is given in figure 7.

---

[20] Note that the cut-off used in Schulz (2006) was set at 50%, so these participants would have acquired T-to-C movement by her criteria. Also note that even though 50% correct is often associated with chance performance, this is not the case in a multiple-choice task with four options (as in the current study). That is, chance level performance on such a task would be 25% correct (i.e. 1 correct choice out of 4 choices per each multiple-choice item).

Figure 7. Complex questions: Total responses for choices A, B, C and D by condition, Group F2-3, n=47
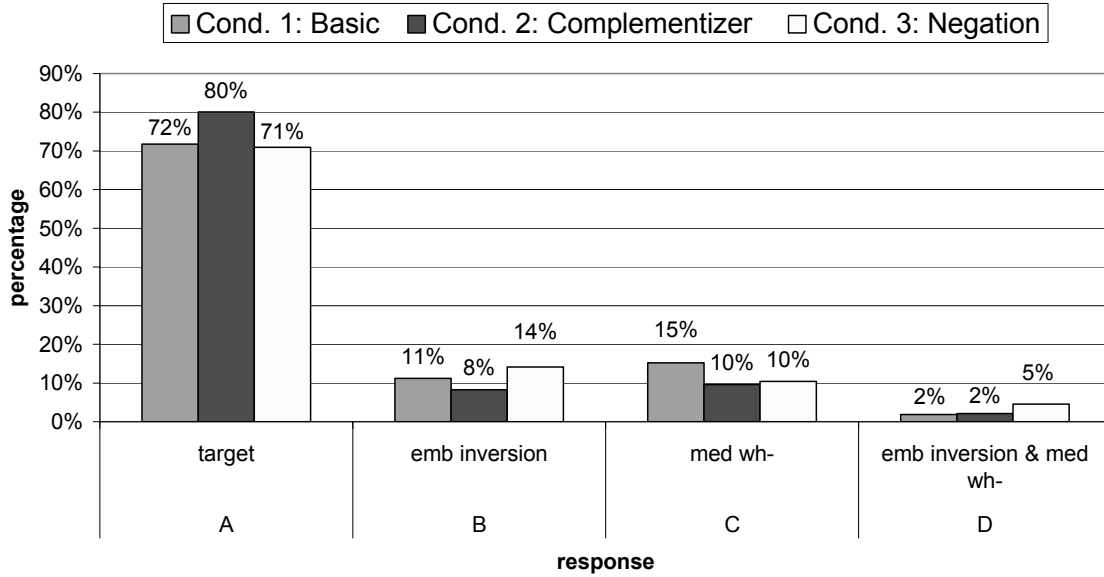
Recall that **R4** aims to investigate whether *medial wh-* can co-occur with an overt complementizer. Thus, in *condition 2* all items had an overt complementizer. Figure 7 indicates that *medial wh-* can co-occur with an overt complementizer, yet the occurrence rate is lower compared to the *basic* condition (i.e. the condition without an overt complementizer). The two choices relevant to this research question are *C* and *D*. In *choice C* the *medial wh-* rate is 15% in the *basic* condition and 10% in the *complementizer* condition. A logistic regression revealed that this difference is significant ($p<0.02$). In *choice D*, on the other hand, there is no difference between condition one and condition two (i.e. 2% for both conditions). This is most likely due to the overall very low percentage of responses with this choice across all conditions.

Turning to **R5**, we need to examine the results from *condition 3*, in which all items contained negation in the matrix clause. As figure 7 shows, the occurrence of *[+] medial wh-, [-] embedded inversion* (*choice C*) is lower in the *negation* condition than in the *basic* condition. A logistic regression revealed that this difference is also significant ($p<0.05$). Thus, the participants are less likely to accept *medial wh-* as grammatical when the matrix clause of the complex question is negated. Interestingly, however, the occurrence rate of the *[+] medial wh-, [+] embedded inversion* option (*choice D*) is higher in the *negation* condition (5%) than in the other two conditions

(about 2% each). Although very small, this difference is significant between conditions 1 and 3 (p<0.04), and thus it is worth noting that the presence of negation in the matrix clause somehow makes the participants select the *indirect medial wh- dependency choice D* slightly more often.

Apart from the five research questions postulated earlier, the results also showed an interesting asymmetry between arguments and adjuncts. Recall that the experimental design included 4 object and 4 adjunct wh- extractions in each of the three conditions (16 objects vs. 16 adjuncts in total). The object wh- words were *what* and *who*, whereas in adjunct extractions the wh- words were *when* and *where*. While no specific predictions about the distribution between arguments and adjuncts were made with regards to the two L2 learner populations, systematic argument-adjunct asymmetries have been observed cross-linguistically (see for example Huang 1982 and Rizzi 1990, among others). As such, the existence of such facts in L2 acquisition is particularly interesting. Figure 8 shows the distribution of arguments (objects) versus adjuncts for the three experimental conditions in the responses of group F2-3.

Figure 8. Complex questions: Arguments vs. adjuncts
Group F2-3, n=47
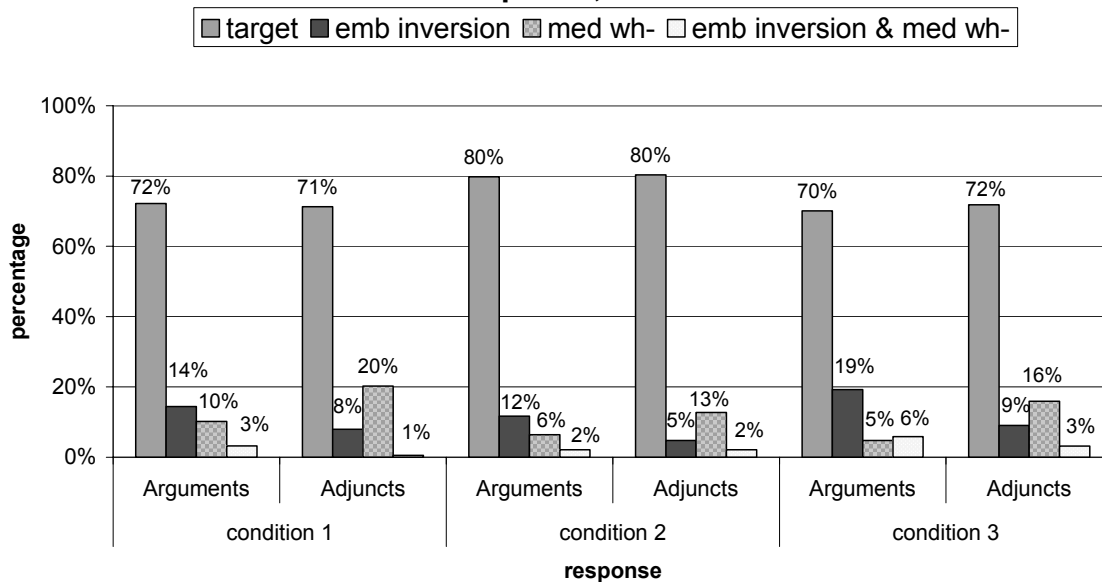
As these data indicate, the percentage correct answers for arguments versus adjuncts is almost identical in each condition (i.e. 72% vs. 71% for *condition 1*; 80% for *condition 2*; and 70% vs. 72% for *condition 3*). The distribution of errors, on the other hand,

consistently shows different patterns in argument versus adjunct contexts. The rate of occurrence of *medial wh-* (*choice C*) in all three conditions is higher with adjuncts than with arguments. A logistic regression revealed that this difference is significant (p<0.006 for *condition 1*, p<0.03 for *condition 2*, p<0.001 for *condition 3*). Interestingly, the reverse pattern can be observed for the *[+]embedded inversion, [-]medial wh-* option (*choice B*). The rate of occurrence of arguments in this option is higher in all three conditions. These differences are also significant (p<0.05 for *condition 1*, p<0.02 for *condition 2*, p<0.01 for *condition 3*). Overall, the data show a pattern in which the non-target *medial wh-* option (*choice C*) is more likely to occur with adjuncts, whereas the non-target embedded inversion option (*choice B*) is more likely to occur with arguments. A more detailed discussion and explanation as to why this asymmetry between adjuncts and arguments exists will be pursued in section 3.2.4.6 of this chapter.

To conclude the presentation of the group results for group F2-3, the key findings were as follows: 1) *medial wh-* was attested in the interlanguage of these learners; 2) *wh-scope marking* had higher rates than *wh- copying*; 3) the *medial wh-* representations were almost exclusively of the *direct dependency* type, which poses a learnability problem; 4) an argument-adjunct asymmetry was observed.

In order to gain a more thorough understanding of the data, the rest of this section focuses on the individual results. The distribution of the responses with regards to the two *[+]medial wh-* options of the test (choices *C* and *D*) is provided in table 7.

Table 7. Group F2-3: Individual medial wh- acceptance rates

Participant	CHOICE C [+]MED. WH-, [-]EMB. INV.		CHOICE D [+]MED. WH-, [+]EMB. INV.		SUM OF CHOICE C AND CHOICE D	
	number of responses out of 24[21]	%	number of responses out of 24	%	number of responses out of 24	%
21	7	29%	5	21%	12	50%
13	7	29%	3	13%	10	42%
2	7	29%	2	8%	9	38%
18	7	29%	2	8%	9	38%
17	6	25%	1	4%	7	29%
34	6	25%	0	0%	6	25%

[21] The vast majority of participants provided an answer to all 24 experimental items. However, there were a few who omitted some answers (usually not more than two); in those cases the percentages in the table are calculated accordingly (i.e. dividing by a number smaller than 24).

Table 7. Continued

Participant	CHOICE C [+]MED. WH-, [-]EMB. INV.		CHOICE D [+]MED. WH-, [+]EMB. INV.		SUM OF CHOICE C AND CHOICE D	
	number of responses out of 24	%	number of responses out of 24	%	number of responses out of 24	%
40	6	25%	0	0%	6	25%
43	5	21%	1	4%	6	25%
3	5	21%	0	0%	5	21%
37	5	21%	0	0%	5	21%
22	4	17%	1	4%	5	21%
6	4	17%	0	0%	4	17%
7	4	17%	0	0%	4	17%
33	4	17%	0	0%	4	17%
41	4	17%	0	0%	4	17%
19	3	13%	2	8%	5	21%
47	3	13%	2	8%	5	21%
20	3	13%	1	4%	4	17%
9	3	13%	0	0%	3	13%
30	3	13%	0	0%	3	13%
36	3	13%	0	0%	3	13%
12	**2**	**8%**	**4**	**17%**	**6**	**25%**
35	**2**	**8%**	**3**	**13%**	**5**	**21%**
11	2	8%	2	8%	4	17%
15	2	8%	1	4%	3	13%
5	2	8%	0	0%	2	8%
14	2	8%	0	0%	2	8%
16	2	8%	0	0%	2	8%
31	2	8%	0	0%	2	8%
39	2	8%	0	0%	2	8%
45	2	8%	0	0%	2	8%
46	2	8%	0	0%	2	8%
10	1	4%	1	4%	2	8%
42	1	4%	1	4%	2	8%
1	1	4%	0	0%	1	4%
4	1	4%	0	0%	1	4%
23	1	4%	0	0%	1	4%
24	1	4%	0	0%	1	4%
25	1	4%	0	0%	1	4%
26	1	4%	0	0%	1	4%
27	1	4%	0	0%	1	4%
28	1	4%	0	0%	1	4%
38	1	4%	0	0%	1	4%
8	0	0%	0	0%	0	0%
29	0	0%	0	0%	0	0%
32	0	0%	0	0%	0	0%
44	0	0%	0	0%	0	0%
mean	**3**	**12%**	**1**	**3%**	**3**	**15%**

As indicated by table 7, the *medial wh-* acceptance rates vary among the participants. While the total average acceptance of *medial wh-* is 15% (choices *C + D*), more than 20 participants' individual rates are higher, and a few of them reach levels of about 40-50%. This provides further evidence that *medial wh-* is a robust phenomenon in the interlanguage of these learners (i.e. cannot be considered noise).

An interesting question with regards to individual patterns concerns the dichotomy between the *direct* and the *indirect dependency*. Recall that the *direct dependency* (*choice C*) in the group results constituted 12% of the total responses. The *indirect dependency* (*choice D*), on the other hand, was attested at an average of only 3%. However, it is still possible that on the individual level there might be participants who consistently accept an *indirect dependency* (*choice D*). As table 7 shows, this scenario is not supported by the data: 43 out of the 47 participants accept the *indirect medial wh-dependency* (*choice D*) at a rate of 8% or less. The remaining four participants (21, 13, 12 and 35) accept the *indirect dependency* between 13% and 21% of the time. Two of these participants (13 and 21, in bold type) have not reliably acquired the English inversion rule, as established by the criteria for control questions discussed earlier. Therefore, one way of addressing their performance could be to propose that they accept an *indirect dependency* because they do not discriminate between inversion and lack of inversion in both matrix and embedded clauses. The other two participants (12 and 35, in bold type) have reliably acquired the target inversion rule in simple questions, so there is some evidence that they do indeed have a preference for an *indirect dependency*; however, it should be noted that they also accept utterances with a *direct dependency* at an only marginally lower rate, and thus it is possible that both *medial wh-* representations are active in their interlanguage. Also note that participant 12 has a very high score of inversion acceptance in embedded clauses (see figure 11 below) and thus it could be argued that this learner overgeneralizes the T-to-C movement rule. To summarize, the vast majority of the participants favour a *direct dependency* representation, which strongly supports the observations made in the group results.

Another issue that should be pursued with regards to the individual data is the distribution of responses in the *complementizer* and *negation* conditions. As the group

results indicated, *medial wh-* utterances (*choice C*) with complementizer *that* and with matrix negation were accepted as grammatical at a lower rate (both at 10%) than the equivalent *medial wh-* utterances in the *basic* condition (15%). While these differences proved to be significant, neither negation nor an overtly expressed complementizer seemed to produce a complete blocking effect on *medial wh-* acceptance (i.e. complete rejection of utterances with both *medial wh-* and an overt complementizer or with both *medial wh-* and matrix negation). However, it is possible that at the individual level there are participants who consistently accept *medial wh-* as grammatical in the *basic* condition and at the same time consistently reject it in the *complementizer* condition. Such behaviour would be consistent with Thornton's (1990) analysis of *medial wh-* words as complementizers. As for negation, if there are participants who consistently accept *medial wh-* in the *basic* condition and at the same time consistently reject it in the *negation* condition, then such participants would show evidence of an L2 grammar consistent with German *medial wh-* options (recall that negation blocks *medial wh-* in German, as demonstrated earlier in examples (7)-(8).

To begin with the question regarding complementizers, figure 9 shows the total individual acceptance rate of *medial wh-* utterances (choices *C + D*) for the *basic* condition and the *complementizer* condition.

Figure 9. Individual med wh- acceptance rates: Condition 1 vs. 2
Group F2-3, n=47
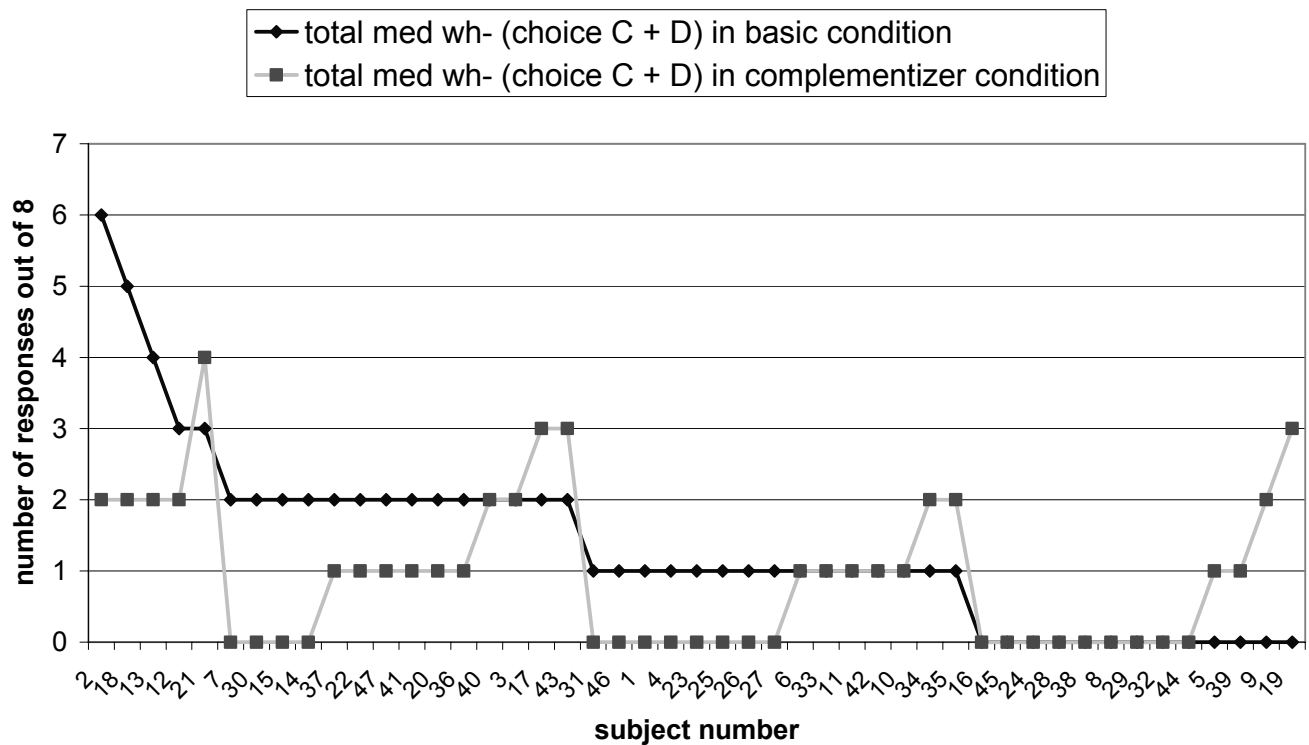
The data indicates that there are no participants who accept *medial wh-* at a high rate in the *basic* condition while consistently rejecting it in the *complementizer* condition. Participants 2, 18, and, to some extent, 13, show a trend in this direction but the pattern is not absolute; that is, these participants fail to fully reject the co-occurrence of a *medial wh-* word and a complementizer. Since this is only a suggestive trend, and since only 3 of the 47 participants show it, it cannot be pursued any further. At the same time, it is important to note that most participants seem to have a sensitivity to the presence of an overt complementizer, which makes them accept these utterances as grammatical at a lower rate. Overall, the individual results support the pattern observed with the group data: although an overt complementizer lowers the rate at which *medial wh-* is accepted by the learners, it does not produce a complete blocking effect. These results are different from Thornton's (1990) child L1 data, where an overt complementizer never co-occurred with *medial wh-*. As such, Thornton's analysis of medal wh- words as complementizers cannot be applied to the L2 data. However, an account in terms of a limited blocking effect of

the overt complementizer over the acceptance of *medial wh-* seems appropriate and will be pursued in section 3.2.4.2 of this chapter.

Turning to the comparison between the *basic* and the *negation* conditions, figure 10 indicates a pattern similar to the one in figure 9.

Figure 10. Individual med wh- acceptance rates: condition 1 vs. 3
Groups F2-3, n= 47
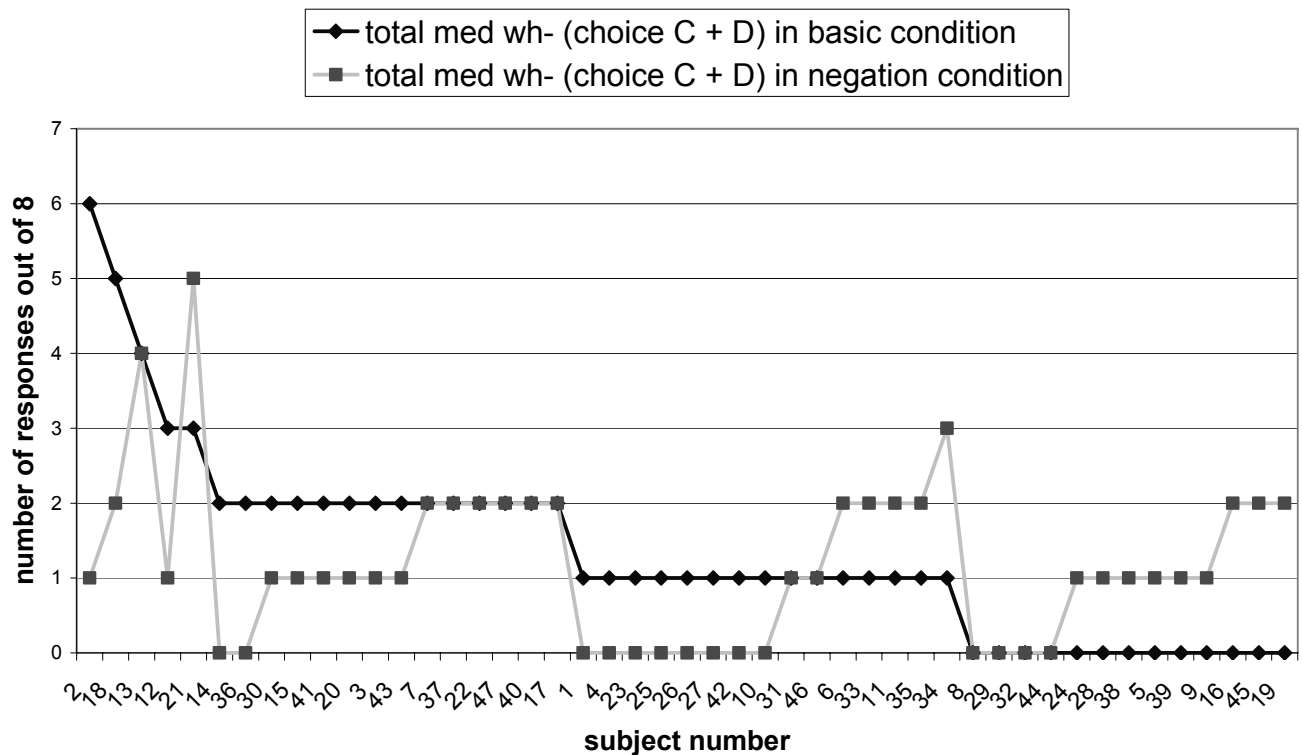
There are no participants who accept *medial wh-* as grammatical at a high rate in the *basic* condition and reject it consistently in the *negation* condition. Participants 2 and 18 again show a trend in this direction, but the pattern is not strong enough to warrant a particular analysis and only two participants out of 47 cannot be indicative of an overall trend. Note that, in fact, for 15 participants, the rate of acceptance of *medial wh-* is higher in the *negation* condition than in the basic one. However, the difference is by only one answer for most of these participants, and as such does not warrant any further consideration.

As indicated by the group results, negation does show a limited blocking effect (i.e. the acceptance rate of *medial wh-* is generally lower in matrix negation contexts) and

this finding is generally supported at the individual level. While the effect does not constitute full intervention as in the German and French examples presented earlier (see (7)-(9)), it seems to be a robust trend in the learners' interlanguage and will be discussed further in section 3.2.4.4.

Finally, it is also interesting to determine whether there are participants who overgeneralize the inversion rule (i.e. apply inversion consistently in both simple questions and in embedded clauses) as well as participants who have not acquired the inversion rule at all (i.e. almost never invert in either simple questions or embedded clauses). These results are given in figure 11.

Figure 11. Individual inversion rates in simple questions and embedded clauses, Group F2-3, n=47
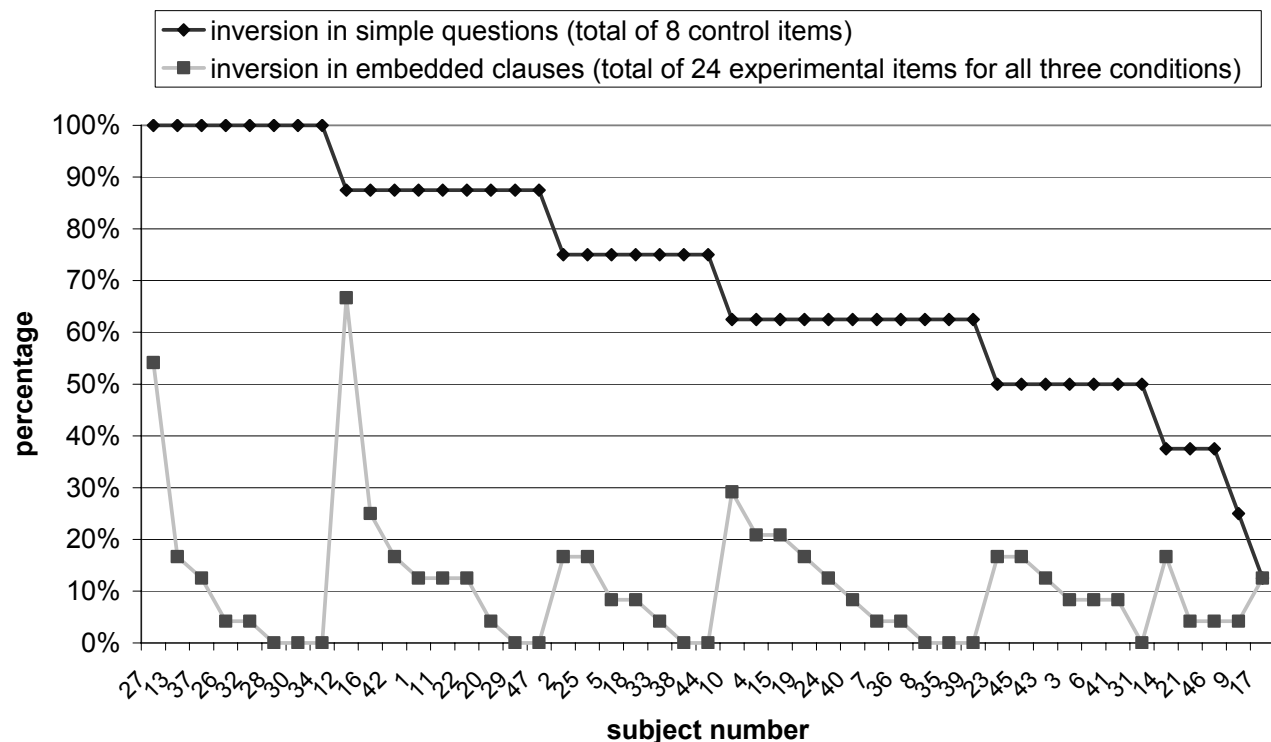
Participants 27, 12 and 10 show some evidence of overgeneralizing the inversion rule (i.e. relatively high rates of inversion responses in both matrix and embedded contexts), while participants 14, 21, 46, 9 and 17 show evidence of very low level acquisition of inversion in general (i.e. low rates of inversion responses in both matrix and embedded contexts). Overall, however, most of the participants seem to discriminate between simple and

complex questions, and therefore their inversion rates are higher in the former and lower in the latter, as already anticipated based on the group results.

This concludes the presentation of the results of the two groups of participants who completed the long version of the written experimental task. An overview of these results, the issues that they raise, and the predictions with regards to the remaining four groups of participants will be given in the intermediate summary section below. Before proceeding with this, however, I will point out how the group and the individual results presented in this section relate to each other. Essentially the observations made based on the mean scores were well supported and complemented by the individual results. The pattern of preference for *direct dependency* in the group *medial wh-* acceptance rates was also confirmed at the individual level. The co-existence of *medial wh-* representations with the standard LD wh- movement involving deletion of all intermediate copies was also attested in both the group and the individual results. The individual results offered a further insight into this finding by showing that there were no participants who accept *medial wh-* exclusively at the expense of the target LD wh- movement structure. The reverse pattern, however, was attested: there were learners who never accepted *medial wh-* and always opted for standard long-distance wh- movement.[22] This suggests that even though the target construction may be competing with *medial wh-* representations, the former is the winning candidate (for more details on this idea, see section 3.2.4). Finally, the observation that an overt complementizer and matrix negation in *medial wh-* constructions cause lower *medial wh-* acceptance rates but no complete blocking effect (i.e. no complete rejection of such utterances) was confirmed by both the mean scores and the individual patterns. Detailed explanations of these effects will be offered in the general discussion below. The conclusion at this point, however, is that the group and the individual results for group F2-3 are consistent with each other and offer cumulative support for the observations made so far.

---

[22] Note that some of these learners may have also chosen the embedded inversion option without *medial wh-* (*choice B*), but this option is considered standard LD wh- movement as it does not contain a wh- word in intermediate position.

3.2.3.4.   Intermediate summary

Before continuing with the results of the remaining four groups of French learners of L2 English, it is important to highlight some of the issues that the data presented so far raise. Since *medial wh-* was attested at a rate of 15% (choices *C* and *D*, figure 2) in the interlanguage of group F2-3, the question that arises is how to account for the existence of such utterances. One possibility is that the *medial wh-* word results from incremental processing effects. That is, due to processing pressures, especially during the early stages of acquisition, the *medial wh-* word provides an overt intermediate step in the derivation, which essentially shortens the dependency. This can be related to the idea of derivational complexity whereby more complex derivations (i.e. LD wh- movement) are more difficult to process than their less complex counterparts (i.e. short wh- movement). The intermediate stop that *medial wh-* words constitute in the long distance dependency (*direct dependency*, *choice C*) as well as splitting the structure into two sequential questions (*indirect dependency*, *choice D*) are both strategies that may provide an escape hatch in the cases where the standard LD structure with deletion of all intermediate copies is difficult to process due to, for example, memory load constraints. Phillips (1996) argues that derivational complexity views are compatible with the idea of a very strong interrelationship between grammatical competence and processing. Thus, in both processing and grammatical terms, derivational complexity can be applied to cases of stopping or slowing down at an intermediate step of the derivation or opting for some other simplified derivation of the original structure. I will revisit this idea later on in this chapter and in chapter 4.

To return to the results of group F2-3, recall that a subgroup of 24 participants who had reliably acquired T-to-C movement in simple questions was identified; in the results for complex questions of this subgroup, the *direct dependency*, which constitutes a representation unavailable in the L1 and the L2, was attested at a rate of 11% (*choice C*, figure 6). While this percentage is relatively low, it is robust in the sense that it cannot be discounted as noise and needs to be accounted for. Other optional phenomena occurring at similar relatively low rates have been found to have important implications in the acquisition literature. For example, low rates of root infinitives in child L1 Spanish, Catalan and Italian initially led some researchers (e.g. Guasti 1993/1994) to propose that

an RI developmental stage does not exist in these languages (in contrast to German, Dutch, Swedish, etc., in which RI rates in child language were much higher). However, Liceras, Perales and Bel (2007, see also references therein) found that an RI stage does exist in Spanish, Catalan and Basque and in fact the lower rates of RIs attested in these null subject languages have important implications in terms of a universal typology of root infinitives in child language.

Another question that arises from the results of group F2-3, which was placed at the (low)-intermediate level of proficiency, is whether the lower level group F1 would have a higher rate of *medial wh-* acceptance. On the other hand, the percentage of *medial wh-* utterances can be expected to decrease with the higher proficiency groups F4, F5 and F6. As the results below will demonstrate, these predictions will be borne out and thus *medial wh-* will be shown to be a phenomenon that occurs only at the early stages of L2 development. An explanation for this fact in terms of an input-driven UG account coupled with probabilistic learning will be proposed in the discussion section below. This account will take into consideration both the reasons for existence of *medial wh-* utterances in the L2 grammar and the mechanisms that allow the participants to "unlearn" such structures and converge with the target.

Finally, the results of group F2-3 showed an argument-adjunct asymmetry as well as lower rates of acceptance of *medial wh-* structures with overt complementizer and matrix negation, compared to their counterparts without a complementizer and negation. An account of these facts will be proposed after the results from the remaining four groups of French learners of L2 English are presented.

3.2.3.5.   *Group F1 (high beginner, n=28)*

The results of group F1 generally support the results of group F2-3, although the patterns show lower significance levels. This can be attributed to the fact that F1 was the lowest proficiency group in the experiment, and some of the participants may have occasionally resorted to guessing (despite the instructions not to do so). Nonetheless, the results are systematic and offer further evidence for the claims made earlier. Most importantly, *medial wh-* (choices *C* and *D*) was attested in the interlanguage of these speakers, as shown in figure 12.

**Figure 12. Complex questions: Total responses
by choices A, B, C and D, Group F1, n=28**
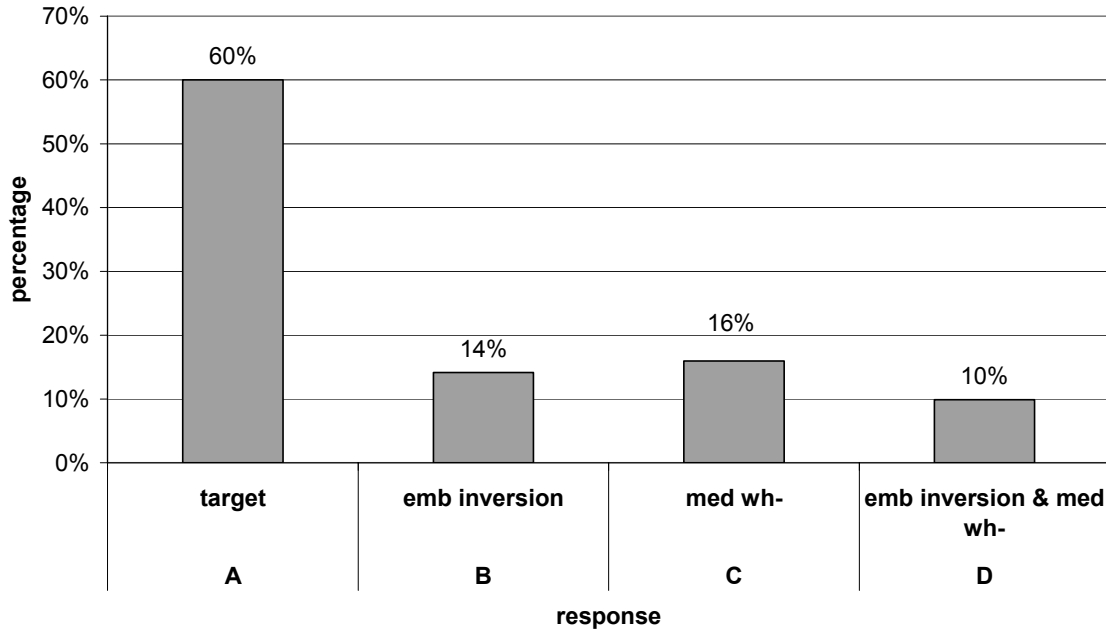
As the data indicate, the target representation (choice *A*) accounted for the great majority of the total responses, similarly to group F2-3. Inversion in the embedded clause (*choice B*) was attested at a similar rate as *medial wh-* with and without embedded inversion (choices *C* and *D*). As expected, due to the lower proficiency level, the overall accuracy rate (i.e. choice *A*) of this group was lower than that of group F2-3 (60% vs. 74%, respectively). On the other hand, the error rates of group F1 were higher than those of group F2-3 (14% vs. 11% on *choice B*, 16% vs. 12% on *choice C*, and 10% vs. 3% on *choice D*, respectively). A logistic regression showed that these differences are significant for choices *A* ($p<0.0001$), *B* ($p<0.03$), and *D* ($p<0.0001$), and close to significant for *choice C* ($p<0.08$).[23]

As for **R2**, the research question regarding the type of *medial wh-*, the scope marking construction was again attested at a higher rate than the *copying* construction, as shown in figure 13. Even though the difference is not significant, the pattern supports the one observed in group F2-3.

---

[23] Recall that group F1 received a shorter version of the test where *condition 3* was omitted. Therefore, the statistical comparison with group F2-3 is only for conditions 1 and 2.

Figure 13. Total wh- copy vs. wh- scope marking
Group F1, n=28
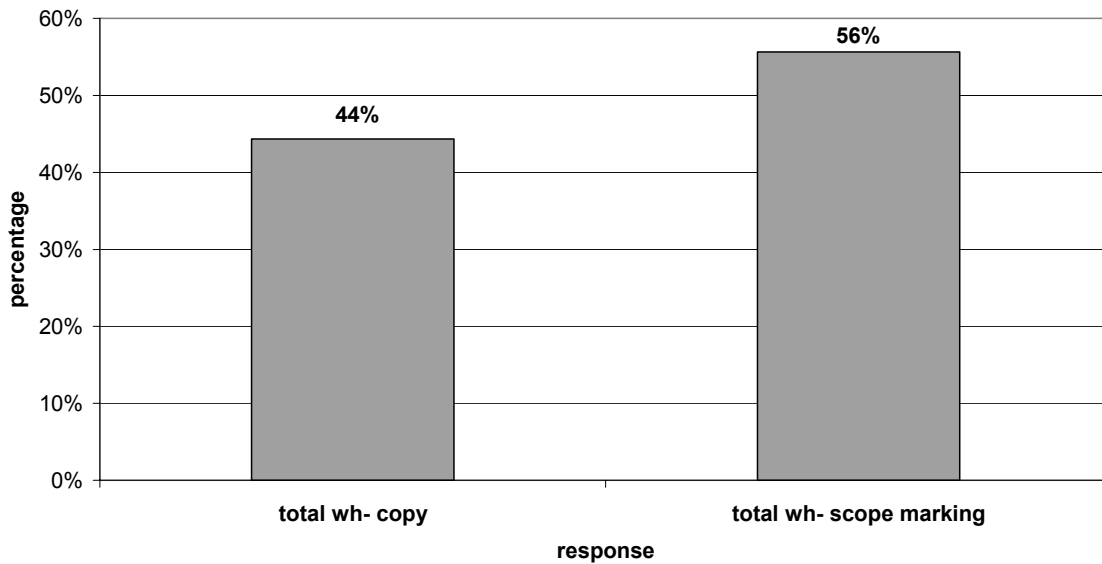
Interestingly, the question about the nature of the dependency, **R3**, receives a different answer with this group than with group F2-3. Recall that group F1 received a shorter version of the test where the *negation* condition was omitted, the number of control questions was reduced from 8 to 6, and the fillers were reduced from 8 to 4. Because of the different number of control questions, the criteria for a participant to be considered to have reliably acquired the English matrix inversion rule had to be adjusted. Thus, only learners who achieved a score of 5 out of 6 (83%) or higher were retained in investigating **R3**, and the rest were excluded. A total of 9 of the 28 participants met these strict criteria, and their responses to the experimental items are shown in figure 14.[24]

---

[24] The figures showing the scores on simple questions for this group (equivalent to figures 3 and 4 for group F2-3) are omitted due to space considerations.

**Fig. 14. Complex questions: Total responses by choice A, B, C, D
Group F1, n=9 (participants who have acquired T-to-C movement)**
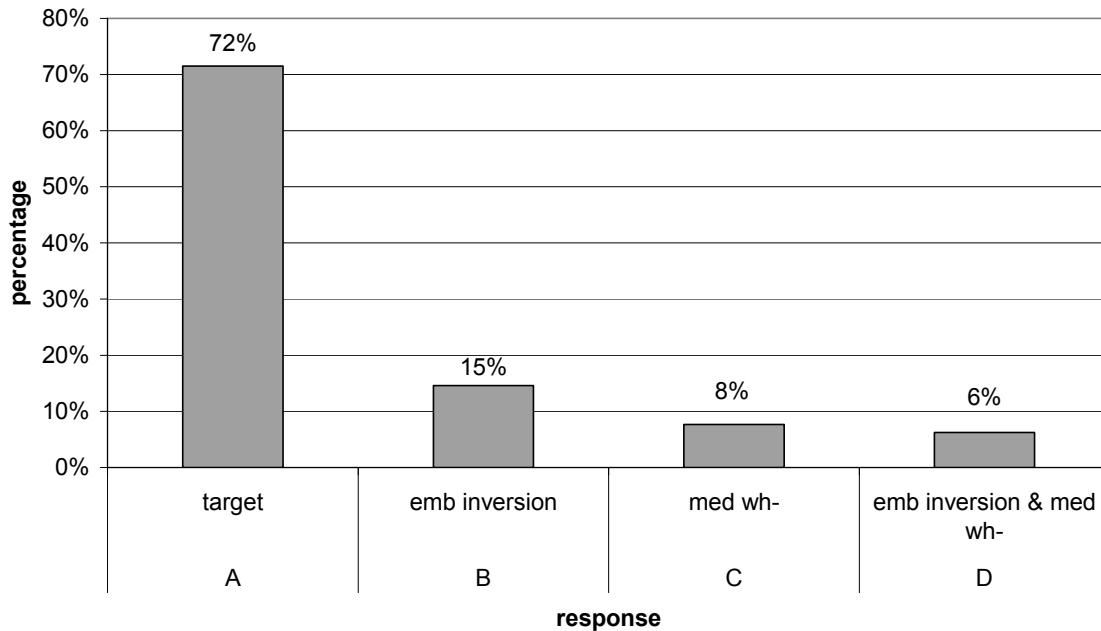
As the above data indicate, the 9 participants who have reliably acquired the T-to-C movement rule generally accept *medial wh-* to a lower degree than the remaining 19 participants in this group. However, following the pattern observed for the entire group, these participants opt for the *direct dependency* representation (*choice C*) and the *indirect dependency* (*choice D*) at a similar rate (cf. figure 14 and figure 12). Considering that the participants in the slightly more advanced group F2-3 showed an overwhelming preference for the *direct dependency* representation (*choice C*), the fact that the participants in group F1 attach almost equal value to the both options (*C* and *D*) is very interesting. It suggests either that these participants sometimes overgeneralize the inversion rule and apply it to embedded clauses, or that the *direct* and the *indirect* dependencies are competing representations with equal status in the interlanguage grammar at this level of proficiency. To investigate this question, a juxtaposition of the individual percentages of inversion on simple (control) questions and embedded clauses (*choice B*) is given in figure 15.

Figure 15. Individual inversion rates in simple questions and embedded clauses, Group F1, n=28
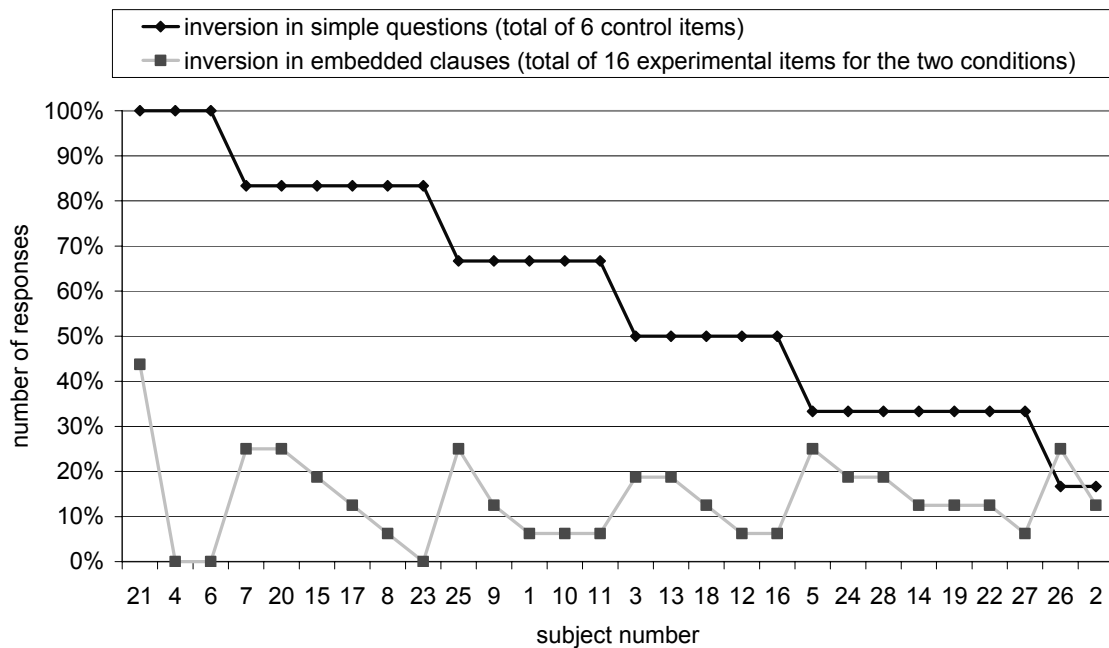
These data indicate that application of inversion in embedded contexts is at a relatively low level in comparison with inversion in simple questions. More than half of the 28 total participants in this group invert in the embedded clause (*choice B*) between 0% and 20% of the time, and at the same time apply the inversion rule in simple questions between 50% and 100%.[25] Particularly notable are participants 4, 6 and 23 whose acceptance of inversion in embedded clauses is 0%, while their acceptance of target inversion on the simple (control) questions is between 67% and 100%. On the other hand, 9 of the 28 participants accept inversion in both matrix and embedded clauses at a very similar rate (participants 5, 24, 28, 14, 19, 22, 27, 26 and 2). Note, however, that all these participants are placed at the rightmost end of figure 15, where the overall inversion patterns are very low (i.e. 33% or lower for simple questions and 25% or lower for the embedded clauses of complex questions). Thus, these 9 participants show evidence of systematic lack of inversion in matrix and embedded contexts rather than overgeneralization. Overall, these results show that the overgeneralization hypothesis has to be rejected in favour of the idea that both the *indirect* and the *direct dependency* representations (choices *C* and *D*) co-

---

[25] Recall that a similar but even stronger pattern was observed with the more advanced group F2-3 (cf. figure 15 and figure 11).

exist in the interlanguage grammar of the learners at this level of proficiency (further consideration to **R3** will be given at the end of this section).

The next research question, **R4**, regarding the acceptance of a complementizer co-occurring with *medial wh-*, receives an ambivalent answer. The comparison between conditions 1 and 2 provided in figure 16 shows that the rates of acceptance of *medial wh-* in both conditions are very similar.

Figure 16. Complex questions: total number of responses for choices A, B, C and D by condition, Group F1, n=28
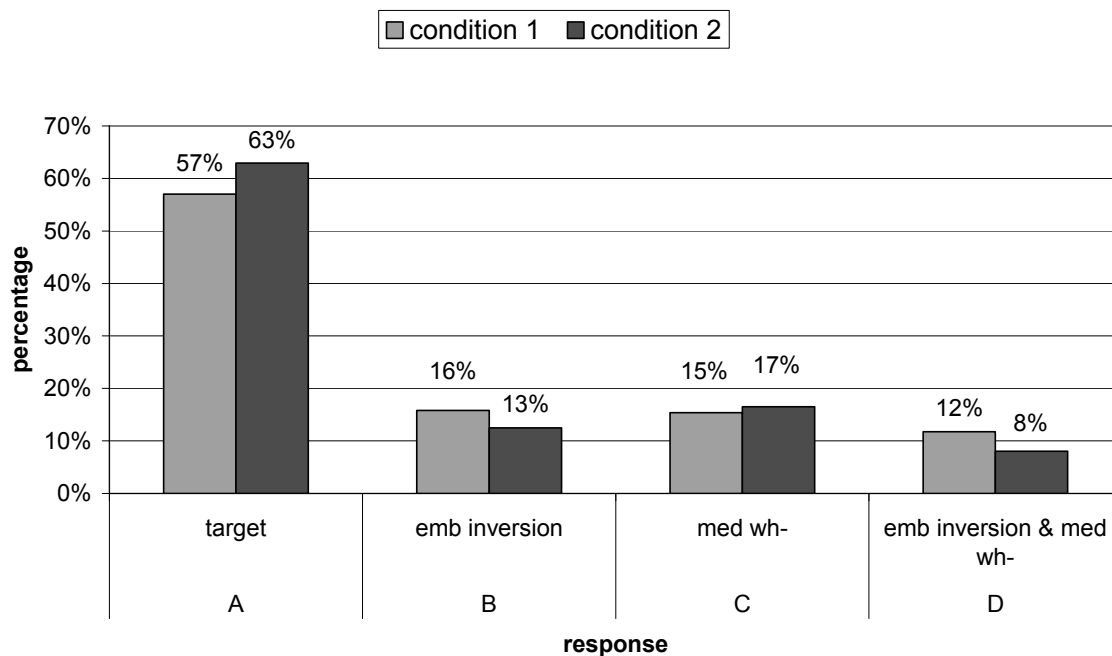
Thus, it seems that this group of learners does not distinguish between utterances with and without an overt complementizer. I assume that this is due to their lower level of proficiency, which causes the overall data to be less systematic and the distinctions between the conditions more difficult to emerge. Note, however, that these learners' overall accuracy increases in *condition 2*, which is a trend in the right direction (i.e. similar to group F2-3).

The next research question, **R5**, cannot be examined in the results of this group because the *negation* condition was not included in the version of the test that they received. As for the argument versus adjunct asymmetry which was observed in the results of group F2-3, the responses of group F1 are not significant in this respect.

However, as figure 17 indicates, there is a trend in *condition 1* where the *direct dependency medial wh-* representation occurs at a lower rate with arguments than with adjuncts.

Figure 17. Complex questions: Arguments vs. adjuncts
Group F1, n=28
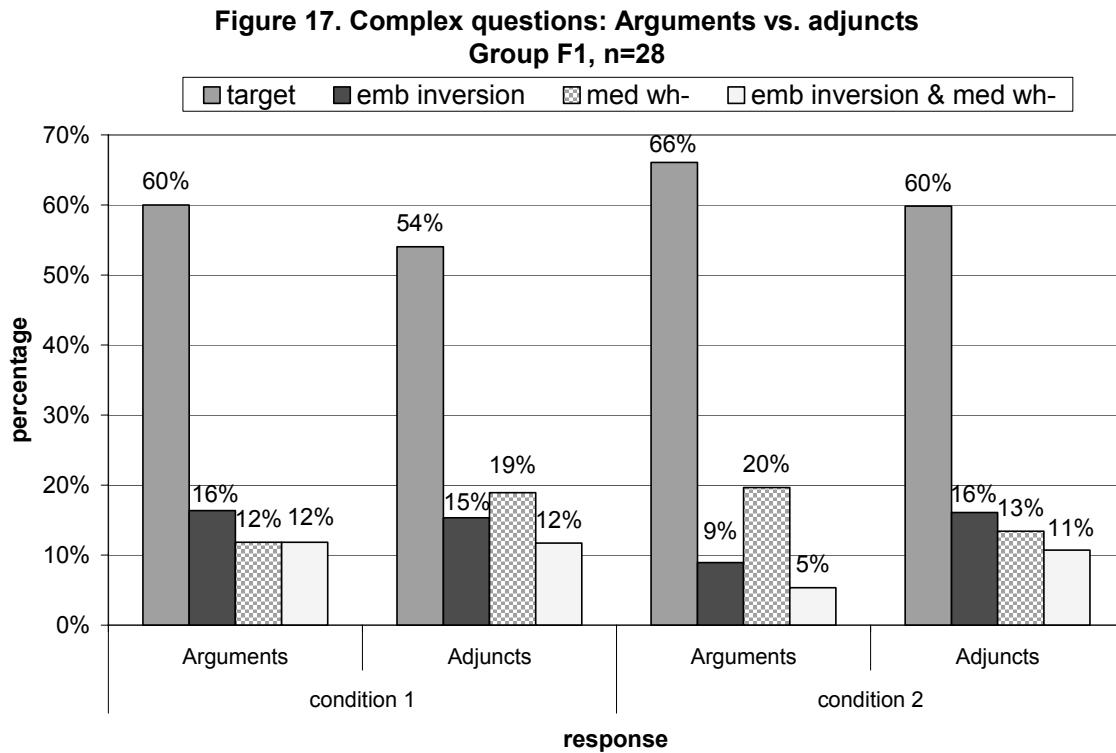
Surprisingly, the opposite trend (not significant) is observed with the equivalent responses (*choice C*) in *condition 2*. It is difficult to propose an explanation for this contradiction, apart from suggesting that with a slightly higher proficiency level and a larger sample size, this trend in *condition 2* would regularize (i.e. reverse) and mirror the highly significant data from group F2-3.

Finally, an overview of the individual *medial wh-* acceptance rates in group F1 is given in table 8.

Table 8. Group F1: Individual medial wh- acceptance rates

Participant	CHOICE C [+]MED. WH-, [-]EMB. INV.		CHOICE D [+]MED. WH-, [+]EMB. INV.		SUM OF CHOICE C AND CHOICE D	
	number of responses out of 16[26]	%	number of responses out of 16	%	number of responses out of 16	%
19	7	47%	6	40%	13	87%
1	6	38%	1	6%	7	44%
16	6	38%	1	6%	7	44%
14	6	38%	4	25%	10	63%
22	6	38%	0	0%	6	38%
12	5	31%	1	6%	6	38%
28	5	31%	2	13%	7	44%
24	4	27%	2	13%	6	40%
26	4	27%	1	7%	5	33%
4	3	19%	0	0%	3	19%
2	3	19%	0	0%	3	19%
17	2	13%	6	38%	8	50%
23	2	13%	0	0%	2	13%
18	2	13%	1	6%	3	19%
21	1	6%	2	13%	3	19%
6	1	6%	0	0%	1	6%
7	1	6%	1	6%	2	13%
15	1	6%	0	0%	1	6%
25	1	6%	2	13%	3	19%
10	1	6%	0	0%	1	6%
11	1	6%	0	0%	1	6%
13	1	6%	3	19%	4	25%
5	1	6%	9	56%	10	63%
27	1	6%	2	13%	3	19%
8	0	0%	0	0%	0	0%
20	0	0%	0	0%	0	0%
9	0	0%	0	0%	0	0%
3	0	0%	0	0%	0	0%
mean	3	16%	2	10%	4	26%

As these data indicate, the *medial wh-* acceptance rates in group F1 vary among the participants, similarly to the individual data for group F2-3. The total average acceptance of *medial wh-* is 26% (choices *C + D*), however, the individual rates of participants 19, 14, 17 and 5 are between 50% and 87%. On the whole, this group, which is at a lower proficiency level, has a higher *medial wh-* acceptance rate than group F2-3. This points to

---

[26] The vast majority of participants provided an answer to all 16 experimental items. However, there were a few who omitted one answer; in these cases the percentages in the table are calculated accordingly (i.e. dividing by 15, not by 16).

a correlation between proficiency level and presence of *medial wh-* constructions in the interlanguage grammar, a finding that will be further reinforced when the results from the other groups are presented.

Another important comparative finding is that the *indirect dependency* was virtually non-existent in the results of group F2-3, while the participants in group F1 showed equal acceptance rates of both the *direct* and the *indirect* dependencies. A key question that emerges with regards to the co-existence of the two dependencies in this group is whether they are distributed equally in the grammar of the individual participants, or if some participants favour the *direct* over the *indirect* one and *vice versa*. As table 8 shows, there are a three participants (1, 16, 22) who very strongly prefer the *direct dependency* over the *indirect dependency*; these participants' scores are 38% vs. 6%, 38% vs. 6%, and 38% vs. 0%, for the *choice C* versus *choice D* responses, respectively. Two of these participants (1 and 16) have inversion rates of 50% and 67% on the control questions, and thus, even though they do not meet the strict cut-off level of 83%, it is still likely that their high preference for [+] *medial wh-,* [–] *embedded inversion* is indicative of a *direct dependency* representation. Participant 22, on the other hand, has a very low level of inversion in both matrix and embedded contexts (see figure 15), and thus it is unclear whether his highly-attested *choice C* responses really represent a *direct medial wh- dependency* or are simply a consequence of a general lack of knowledge of the English T-to-C movement rule. Five other participants (12, 28, 26, 4, 2) show what looks like a preference for the *direct* over the *indirect dependency* that needs to be investigated further. Participant 4 has acquired the T-to-C movement rule perfectly (100% inversion in simple questions) and at the same time never applies this rule to embedded clauses (0% inversion in embedded clauses). Thus, this learner's preference for *choice C* over *choice D* responses is a firm indication of a *direct dependency medial wh-* representation. Participant 12 may also be showing evidence of a *direct dependency* representation. Although this learner's inversion rate on simple questions was not very high (50%), his inversion rates in embedded clauses were very low (only 6%), and thus it can be assumed that the learner distinguishes between the two. Participants 28, 26 and 2, on the other hand have low rates of inversion in both matrix and embedded clauses, and thus it is

unclear whether their preference for *choice C* over *choice D* answers reflects a *direct dependency* representation or an overall lack of knowledge of inversion.

The reverse individual pattern is also attested. Participants 17 and 5 have very high rates of *choice D* versus low rates of choice *C* responses. As figure 15 indicates, participant 17 has acquired T-to-C movement in simple questions and also knows that T-to-C movement does not apply in embedded clauses (83% simple vs. 13% embedded inversion rate). Consequently, the fact that this learner has high rates of *choice D* responses and low rates of choice *C* responses offers evidence in favour of the *indirect dependency* (sequential questions) representation in the L2 grammar. Participant 5, on the other hand, has very low inversion rates overall, and does not seem to distinguish between matrix and embedded contexts (see figure 15). Thus, it is unclear whether this learner's higher acceptance of *choice D* over choice *C* responses really reflects a preference for an *indirect dependency* or lack of knowledge of inversion.

Finally, there are also participants (19 and 14) who seem to favour both dependencies. As figure 15 indicates, these two participants are ambivalent with regards to the T-to-C movement rule in both matrix and embedded questions. Thus, it is difficult to decide if they really treat both dependencies on a par or if they simply accept both *choice C* and *choice D* because they do not distinguish between inversion and lack of inversion.

To recapitulate, the results of group F1 are less systematic than the ones for group F2-3. Nonetheless, they offer two important new findings: *medial wh-* acceptance rates are affected by proficiency; and both the *indirect* and the *direct* dependencies are attested with these lower-level learners.

3.2.3.6.   Group F4 (high intermediate, n=24)

Group F4 received a short version of the test, with the *negation* condition omitted. As figure 18 indicates, the results of this group are consistent with the results of group F2-3, which was at the previous proficiency level.

Figure 18. Complex questions: total responses for choices A, B, C and D by condition, Group F4, n=24
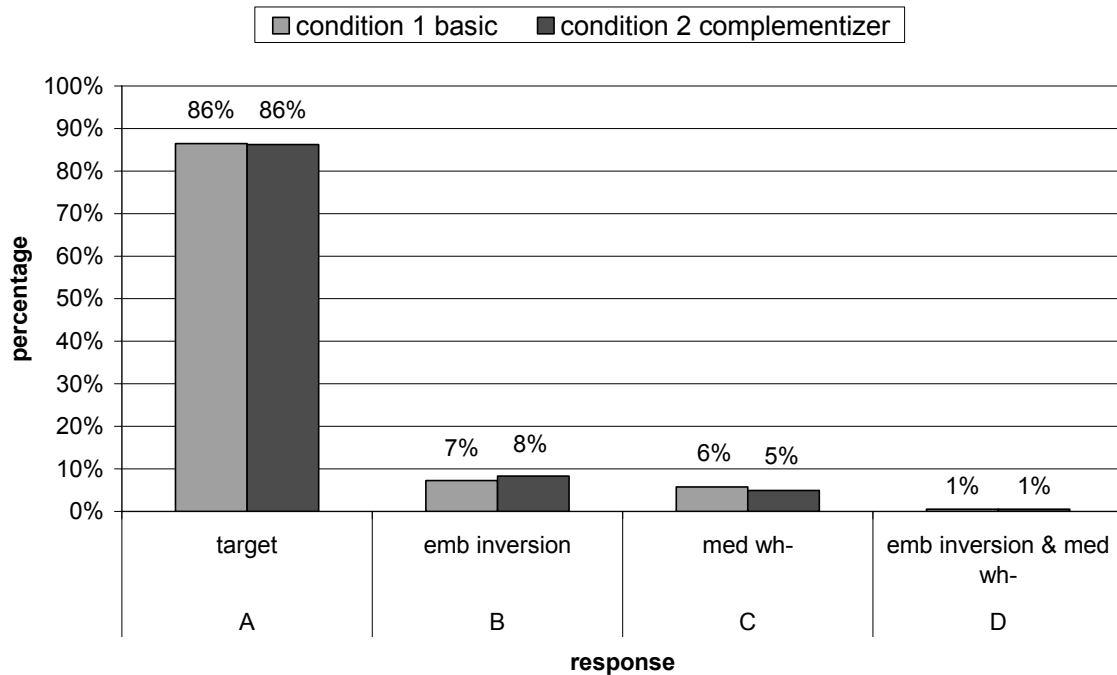
While the overall distribution of the responses and the ensuing patterns are the same in group F2-3 and F4, the latter shows lower rates for the non-target options and higher accuracy with the target construction. These differences (cf. figure 7) are significant for choices *A*, *C*, and *D* ($p<0.0001$, $p<0.0001$, $p<0.03$, respectively).[27] This confirms the correlation between level of proficiency and performance on the experimental task, which, as mentioned earlier, provides independent evidence of the validity of the placement test administered to all participants. Further evidence for this correlation will come from the results of groups F5 and F6, which will be presented in the following sections.

With regards to *medial wh-* representations, figure 18 shows that they are attested in the responses of group F4. However, the percentages become lower with the increased proficiency level, and thus pursuing the questions regarding the type of *medial wh-*, the nature of the dependency, the possibility of co-occurrence with an overt complementizer, and the argument-adjunct asymmetry becomes irrelevant. The crucial finding that is starting to emerge at this level, though, (a finding that will be reinforced by the results of

---

[27] Since group F4 received a short version of the test in which the negation condition was omitted, the statistical comparisons with group F2-3 are based on conditions 1 and 2 only.

the two remaining groups) is that *medial wh-* is not a representation that persists in the learners' interlanguage. Recall that *direct dependency medial wh-* constructions are not part of the participants' NL or TL. Thus, it seems that with enough experience the learners are increasingly able to realize that such constructions are not licensed in the L2 and are starting to exclude them from the interlanguage grammar. Interestingly, this process seems to be driven by indirect negative evidence, that is, lack of such constructions in the target input. This fact will be taken into account in the overall discussion of the results, and more specifically, in proposing an explanation for the learnability problem posed by *medial wh-* in the interlanguage of French learners of L2 English (section 3.2.4).

3.2.3.7.   Group F5 (low advanced)

Group F5 was given a short version of the experimental task, with the *complementizer* condition omitted. Based on the data presented so far, the main prediction with regards to the results of these participants was that the occurrence of *medial wh-* in their interlanguage would decrease further. As figure 19 indicates, this prediction was borne out.

Figure 19. Complex questions: total responses for choices A, B, C and D by condition, Group F5, n=18.
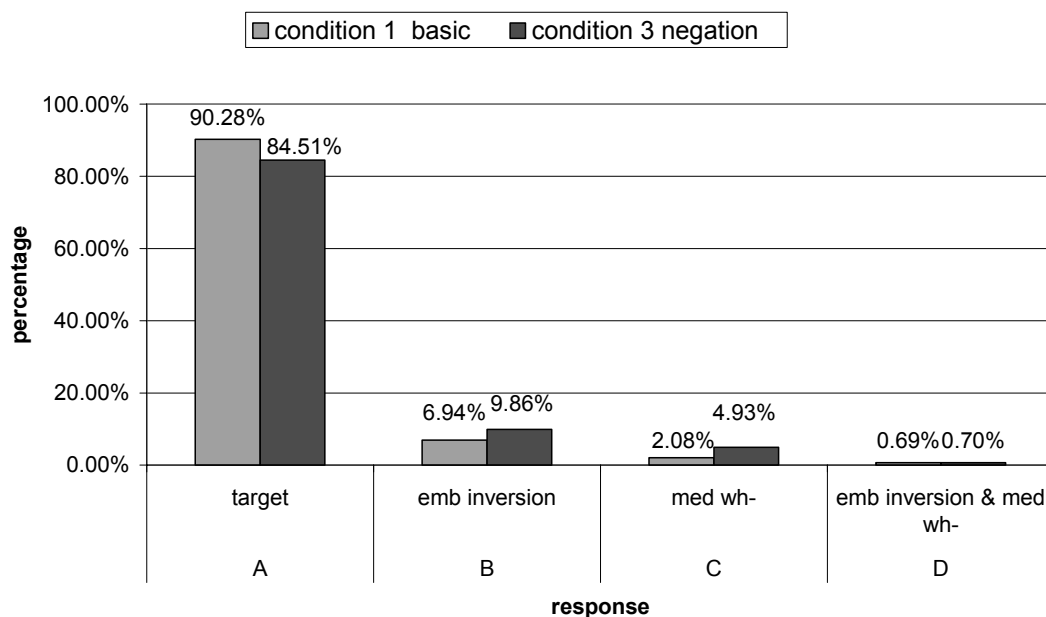
While only the difference in acceptance of *choice C* reaches a value close to statistical significance in comparison with the previous proficiency group (p<0.085), the pattern of increased accuracy and reduced error rates is uncontroversial across all groups.

Once again, the discussion of issues such as the type of *medial wh-*, the nature of the dependency, a potential negation intervention effect on *medial wh-*, and the argument-adjunct asymmetry is not possible in a population with such low error rates. The important finding, however, is that this group has almost excluded *medial wh-* representations from its interlanguage grammar.

3.2.3.8.   *Group F6 (advanced, n=13)*

The most advanced group in the experiment, group F6, was also given a short version of the test with the *complementizer* condition omitted. The inverse correlation between proficiency level and *medial wh-* is strengthened even further, as indicated by figure 20 (differences with figure 19 are significant for *choice A, p<0.005*, and for *choice B, p<0.01*, but do not reach significance for *choices C* and *D* due to the overall very low percentages attested for these responses at this level).

Figure 20. Complex questions: total responses for choices A, B, C and D by condition, Group F6, n=13
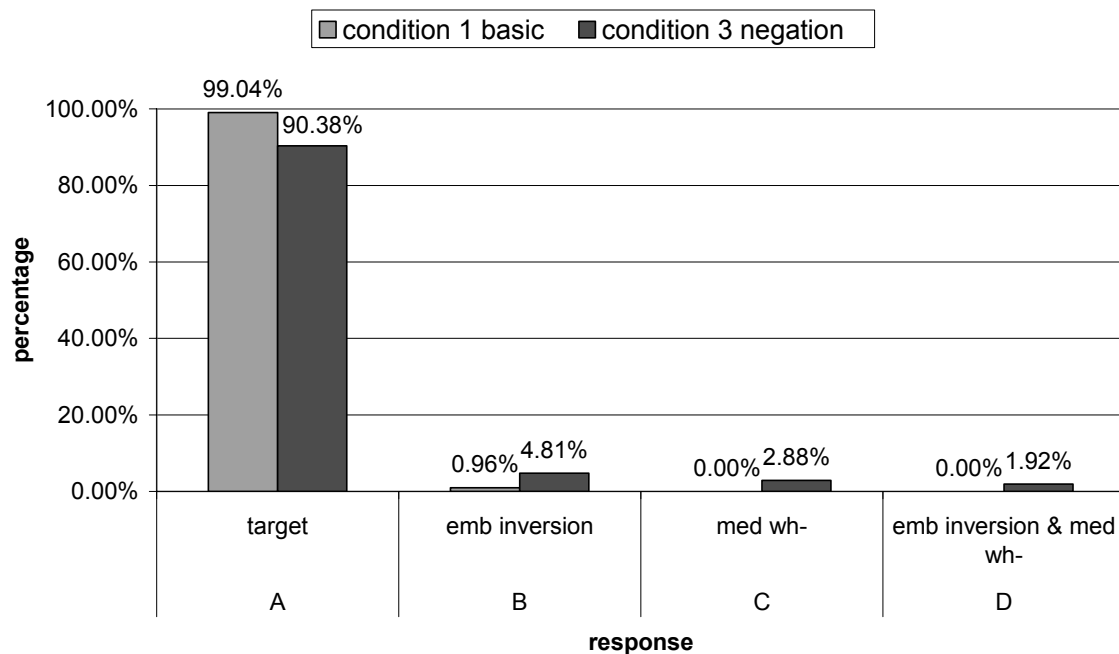
The participants in this group have essentially eliminated *medial wh-* from their interlanguage grammars and, in fact, their results converge with the native speaker data presented in figure 1 (i.e. the differences between the two groups are not significant[28]).

3.2.4. Discussion

The main goal of *experiment 1* was to determine if *medial wh-* representations exist in the L2 English of French speakers. The experiment was successful in revealing that such non-target constructions are present in the interlanguage of this population. The results from the lowest proficiency group F1 showed evidence of both the *direct* and the *indirect dependency* (sequential question) configuration. The higher level group F2-3, on the other hand, showed very strong evidence of the *direct dependency* configuration, which constitutes a learnability problem for this participant population. That is, such *medial wh-* constructions are indeed ungrammatical in both the L1 and the L2, yet attested in other natural languages, of which the learners report no knowledge. These findings raise several important questions: Why do such representations occur in the interlanguage grammar? What is their syntactic and psychological status? What insights do they offer in terms of second language acquisition theory and processes?

3.2.4.1.   Competing representations

Since *medial wh-* was attested mostly in groups F1 and F2-3, and its occurrence decreased significantly with the more advanced learners, there is evidence that the phenomenon disappears from the interlanguage with the increase of proficiency. Interestingly, even with the lowest proficiency group, *medial wh-* was not pervasive and was decidedly outranked by the target construction involving long-distance wh-movement of a single wh- phrase and deletion of all intermediate copies. This indicates that *medial wh-* is not an error that low proficiency learners make prior to acquiring long-distance wh- movement in the L2, but rather that the two constructions co-exist. Thus, the

---

[28] Although the logistic regression output comparing groups N(ative) and F6 showed that the difference in *choice D* (0% vs. 1.92%, respectively) is significant, this result is not meaningful. That is, the mathematical significance value is based on the difference between two tokens of this type of response for the entire group F6 and zero tokens for group N, and as such cannot be considered valid. This is also confirmed by the logistic regression estimates table included in appendix 5.

overall picture that emerges from the data is one in which different representations are competing in the participants' interlanguage. The target long-distance option seems to be "winning" the competition at all levels of proficiency. This can be explained by the fact that long-distance wh- movement mechanisms exist in both the NL and the TL, and thus the learners make extensive use of them. In addition to the target construction, however, the less experienced L2 learners also hypothesize the existence of a *medial wh-* representation supported by UG but not instantiated in either the L1 or the L2. In other words, during the early stages of acquisition, the learners seem to be making conjectures about the L2 which are not supported by the target input, but are within the hypothesis space of Universal Grammar.

Apart from *medial wh-*, another competing representation attested in the data was the non-target option involving inversion in the embedded clause (*choice B*). Non-target inversion on its own is not of key interest to this dissertation, and as such, the responses containing *choice B* in the test were not discussed extensively in the previous sections. However, the existence of this additional competing representation (at similar rates as *medial wh-*) needs to be accounted for. There are several possibilities with regards to this construction. First, it could be that learners overgeneralize the English T-to-C movement rule in matrix clauses and apply it to embedded clauses as well. This was in fact attested with a handful of learners who showed high inversion rates in both the simple (control) questions and in the embedded clauses in the experimental questions. However, this explanation, could not fully account for the data, as shown by the individual results of groups F1 and F2-3. In addition, if overgeneralization of inversion played such a prominent role, then the participants should have also opted for the *choice D* error to a higher extent because it also contained inversion in the second clause. Instead of doing this, however, the participants generally chose *medial wh-* without embedded inversion (*choice C*), and the option containing *medial wh-* with embedded inversion (choice *D*) was attested at a much lower level, often close to zero in group F2-3.

A second potential explanation for the presence of the embedded inversion representation (*choice B*) comes from L1 transfer. Although, none of the examples of French complex questions given earlier in (11)-(15) and in (19)-(20) contained inversion

in the embedded clause, this option is in principle possible in (Quebec) French, as illustrated in the contrast between (27)*a* and (27)*b*.

(27) a. *Que penses-tu   que   je doits faire?*
        what think-you   that   I   have to do
        "What do you think I should do?"
    b. *Que penses-tu   que   **doits-je** faire?*[29]
        what think-you   that   have-I   to do
        "What do you think I should do?"

In addition to (Quebec) French, embedded inversion is attested in other natural languages (e.g. Spanish, some varieties of Irish English, etc.) and is thus within the possible hypothesis space of Universal Grammar. By this token, an explanation in terms of the L2 learners accessing UG is possible as well.

To recapitulate, the data from *experiment 1* fit very well in a theory where different competing representations, constrained by the principles of UG but not necessarily instantiated in the native language or the target input, co-exist in the interlanguage grammar. Gradually, with enough experience, the learners are able to eliminate the representations that are not licensed in the TL and retain the ones that are licensed. One such approach has been proposed by Yang (2002, 2004, 2008) for child language acquisition (see also Roeper 2000) and can be adopted successfully for L2A as well. In this model, dubbed 'variational', Yang combines insights from both the innateness view and from approaches relying on the role of non-linguistic learning mechanisms, and more specifically *probabilistic learning*. The proposal is illustrated schematically in (28).

(28) For an input sentence, *s,* the child:
    (i)   with probability P_i selects a grammar G_i
    (ii)   analyzes *s with G_i*
    (iii)   if successful, reward G_i by increasing P_i,
            otherwise punish G_i by decreasing P_i.

Essentially, the idea is that competing representations have different probabilities, which are largely determined by the target input. Eventually, representations that are not rewarded by the input are "punished" by the learners through a decrease in their

---

[29] Many thanks to Sylvie Rondot and Nicolas Talbot for providing me with these data.

probability values. Thus, the language learner's grammar gradually converges with the TL.

Note that Yang is assuming multiple grammars, whereas I assume one interlanguage grammar with competing representations; nonetheless, the approach is compatible with the present data. As Yang (2004) explains, the model is based on insights from biological evolution in the sense that grammatical representations compete with each other in a Darwinian fashion; as such, the process of language 'growth' is necessarily constrained by the biological endowment represented by UG. At the same time, acquisition also entails probabilistic learning mechanisms, which have received strong support in recent studies (see, for example, Saffran, Aslin and Newport 1996, Aslin, Saffran and Newport 1998). Thus, in addition to the biological language endowment, the language learner computes statistical information provided by the input and shapes the target grammar (or chooses a target grammar) accordingly.

Adopting this approach helps account for the data presented above in a very natural way. I propose that the French speaking learners of L2 English rely both on their innate, subconscious linguistic knowledge and on non-linguistic, non-modular learning mechanisms. They favour the target representation of long-distance wh- movement because of three independent factors: the presence of this representation in their L1; the presence of this representation in the target input; and the fact that this representation is a licensed UG option. The combination of these three factors has a powerful impact in making the learners attach a high probability value to this representation in the interlanguage grammar; in practical terms, this means that in the experimental task they select this option at the highest rate. At the same time, the less experienced learners, who have not received enough target input to rule out other representations possible within the UG hypothesis space, attach a significant but lower value to such options. Because the *direct dependency medial wh-* representation is not supported by either the native language or the target input, its existence in the interlanguage grammar is based solely on the learners' innate language instincts. Thus, the probability attached to it is much lower, and this representation is "punished" in the interlanguage grammar by gradual lowering of its occurrence until it is completely eliminated at the higher levels of proficiency.

As mentioned earlier, the learners rely on indirect negative evidence in excluding the non-target *medial wh-* representation from the L2. That is, eventually they realize it is not a viable option because it is never instantiated in the input. This idea has been advanced consistently in generative grammar (e.g. Chomsky 1957, Chomsky 1981), however, a precise implementation comes from more recent statistical modelling. For example, using a Bayesian framework, Tenenbaum and Griffiths (2001) discuss hypothesis testing in the learning process as follows:

> The learner's knowledge about the consequential region is represented as a probability distribution $p(h|x)$ over an a priori-specified hypothesis space H of possible consequential regions $h \in H$. H forms a set of exhaustive and mutually exclusive possibilities; that is, one and only one element of H is assumed to be the true consequential region for C[onsequence] (although the different candidate regions represented in H may overlap arbitrarily in the stimuli that they include). The learner's background knowledge, which may include both domain-specific and domain-general components, will often translate into constraints on which subsets of objects belong to H. […] At all times, the learner's knowledge about the consequential region consists of a probability distribution over H. Prior to observing x, this distribution is the prior probability $p(h)$; after observing x, it is the posterior probability $p(h|x)$. As probabilities, $p(h)$ and $p(h|x)$ are numbers between 0 and 1 reflecting the learner's degree of belief that h is in fact the true consequential region corresponding to C. […] The probability of any h that does not contain x will be zero, because it cannot be the true consequential region if it does not contain the one observed example. (p. 631)

Essentially, in this model learners formulate hypotheses within a given hypothesis space defined by domain-specific or domain-general components. Under the account adopted in this dissertation, the L2 learners' hypothesis space is restricted by UG. Initially, learners postulate hypothesis h that phenomenon x (i.e. *medial wh-*) is a possible representation in the target grammar. Prior to observing *medial wh-* in the input, the learners assign a probability p to their hypothesis based on the degree of belief that h is true; thus, the prior probability is $p(h)$. Assuming that x (*medial wh-*) is observed in the input, it's posterior probability would be $p(h|x)$; however, since the probability of any h that does not contain x is zero, the *medial wh-* representation is eliminated from the target grammar. In other words, after enough experience with the L2, the learners observe that their initial hypothesis about the existence of *medial wh-* in the target grammar is not supported by any tokens of this representation, and thus they bring down its probability to zero. At the same time, since the target LD wh- movement representation is supported by the input (and by their L1), its probability is gradually increased until it becomes the only available option in the grammar.

It is important to point out that there is no grammatical trigger (i.e. direct evidence) responsible for eliminating *medial wh-* constructions from the L2 English of the participants; nor does it seem that there is a trigger that activates the target long-distance wh- movement option, which was attested at a high rate in the interlanguage grammar of even the lowest proficiency group. Interestingly, Yang (2004) claims that triggering proposals (e.g. Gibson and Wexler 1994; Nowak 2001, among others) should be dispensed with because, among other things, they generally presuppose a sudden switch from one grammar to another and lack of long-term co-existence of different grammatical representations. Yang argues that the idea of triggering should be replaced entirely by probabilistic learning, which better reflects the process of acquisition.

While the data presented above is compatible with this line of thought and *medial wh-* is a candidate for a construction unaffected by triggering, it is possible that triggering may play a higher role than probabilistic learning in the case of other acquisition phenomena (for discussion of morphosyntactic triggers in L2 and their differences from L1 triggers see, for example, Vainikka and Young-Scholten 1998, Zobl and Liceras 1994, Liceras and Díaz 2001, among others). Thus, the extent to which Yang's proposal in this respect applies to L2 acquisition remains to be tested by future research.

Another potential modification of Yang's approach in applying it to L2 acquisition would be to include more than just linguistic innateness and probabilistic learning. It could be argued that adding other dimensions may make the overall picture convoluted, and lessen the model's predictive power. At the same time, in applying the approach to L2 acquisition, it is impossible to overlook the multitude and complexity of other important factors at play. For example, L1 transfer always has the potential of playing a prominent role and this idea was already integrated in the discussion of the *choice A and B* representations above (see below for more discussion of L1 transfer). In addition, I assume that general cognitive learning mechanisms other than calculation of statistical probabilities may also play an important part. It is beyond the scope of this dissertation to discuss these in any more detail and they are not directly pertinent to the data of the current study. However, allowing for such mechanisms to be incorporated will likely make the approach more generally applicable to other L2 phenomena.

Note that many acquisition theories assume use of non-linguistic learning mechanisms which lead to competition of different forms and representations. For example, the well-known Competition Model (CM) grounded in connectionism and developed by MacWhinney (1987, 2001, 2007, MacWhinney and Bates 1989, and references therein) stresses the use of universal cognitive strategies and assumes that both L1 and L2 learners engage in complex analysis of the distributional properties of the input, which is crucial in shaping learning and in processing language. Another (related) approach that has been gaining considerable attention in the recent literature is emergentism (see O'Grady 2002, 2005, 2008, and references therein). While the specific emergentist proposals and assumptions vary, proponents of this approach generally view language acquisition in terms of competing basic non-linguistic factors coming from properties of the input, processing, working memory, pragmatics, learning strategies, etc. What these approaches have in common is that they strongly oppose Chomsky's innateness views and the ensuing assumption of domain-specific linguistic constraints operating in the computational system.

The debate between strictly linguistic versus strictly general-cognitive views of language and language acquisition remains largely unresolved, and there seems to be compelling evidence for both. Thus, the general approach to L2 acquisition that I develop in accounting for the L2 data presented above is hybrid in the sense that it aims to benefit from both sides of the debate. On the one hand, the very nature of the phenomenon that I investigate, a construction unavailable in the NL and the TL but attested cross-linguistically, is explained most naturally if the nativist insights provided by UG are preserved. On the other hand, while assuming that UG plays a role in L2 acquisition, it seems difficult to disregard the role of general cognitive processes whose presence in second language acquisition is becoming more and more widely-acknowledged in the literature.[30]

It should be pointed out that in combining elements from competing frameworks, hybrid approaches run the risk of "bring[ing] together strange bed fellows" (Herschensohn 2004, p. 32), and thus they must be carefully balanced. Achieving this

---

[30] Some of the problems that a UG based L2 approach not incorporating domain-general mechanisms might face include: 1) difficulty in explaining variability within and across participants; 2) disappearance of certain L2 phenomena and fossilization of others; 3) task-induced differences in L2 performance.

non-trivial goal, however, has the potential of making such proposals increasingly appropriate in L2 acquisition, where the processes are influenced by a multitude of competing linguistic, cognitive and social factors.[31] While it is often argued that modular and non-modular views of language are essentially irreconcilable, since the inception of the Minimalist Program, Chomsky himself has been increasingly emphasizing the role of non-linguistic factors in the architecture of language, in addition to UG. In *Three Factors in Language Design* (2005a) he accords a prominent role to the genetic endowment (UG), language experience, and "language- or even organism-independent" mechanisms and processes (p. 1). In developing a general approach to second language acquisition and a specific proposal that adequately accounts for the L2 data presented above, this dissertation incorporates the three factors in language design proposed by Chomsky as follows: the genetic endowment is invoked in explaining the existence of otherwise paradoxical *medial wh-* structures in the L2; the language experience is implicated in the inverse correlation between the level of learner proficiency and the rate of occurrence of *medial wh-*; and the non-linguistic (probabilistic learning) mechanisms are instantiated in the ability of the learners to completely eliminate these non-target utterances from their interlanguage (i.e. to bring down their probability in the interlanguage grammar to zero).

Having outlined an approach that accounts for the overall data, in what follows, I will return to some specific results which require further examination.

3.2.4.2. Co-occurrence with complementizer

One of the issues that this dissertation investigates is whether *medial wh-* can co-occur with overt complementizer *that* in the learners' interlanguage. As already mentioned, this is based on Thronton's (1990) observation that *medial wh-* never co-occurred with *that* in children's L1 English utterances. Thus, Thornton proposed that children in fact treat *medial wh-* words as complementizers (see chapter 3 for details). As the data presented above indicated, group F1 did not show sensitivity to the presence or absence of an overt complementizer in *medial wh-* utterances. Group F2-3, on the other hand, showed a small but significant difference in accepting *medial wh-* without an overt

---

[31] For an example and an in-depth discussion of a recent hybrid model grounded in processing but incorporating both minimalist and non-modular (connectionist) views, see Truscott and Sharwood Smith (2004) and the following commentaries.

complementizer more frequently than with a complementizer (15% vs. 10%, respectively). If these L2 learners were misanalysing wh- words as complementizers, as Thornton proposed for the child L1 learners, we would expect this difference to be greater (i.e. the percentage in the *complementizer* condition should be close to zero). Thus, the *medial wh- as a complementizer hypothesis* cannot be pursued successfully for the present data. At the same time, the limited blocking effect that an overt complementizer has on the learners' acceptance of *medial wh-* utterances needs to be accounted for. There are several possibilities in this respect. First, L2 learners could be postulating a doubly-filled complementizer structure in accepting *medial wh-* utterances in *condition 2* (e.g. What do you think **that where** he should go?). This is a plausible scenario because in some varieties of Quebec French utterances as in (29) have been analysed as doubly-filled complementizers (e.g. Rizzi and Shlonsky 2005).[32]

(29) ***Quel*** *garçon* ***que*** *tu as vu?*   *(Quebec French)*
     which   boy   that you have seen
     'Which boy have you seen?'   (adapted from Rizzi and Shlonsky 2005)

It is important to point out, however, that the word order of the doubly-filled complementizer construction differs from the word order in the experimental items of *condition 2*: the wh- word precedes the complementizer in the former, whereas the reverse order is employed in the latter. Thus, I assume that *medial wh-* acceptance with an overt complementizer in *condition 2* of the experimental task is unlikely to be influenced by the speakers' potential use or awareness of doubly-filled complementizers in their native language.

What makes things more complicated is that even though this is not the case in Quebec French, the word order in which a complementizer precedes a wh- word is attested cross-linguistically, as illustrated by Hungarian in (30)-(32)

(30) *Mit mondtak, hogy kit hívott fel Mari?*
     what-acc   said-3ps   that wh-acc called up Mary-nom
     "Who did he say that Mary had called up?   (Horvath 1998)

---

[32] Apart from questions, some varieties of Quebec French also allow doubly-filled complementizers in relative clauses as in (i) below (thanks to María Luisa Rivero for this point).
   (i) *Joséphine voyait la fille avec qui que Louis était sorti.*   (adapted from Rooryck 1988)

(31) *Kiváncsi   vagyok, hogy ki   ment el.*
    curious   I.am   that   who   went PV[33]
    'I wonder who left.'

(32) *János meghívott   egy   lányt, de nem tudom   hogy kit.*
    János invited   a   girl   but not   know-1SG that   who
    'János invited a girl, but I don't know who.'

                    (adapted from van Craenenbroeck and Lipták 2008)

Thus, it is possible that the L2 learners who accept *medial wh-* utterances with an overt complementizer in *condition 2* of the experimental task are postulating a Hungarian-type grammar, an option allowed for by the approach to L2 acquisition developed above (i.e. a UG-constrained representations not necessarily attested in the L1 or L2). Methodologically, it is unclear how one could provide concrete evidence distinguishing between the above two scenarios. However, it is possible to speculate about them by linking them to a general explanation in terms of markedness.

A definition of markedness can be pursued both cross-linguistically, in typological terms, and within a given language. I am adopting the idea that markedness depends on the availability or frequency of distribution of a given property or construction (i.e. the less frequent the more marked and *vice versa*). If we assume that French-speaking learners of L2 English accept *medial wh-* utterances with overt complementizer *that* because the doubly-filled comp filter does not apply in certain varieties of Quebec French, this could be linked to markedness within the L1. That is, utterances with a doubly filled complementizer in Quebec French are marked; thus, if learners transfer this sensitivity to markedness constraints into L2 English, the fact that *medial wh-* with an overt complementizer is accepted to a lower degree than *medial wh-* without an overt complementizer receives an explanation.

On the other hand, under the second scenario where the L2 learners would be postulating a Hungarian-type grammar, the typological approach to markedness would apply. To my knowledge, the Hungarian word order, in which the overt complementizer precedes the *medial wh-* word, has not been reported for other languages that allow *medial wh-* representations. As such, this would be a typologically marked option of a *medial wh-* construction. Of course, the participants in this experiment have no explicit knowledge of typological markedness, however, it is possible to assume that such

---

[33] PV=verbal modifier.

markedness is linked to some yet to be determined universal properties. That is, if markedness, both crosslinguistic and within language, is UG-constrained, then it is plausible that L2 learners exhibit an abstract sensitivity to markedness, without knowing explicitly particular intra- and cross-linguistic properties. This assumption, combined with the general approach developed in this dissertation gives us reason to expect that a Hungarian type *medial wh-* construction with an overt complementizer preceding the intermediate wh- word would be accepted by the L2 learners; crucially, however, due to the marked status of such a construction under UG, the probability learners attach to its existence in the L2 grammar would be lower than the probability of the less marked *medial wh-* construction without an overt complementizer. A more elaborate proposal along these lines is left for future research.

Apart from the question about co-occurrence with complementizer, adopting a typological markedness framework also allows us to account for another *medial wh-* asymmetry obtained in this study. The difference of acceptance in *wh- copying* versus *wh- scope marking* nicely correlates with the cross-linguistic distribution of the two constructions. That is, *wh- copying* occurs less frequently than *wh- scope marking* in the English interlanguage of the French learners, mirroring the distribution of the two constructions in the languages of the world (see chapter 2).

3.2.4.3. The Subset Principle

As mentioned in chapter 2, *medial wh-* constructions are less frequent in the languages of the world and seem to always co-exist with long-distance wh- movement with deletion of all intermediate copies or with wh- *in situ* (again without intermediate copies). This puts *medial wh-* constructions in a superset-subset relation: the superset is the case of languages where both standard LD wh- movement and *medial wh-* or both wh- *in situ* and *medial wh-* can apply; the subset, on the other hand, would be languages where only standard LD wh- movement or only wh- *in situ* (i.e. without *medial wh-*) can apply. According to the Subset Principle (Berwick 1985), learners initially adopt the most restrictive grammar (subset) and only subsequent input may cause restructuring or expansion of the grammar to the less restrictive option (superset). However, the data in *experiment 1* show the opposite of what the Subset Principle would predict: the L2

learners initially embrace the superset grammar where standard LD wh- movement and *medial wh-* are competing representations (with different probability values); subsequently, at the advanced stages of acquisition, the L2 learners restrict their grammar to the subset option (for further claims against the Subset Principle in L2, see White 1989b, Belikova 2008, among others).

3.2.4.4. The negation asymmetry

Recall that the possible effects of negation in the matrix clause on the acceptance rates of *medial wh-* in L2 English were investigated in *condition 3*. As mentioned earlier in this chapter, negation has different intervention effects cross-linguistically. With regards to *medial wh-*, it has been established that negation blocks such constructions in German, as illustrated in the contrast between (33)*a* and *b*; however, the grammaticality of the standard LD movement construction in which the wh- phrase appears only in the matrix clause is unaffected by negation, as in (34). More generally, a negation intervention effect also applies to the possibility of wh- movement in simple questions in French, as illustrated in (35)-(36) (repeated for convenience).

(33) a. **Wen   glaubst du   nicht, wen   sie   liebt?*
       whom   believe you   not   whom   she   loves
       'Who don't you think (that) she loves?'
   b. *Wen   glaubst du   wen   sie   liebt?*
       whom   believe you   whom   she   loves
       'Who don't you think (that) she loves?'

(34) *Wen   glaubst du   nicht,   dass sie   liebt?*
     whom   believe you   not   that she   loves
     'Who don't you think (that) she loves?'

 Wh-*in situ*:

(35) **Il ne   mange   pas   quoi?*
     He NEG eat   NEG what
     'What doesn't he eat?'

 Overt *wh*-movement:

(36) *Qu'est-ce qu'il   ne   mange   pas?*
     What is-it that-he NEG eat   NEG
     'What doesn't he eat?'   (adapted from Oiry and Demirdache 2006)

The *negation* condition in the experimental task contained English sentences roughly equivalent to the example in (33)*a* above (e.g. **What don't you think what he is buying these days?*). Similarly to the results for overt complementizers, the *negation*

condition yielded significantly lower acceptance of *medial wh-* constructions in comparison with the *basic* condition (15% vs. 10%, respectively, for *choice C*) in group F2-3. However, once again, a complete blocking effect was not observed (i.e. presence of matrix negation did not cause the participants to reject *medial wh-* utterances completely). One possibility of accounting for these data is by proposing that negation constitutes a barrier not in the overt syntax but at LF. Such an approach has been adopted with regards to *wh- scope marking* in German by Beck (1996) and Beck and Berman (2000).

The proposal is that in the *wh- scope marking* construction the semantically empty *wh- scope* marker is replaced at LF by the wh- phrase (i.e. the wh- phrase moves to the matrix CP at LF). This is motivated by Chomsky (1986) who argues that expletives must be removed at LF and replaced by a semantically interpreted expression. In the long-distance wh- movement counterpart construction, however, movement of the wh- phrase is overt. Since negation poses a barrier at LF but not in the syntax, the *wh- scope marking* construction is ungrammatical with matrix negation, as in (33)*a*, whereas its LD wh-movement counterpart is not, as in (34). This proposal finds further support from other languages such as Korean, where a negation asymmetry applies between wh *-in situ* and scrambled *wh*-phrases. As Beck and Kim (1997), point out, scrambling across negative polarity items applies in the overt syntax, and thus no negation intervention effects apply; wh- *in situ*, on the other hand, involves LF movement of the wh- phrase and negative polarity items in such cases cause ungrammaticality. Finally, as pointed out by Oiry and Demirdache (2006), the proposal also applies to French where a wh- *in situ* question is licensed when there is no negation but ungrammatical with negation. Once again, assuming LF movement of the wh- phrase in such questions, negation poses a barrier. Oiry and Demirdache (2006) summarize the three cases discussed above as follows:

(37) Negation effects at LF and in overt movement (Oiry and Demirdache 2006)

Negation licit	Negation illicit
Long overt movement in German	Partial *wh*-movement in German
Long overt movement in French	*Wh-in situ* in French
Scrambled *wh-* in Korean	*Wh-in situ* in Korean
→ *Negation does not block overt movement*	→ *Negation blocks covert movement*

Based on these generalizations, the authors adopt the LF approach to negation intervention effects in the *wh- scope marking* (partial movement) constructions in child French L1 acquisition. I also adopt this idea in explaining the data of experiment 1 and apply it to both *wh- scope marking* and *wh- copying* constructions with matrix negation in the interlanguage grammar of the learners of L2 English (see Felser 2004 for discussion of the relationship between negation effects and LF in *wh- copying* in German). Generally, under this approach *medial wh-* utterances with a negated matrix clause would be well-formed in the syntax but not in the LF in the interlanguage of L2 participants who accept *medial wh-* constructions. This contrast between a well-formed syntactic derivation and problematic logical form could be the reason why such *medial wh-* utterances are generally disfavoured, compared to their non-negated counterparts, but at the same time not completely rejected. In other words, since such sentences are not ungrammatical *per se*, the participants still hypothesise that they are possible representations and accept them. At the same time, because these sentences are logically problematic, the learners attach a lower probability value to them than to their non-negated counterparts.[34]

3.2.4.5. Syntactic analysis

The precise syntactic analysis of *wh- copying* and *wh- scope marking* constructions is to this day debatable. A number of proposals vary along several different dimensions, most notably in terms of an *indirect* versus *direct dependency* analysis, and in terms of the question whether the two constructions warrant a unified or separate accounts (see chapter 2 for more details). To account for the *medial wh-* phenomena attested, I adopt a *direct dependency* analysis and assume that *wh- scope marking* and *wh-*

---

[34] It is important to point out that the status of LF movement in current minimalist theory is controversial. On the one hand Chomsky (1995; 2000, and subsequent work; also 1997 MIT Lectures, cited by Bošković 1998) dispenses with LF movement claiming that all movement must take place before Spell-Out. Thus, the distinction between covert and overt syntax disappears and what was previously known as LF movement is reanalysed as pure feature movement (i.e. without movement of the category). On the other hand, Pesetsky (2000) objects to this view and argues that feature movement is not the proper reanalysis of LF movement. Under Pesetsky's proposal, the distinction between covert and overt movement needs to be preserved, while feature movement, a phenomenon in its own right, represents a third type of a movement relation.

With regards to the above data, it seems that either of these accounts could be pursued. On the one hand, it may be possible to suggest that negation is an intervener to pure feature movement. On the other hand, the covert movement analysis could be preserved under Pesetsky's proposal. The precise implementation of either of these two analyses remains beyond the scope of this dissertation.

copying are variants of the same structure in the L2 English of French learners. As mentioned earlier, this assumption capitalizes on the fact that both constructions involve a wh- element in an intermediate position, and that both constructions are ungrammatical in the L1 and the L2 of the population under consideration (provided that such constructions in the L2 do not represent sequential questions).

Hiemstra (1986) was among the first to propose a generative account for *medial wh-*, and also one of the proponents of the view that the *wh- copying* and the *wh- scope marking* constructions can be captured by a unified analysis. Her proposal is based on data from Frisian, which allows three options with regards to complex wh- question formation: standard LD wh- movement with the wh- phrase appearing only in the matrix clause, as in (38); *wh- scope marking*, as in (39); and *wh- copying*, as in (40).

(38) *Wa   tinke jo   dat   ik sjoen haw*
    who think you   that   I   seen have
    'Who do you think that I have seen'

(39) *Wat   tinke jo   wa't   ik sjoen haw*
    what think you who that-cl I   seen have
    'Who do you think that I have seen?'

(40) *Wa   tinke jo   wa't   ik sjoen haw*
    who think you who that-cl I   seen have
    'Who do you think that I have seen?'

To generate these three options, Hiemstra assumes what would be an equivalent to a *direct dependency* approach[35] in which different elements of the wh- phrase (e.g. wh- feature, *phi* features, phonetic matrix) may move independently of one another. By this token, the *wh- scope marking* (partial movement) construction in (39) is generated when the wh- phrase moves from its basic position to the CP of the embedded clause; from that position, the wh- feature separates and raises to the specifier of CP in the matrix clause where it is realized phonetically as the unmarked/default wh- word *wat* "what". The *wh- copying* construction in (40) is also generated by successive cyclic application of wh- movement: when the wh- feature moves to the specifier of the matrix CP, the *phi* features are pied-piped and also move to that position but a copy of the phonetic matrix is left behind (i.e. pronounced in an intermediate position). Finally, LD wh- movement without

---

[35] The terms *indirect* and *direct dependency* emerge in the subsequent literature.

a *medial wh-* word is generated by moving the wh- feature, the *phi* features and the phonetic matrix to the specifier of the matrix CP.

While Hiemstra's (1986) account was conceived within the Government and Binding framework, a more recent analysis in terms of wh- feature separation was developed independently[36] by Cheng (2000) for German. Cheng's proposal is framed in minimalist terms (Chomsky 1995) and focuses on the *wh- scope marking* construction only, but can in principle be extended to *wh- copying* as well (see, however, Felser 2004 for a contrary argument). The account is based on Chomsky's (1995) idea that wh-movement is a two-step process involving feature movement and category movement. While feature movement results from a feature checking operation, category movement (also known as generalized pied-piping) is motivated by reasons of PF convergence. An additional assumption is that category movement is always followed by a repair strategy that prevents the features from being scattered (i.e. the category and the feature bundle stay together). After the repair strategy takes place, subsequent operations are based on the output of that repair. Cheng's (2000) proposal regarding *wh- scope marking* is that first the wh- feature moves to the lower CP. This is followed by movement of the category in order for the repair to take place. Subsequently, the feature bundle undergoes another movement to the matrix CP checking the [+wh] feature of C^0 and appearing at spell out as the default PF form *was* "what"; the real wh-word, however, is left behind and pronounced in the embedded CP. A derivation of the example in (41) is illustrated in (42).

(41) Was$_i$ glaubt [$_{IP}$ Hans [$_{CP}$ wen$_i$ [$_{IP}$ Jakob t$_i$ anruft]]] ?
    WH   thinks   Hans   whom Jakob   is calling
    "Whom does Hans think that Jakob is calling?"
(42) [$_{CP}$ [FF] glaubt [$_{IP}$ Hans [$_{CP}$ wen [$_{C^0}$ t [FF]] [$_{IP}$ Jakob t$_{wen}$ anruft]]]]
                                    copy   Cheng (2000)

As Cheng points out, the movement of the feature bundle leaves a copy behind, similarly to movement of categories, and thus the wh- phrase that appears in intermediate position is not left without its feature bundle. The way Cheng derives the difference between *wh-scope marking* and full LD movement is as follows: in the former case, the repair strategy takes place in the embedded CP and the category is spelled out in it; in the latter case, the

---

[36] See Cheng (2000) p. 81, footnote 2.

repair strategy takes place in the matrix CP and thus the real wh- phrase (i.e. not the just a phonological realization of a wh- feature) is spelled out there.

In accounting for *medial wh-* utterances in L2 English both Gutierrez (2005) and Schulz (2006) assume a feature separability analysis. The former adopts Hiemstra's (1986) account applying it to the *wh- copying* and wh- scoping constructions produced in L2 English by Spanish/Basque bilinguals. The latter author, on the other hand, adopts Cheng's (2000) account focusing on *wh- scope marking* structures in the L2 English of Japanese and German speakers. In this dissertation I will also adopt feature separability as an explanation for both the *wh- scope marking* and the *wh- copying* constructions attested in the English interlanguage of (Canadian) French speakers. The account is conceptually appealing and allows for a unified treatment of the two phenomena. It is important to point out that although most contemporary analyses treat *wh- scope marking* and *wh- copying* separately, a unified account in L2 is more desirable since the L2 data collected so far have not shown that the two constructions have different properties warranting two separate analyses. A simplified derivation of a *wh- copying* and a *wh- scope marking* utterance is provided in (43) and (44), respectively.

(43) Who do you think who John kissed?
[$_{CP}$ **who** [$_C$ do [$_{TP}$ you [$_v$ think [$_{CP}$ **who** [$_{TP}$ John [$_{vP}$ {John} [$_v$ kissed [$_{VP}$ {kissed} [$_{DP}$ {who}]]]]]]]]]]]

(44) What do you think who John kissed?
[$_{CP}$ **what** [$_C$ do [$_{TP}$ you [$_v$ think [$_{CP}$ **who** [$_{TP}$ John [$_{vP}$ {John} [$_v$ kissed [$_{VP}$ {kissed} [$_{DP}$ {who}]]]]]]]]]]]

In (43), the wh- phrase *who* is merged as an object of the embedded verb *kiss* and undergoes successive cyclical movement to the specifier of the matrix CP where spell out takes place.[37] A copy of the phonetic matrix, including the relevant features is also spelled out in the specifier of the embedded clause. Similarly, in (44), the wh-phrase is merged as an object of the embedded verb and subsequently undergoes movement to the specifier of the embedded CP. The phonetic matrix of the wh- phrase is spelled out in this position, including the relevant *phi* features. The crucial difference is that the wh- feature of the phrase leaves a copy in that position and separates; the wh- feature alone continues

---

[37] Wh- movement to the left periphery is assumed to be caused by an uninterpretable interrogative feature in *C*.

movement to the specifier of the matrix CP where it receives a phonetic realization as a default *what*.

As discussed earlier, utterances including both a complementizer and a *medial wh-* word were accepted as grammatical by some participants in the French speaking population, as illustrated in (45).

(45) **Who** do you think **that who** John kissed?

Such utterances were also attested by the Spanish/Basque speakers of L2 English in Gutierrez's (2005) study. Thus, the idea that a *medial wh-* word might in fact be used as a complementizer in the L2 learners interlanguage was rejected (cf. Thornton 1990 for L1 English). Gutierrez (2005) offers additional arguments against such an analysis using evidence from *D-linked*[38] *medial wh*-phrases attested in some of her participants' productions, as in (46).

(46) **Who** do you think **what person** live in the house?   (participant 79)
                                                                        (Gutierrez 2005)

As Gutierrez points out, the *D-linked* wh-phrase must occupy the SpecCP position in the embedded clause, as it is a maximal projection. Thus, the wh- word cannot be a complementizer.

Having proposed that the *medial wh-* word appears in a specifier position of the embedded CP, the question arises about the precise position of the complementizer *that* which precedes the *medial wh-* word, as in (45). One analysis that seems to give a natural explanation to this question is proposed by Rizzi (1997, and subsequent work) whereby the left periphery is made up of multiple layers. The general proposal is that the complementizer phrase consists of at least a Force Phrase and a Finiteness Phrase (other phrases such as topic and focus have also been identified). Following Gutierrez (2005), I adopt this general idea for the *medial wh-* utterances involving complementizers. A (simplified) sample derivation is provided in (47).

(47) [CP **who** [C do [TP you [vP [v think [VP ...[FORCEP that [CP **who** [TP John [vP {John} [v kissed [VP {kissed} {who}]]]]]]]]]]]

---

[38] The term *D(iscourse)-linked* has been used since Pesetsky (1987) to refer to complex wh- phrases of the type Wh- N (e.g. which person, what boy, etc.); such wh- phrases relate to sets of referents previously established in the discourse and known to both the speaker and the hearer.

As this derivation indicates, the complementizer *that* is located in the ForceP of the left periphery of the embedded clause which is structurally above the *medial wh-* word. Recall that Hungarian also has *medial wh-* structures in which both a wh- word and a complementizer appear in intermediate position. The word order in Hungarian is also such that the complementizer precedes the wh- word, as in (30) above, and Horvath (2000) suggests an account in terms of split CP layers as well.[39]

3.2.4.6. Arguments versus adjuncts

Recall that an argument-adjunct asymmetry was attested in group F2-3, and to some extent in group F1. In essence, in group F2-3 wh- arguments were significantly less likely to appear in a medial position than wh- adjuncts (*choice C* answers were attested as follows: 10% arg. vs. 20% adj. for *condition 1*; 6% arg. vs. 13% adj. for *condition 2*; and 5% arg vs. 16% for *condition 3*). These results are particularly interesting considering that argument-adjunct asymmetries have been attested in a number of different constructions cross-linguistically. One of the earliest proposals with regards to such asymmetries and wh- movement was made by Huang (1982) who observed that adjuncts cannot be extracted from wh- islands, whereas wh- arguments (objects) cause only minor deviance, as illustrated in (48) and (49), respectively.

(48) *How do you wonder which problem to solve?
(49) ?Which problem do you wonder how to solve?

<div align="right">(Huang 1982; adapted from Rizzi 1990)</div>

In perhaps one of the most thorough investigations of such issues in the literature, Rizzi (1990) proposes a GB account relating the different behaviour of arguments and adjuncts to both referential indices and the empty category principle. In essence, the proposal is that arguments are inherently referential, whereas adjuncts are not. This idea is captured in the principle formulated in (50) and is consistent with Chomsky's (1965) original proposal of referential indices.

(50) A referential index must be licensed by a referential *theta* role.   (Rizzi 1990)

---

[39] The precise location of the complementizer and the moved wh- phrase under the split CP layer approach may vary. The implementation of this issue in the most current theory is not pursued in this dissertation. The key point is that under this approach, it is possible for both a complementizer and a wh- word to appear in the same CP, and for the former to precede the latter.

This principle ensures that only positions associated with a referential *theta* role in a given representation can receive a referential index. By this token, the index *k* is licensed in (51) because it is associated with a *theta*-marked position. On the other hand, no referential indexing is allowed in (52) because no referential *theta* role is involved.

(51) Who$_k$ did you see t_k?
(52) How did you behave t?
(53) X binds Y iff
    (i) X c-commands Y
    (ii) X and Y have the same referential index.   (Rizzi 1990)

To summarize, the principle illustrated in (50) yields the representations in (51) and (52). Assuming a definition of binding as in (53), Rizzi points out that an A′ dependency is legitimately formed through the index in the case of the argument; however, another mechanism must be used to establish this relationship in the case of the adjunct. He proposes this mechanism to be government and formulates a simplified version of the ECP, as illustrated in (54).

(54) **ECP:** A nonpronominal empty category must be properly head-governed.

The crucial difference between the two mechanisms is that binding can apply at distance whereas government is strictly local. The examples in (51) and (52) are both grammatical because in the former the wh- trace is properly bound and in the latter the trace is locally head-governed. This is not the case, however, with the LD examples in (48) and (49) above. The former is ill-formed because the trace of the adjunct *how* cannot bear a referential index and cannot be bound long-distance (note that the trace of the object *which problem* is properly governed because it is in the local domain of its operator). The latter example, on the other hand, is grammatical because the adjunct *how* has not moved long-distance and its trace is properly governed (i.e. locally); the object *which problem* has moved long distance, but nonetheless its trace can be properly licensed because it bears a referential index and can establish a long-distance binding relationship.[40]

What are the implications of the above analysis with regards to the *medial wh-* structures observed in the L2 English of (Canadian) French speakers? Keeping in mind that argument-adjunct asymmetries have been attested in a number of different contexts cross-linguistically, and adopting Rizzi's proposal that adjuncts are more sensitive to

---

[40] The slight deviance of (49) is attributed to a subjacency violation.

locality, the data from *experiment 1* receive an explanation. The L2 learners have a sensitivity to the distinction between arguments and adjuncts and follow the generalization that it is more difficult to establish an LD relationship with adjuncts. This idea can be implemented in different ways. Assuming the GB framework, the L2 learners could be applying the binding mechanism to license the LD connection between arguments and their traces (or unpronounced copies). Thus, they are able to establish the long-distance representation more easily with arguments than with adjuncts and a wh-argument is less likely to be spelled out in an intermediate position. With adjuncts, on the other hand, referential indices and binding are unavailable, and thus a chain of government relations must be established. Since a chain must be able to hold locally, a wh- adjunct is spelled out in an intermediate position as an indication of locality. Without specifically addressing *medial wh-* constructions, Rizzi (1990) provides the following summary, which supports the L2 scenario proposed above.

> There are two (nonexclusive) ways to connect an operator and its variable: binding and a chain of government relations. Binding requires identity of referential indices, a formal property now restricted by principle [(50)]. When coindexation and binding are not available, the chain of government relations is the only connecting device. But government relations are intrinsically local. Hence, if a link of the government fails, the connection between operator and variable cannot be established, and the structure is ruled out. The chain as such can cover an unbounded distance, but each link is local; this is the fundamental difference between the two connecting devices. (p. 92)

Essentially, Rizzi suggests that both LD connection mechanisms can be at play at the same time: arguments (objects) are compatible with both, whereas adjuncts can only rely on government, which is intrinsically local. Thus, it is plausible that the L2 learners are more likely to accept *medial wh-* structures where the intermediate wh- element is an adjunct than when it is an argument because the former are more intrinsically local than the latter. As such, *medial wh-* structures in the interlanguage of the L2 learners seem to be an overt expression of their sensitivity to locality constraints.

It is important to point out that almost all learners who accepted *medial wh-* constructions also accepted target LD wh- structures. Thus, as mentioned earlier, both representations exist in the interlanguage grammar. The question that arises is why there should be a need for an overt expression of locality (i.e. pronouncing a *medial wh-* word) only part of time and no need of such overt indication of locality the rest of the time. This question receives an answer in the general proposal about the nature and process of L2

acquisition spelled out earlier. First, recall that in natural languages *medial wh-*constructions co-exist with LD wh- structures where no intermediate copies are pronounced; as such, there is no reason why similar co-existence should not be attested in early interlanguage (i.e. before the learners realize that only one of the two options is licensed by the particular TL). Second, consistent with probabilistic learning and markedness, the LD wh- structures without an intermediate wh- element are more likely to be accepted due to their presence in the L1 and the L2 input, and due to their less marked status cross-linguistically.

As mentioned earlier, the account proposed for the argument-adjunct asymmetry attested in the L2 experimental data is grounded in the GB framework. However, it is also possible to recast it in minimalist terms.[41] According to current minimalist theory (e.g. Newmeyer 2003; Chomsky 2005b) the wh-phrase has an interpretable interrogative feature [iWh] and the head *C* in the left periphery has an uninterpretable wh- feature [uWh]. The latter may or may not have an EPP property responsible for attracting the wh-phrase to the edge of *C*.[42] In case the EPP property is not available, checking the uninterpretable feature on *C* is achieved through (long-distance) Agree (Chomsky 2001) yielding wh- *in situ*, as in French. Alternatively, wh- *in situ* may be achieved through a choice function (Reinhart 1998) or an abstract *Q-feature* operator in CP (Cheng and Rooryck 2000; Mathieu 2004; Hamann 2006). Since in the relevant L2 data long distance wh- movement is assumed (i.e. the whole phrase moves to the matrix CP in the case of *wh- copying*, and only a wh- feature moves to the matrix CP leaving the wh- word in medial position behind in the case of *wh- scope marking*), I adopt the idea that the EPP property is present in *C*. Thus, the wh- phrase (or at least part of it) is attracted to the left periphery of the matrix CP. The idea of locality in minimalism is expressed by the notion of phase (Chomsky 2001; 2005b). Thus, if we again assume Rizzi's idea that two different "connection" devices can apply to the movement of arguments (referentiality, which can hold at distance, and feature-driven wh- movement, which goes through phases), and only one device can apply to the movement of adjuncts (referentiality is

---

[41] The assumptions adopted here are to some extent similar to the ones used by Jakubowicz and Strik (2008) in their account of *medial wh-* structures in child L1 acquisition of French and Dutch.
[42] As Jakubowicz and Strik (2008) point out, Chomsky (2005b) attributes the effect of the EPP to an edge feature EF automatically available for a given lexical item.

unavailable but strictly local phase-driven movement is available), the need to mark locality by spelling out an intermediate copy of wh- adjuncts before proceeding to a new phase can apply once again.[43] A more precise implementation of this idea is left open to future research. The key point, however, is that it seems possible to account for the L2 argument-adjunct asymmetry under the current minimalist model as well as under the GB framework.

Overall, the fact that the L2 learners seem to have such a refined sensitivity to the different movement properties of arguments and adjuncts shows that their interlanguage is constrained at a very deep level by abstract grammatical principles. Since this is a defining characteristic of all natural languages, there is evidence that interlanguage development should be treated on a par with natural language development.

3.2.4.7. The learnability problem

Recall that one of the main findings in the data from the French learners of L2 English was that the *medial wh-* dependency attested is to a very large extent of the *direct* type (exclusively so in group F2-3). Since such constructions are not instantiated in either the L1 or the L2 of the participants, I adopted an approach in which access to UG plays a prominent role. That is, the essential learnability problem posed by *direct dependency medial wh-* constructions in the English interlanguage of French speakers is resolved by proposing that UG is implicated in the L2 acquisition process (along with other, non-linguistic learning mechanisms). However, there are several additional factors that could potentially trivialize the learnability problem. One of them is the possibility of L1 transfer. Even though it was already pointed out that the *direct dependency medial wh-* representation is ungrammatical in the L1, there is one possibility of L1 transfer that needs to be addressed. As noted in section 3.2 (examples (19)-(20)), the French complementizer *qui/que* 'that' are homophonous with the wh- words who *qui* 'who' and

---

[43] Note that Rizzi (2001) eliminates the formal device of the referential index proposed in Rizzi (1990) and discussed above. Nonetheless, he claims it is possible to preserve the essential idea of referentiality by employing the notion of *D-linking* (Pesetsky 1987), and uses reconstruction theory to explain the relevant data. In this dissertation I do not use Rizzi's (2001) approach because it does not seem to apply to the L2 data discussed here as naturally. More specifically, the notion of *D-linking* does not seem pertinent as the wh- phrases used in the experiments reported on here were not *D-linked*. Furthermore, in a recent paper by Friedmann, Belletti and Rizzi (in press) it is suggested that the account in terms of *D-linking* may not be fully satisfactory.

que 'what'. This makes an explanation in terms of L1 transfer, as illustrated in (55), quite plausible. Essentially, the argument would be that due to the homophony in French, the learners misanalyse the English wh- word "what" as a complementizer.

(55) **Que**   penses-tu **que**   j'ai acheté ?
      what   think-you **that**   I have bought

COMP TO MED WH- TRANSFER

"***What** do you think **what** I bought"
Intended meaning: "What do you think that I bought?"

Arguing that the *medial wh-* word results from transfer of a French complementizer is consistent with Thornton's (1990) account of *medial wh-* in child English: recall that Thornton claimed that children produce *medial wh-* utterances in the L1 acquisition of English because they incorrectly categorize the wh- word as a complementizer. However, the results presented above show evidence against the transfer explanation. First, as already discussed, *medial wh-* utterances in *condition 2*, where an overt complementizer was included, were not fully rejected; thus, we would have to postulate that the learners accept a doubly-filled comp representation, which would be problematic because the word order of the experimental items in *condition 2* (comp. followed by wh-) contradicts the word order (wh- followed by comp.) of the doubly-filled complementizers attested in Quebec French (see section 3.2.4.2 above). Second, the argument versus adjunct asymmetry observed in the results of group F2-3, and to some extent group F1, provides even stronger evidence against the L1 transfer scenario. As mentioned earlier, the wh- words for adjunct extractions were *where* and *when*; crucially, these words are not homophonous with complementizers in French and the fact that such adjunct wh- words were attested at a higher rate in *medial wh-* positions than their argument counterparts (see figure 8) makes the L1 transfer hypothesis impossible to pursue.

Apart from L1 transfer, another potential explanation for the *medial wh-* constructions that does not necessarily involve UG can be sought in terms of L2 processing. As mentioned in section 3.2.3.4, it is possible that the participants have the need to shorten the long-distance wh- dependency in complex questions and thus pronounce intermediate copies of the moved wh- words. This would help alleviate the processing burden, which, especially during earlier stages of acquisition, can be very high

(see Schulz 2006 for an overview of processing considerations in L2). Such an explanation seems quite plausible in light of the fact that *medial wh-* is a phenomenon that occurs during the early stages of acquisition and tends to disappear as acquisition progresses (i.e. when the processing load presumably decreases). Thus, it is possible to argue that in order to alleviate the high processing burden, less advanced learners posit *medial wh-*. However, there are several factors, both in terms of the nature of the experimental technique and in terms of the results obtained, that point against an explanation based entirely on processing. First, the design of this experiment is more likely to reflect the learners' grammatical competence rather than their processing profile. That is, an off-line grammaticality judgment task is considered to be a measure of existing grammatical representations rather than independent processing constraints. Note also that the participants had ample time to provide an answer to each item[44], which presumably alleviated the processing burden.

A second source of evidence against a processing analysis comes from the fact that the percentage of *indirect medial wh- dependency* responses attested was at a negligibly low level (3%) in group F2-3. If those learners accepted *medial wh-* utterances as grammatical because the *medial wh-* element helped them shorten the long-distance dependency, then we would expect them to have accepted the *indirect dependency* (*choice D*) much more often. In other words, the most logical and perhaps easiest way of shortening a long-distance dependency is to posit sequential questions which contain two short, local wh- dependencies; yet the participants in group F2-3 almost never chose this option.

Finally, recall that independent support against a processing account comes from Schulz (2006) who investigates *wh- scope marking* using an elicitation experiment in conjunction with an off-line and an on-line grammaticality judgment tasks. Her study was specifically designed to address the question whether *wh- scope marking* in the L2 English of German L1 and Japanese L1 speakers is due to a processing effect and the results from the on-line task suggest that the phenomenon is not related to a processing deficit (see chapter 2 for details).

---

[44] Even though the task was not self-paced (i.e. each item appeared on a screen for a limited time), the time allotted to each item was quite generous, and in fact some participants reported that they had to wait too long before the next item would appear.

Rejecting the L1 transfer and the processing explanations brings us back to the proposal that L2 acquisition is UG constrained. Recall that the *direct dependency medial wh-* constructions are ungrammatical in the L1 as well as the L2; they are not taught, nor are they part of the target input in any other way; they are attested in a number of natural languages, of which the participants report no knowledge; they appear in the L2 English of participants with different L1 backgrounds such as Japanese, Spanish/Basque and French; and finally, they seem to be a phenomenon that the learners overcome in the later stages of acquisition. Taken together, all of these factors are best explained by the idea that L2 learners have access to Universal Grammar. As such, the account developed in section 3.2.4.1 provides a suitable explanation for the findings of *experiment 1*.

Before proceeding with *experiment 2*, it is important to reiterate that rejecting the L1 transfer and the processing scenarios considered above does not mean that a UG-based account can exist in isolation of these factors. In the case of L1 transfer, while such effects do not seem to be possible with regards to the *medial wh-* phenomena observed in the L2 English of French speakers, they were discussed as a likely factor in the learners' acceptance of the target LD construction. In the case of processing, the above discussion provides evidence that its effects alone cannot explain the *medial wh-* structures attested; however, it is still possible that processing effects in combination with access to UG may be responsible for the L2 phenomenon under investigation. Further support for this idea will be offered in the conclusion of this chapter and in chapter 4, where I discuss the spoken production experiments. In what follows immediately, however, I focus on the second written experiment involving Bulgarian L1 learners of L2 English.

3.3   Experiment 2: L1 Bulgarian; L2 English.

The second experiment was conducted with a much smaller population. The purpose was to determine if the results obtained with the French speakers of L2 English can be replicated with speakers of a typologically different L1, Bulgarian. Since the theory developed in section 3.2.4 assumes that UG is implicated in L2 acquisition and that the construction under investigation is not directly affected by L1 transfer, the prediction was that the results of Bulgarian speakers of L2 English would be similar. The research questions for *experiment 2* were the same as for *experiment 1*.

Before providing details on the participants, procedures and results of the experiment, in what follows I show examples of Bulgarian complex questions. Unlike French, where in-situ questions are possible, Bulgarian makes exclusive use of the LD movement mechanism. This is illustrated in object and adjunct extractions in (56) and (57), respectively. An interesting fact about Bulgarian is that it has flexible word order, as illustrated by the *a* versus *b* versions of the examples. Furthermore, as pointed out in the previous chapter, Bulgarian also allows multiple wh- fronting, as in (58). Finally, note that Bulgarian is a pro-drop language, as illustrated by all the examples below.

(56) a. *Koj misliš   (če) Ivan e celunal?*
  who think.2sg   (that) Ivan has kissed
  'Who do you think (that) Ivan has kissed?'

  b. *Koj misliš   (če) e   celunal Ivan?*
  who think.2sg   (that) has kissed   Ivan
  'Who do you think (that) Ivan has kissed?'

(57) a. *Kâde   misliš   (če) Ivan e otišâl?*
  where think.2sg   (that) Ivan has gone
  'Where do you think Ivan went?'

  b. *Kâde   misliš   (če) e   otišâl Ivan?*
  where think.2sg   (that) has gone Ivan
  'Where do you think Ivan went?'

(58) a. *Koj, kâde, koga   misliš   (če) Ivan e   celunal?*
  who where when   think.2sg   (that) Ivan has   kissed
  'Who do you think (that) Ivan has kissed where and when?'

  b. *Koj, kâde, koga   misliš   (če) e   celunal Ivan?*
  who where when   think.2sg   (that) has kissed   Ivan
  'Who do you think (that) Ivan has kissed where and when?'

A comparison of these Bulgarian data with the French examples given earlier in (11)-(15) would show a number of differences in the properties of complex question formation between the two languages. However, these are not discussed in any detail. The key point with regards to the L2 experiments of this chapter is that both languages can use long-distance wh- movement with deletion of intermediate wh- copies, which is an option consistent with the target language, English; at the same time, both Bulgarian and French do not license *medial wh-* representations.

3.3.1. Participants

The total participants in *experiment 2* were 31 students recruited from two different high-schools in Sofia, Bulgaria. They were matched for age with the participants from *experiment 1* (between 15 and 18 years old). As illustrated in table 9, the participants were divided into two proficiency groups: high beginner and advanced.[45]

Table 9. Number of Participants by proficiency level (experiment 2)

Group	Proficiency	Number
BG1	High beginner	17
BG2	Advanced	14
Total		**31**

As with the French participants, administering a separate proficiency test specifically designed for the study was not possible due to logistical considerations. Thus, the proficiency levels used are based on the levels provided by the schools. However, as with the French speakers, the Bulgarian participants' scores on the experimental task also clearly correlate with the proficiency level in which they were placed by their school.

One important difference between the French and the Bulgarian speakers is that the former were ESL learners (i.e. learning English as a *second* language in a country where the dominant language is English) and the latter were EFL learners (i.e. learning English as a *foreign* language in a country where the dominant language is not English). As such, the Bulgarian participants were exposed to English in their daily experiences to a much lower degree than the French-Canadians. It is also safe to hypothesise that the Bulgarian learners learned English in a more structured and less naturalistic setting, whereas many of the French speakers, due to higher degree of contact with English native speakers and opportunities to practice their L2, acquired English in a much more naturalistic way.

Despite these differences between the French and the Bulgarian speakers, comparing their results is possible due to the very nature of the phenomenon which this dissertation investigates. That is, *medial wh-* constructions are unavailable in both L1s and in the target language, and the key goal is to determine if such constructions exist in the English interlanguage of the two populations; thus, the fact that the learning

---

[45] Due to logistical and time constraints, it was not possible to recruit a (low)-intermediate group.

environments were different and that the levels of proficiency could not be matched exactly does not constitute a serious obstacle.

3.3.2. Design and procedures

The Bulgarian participants were given the same language background questionnaire and experimental task as the French speakers. All participants were also subject to equivalent selection criteria as the ones set in *experiment 1*. The two Bulgarian groups were given the long form of the experimental task, (i.e. the version given to group F2-3). The presentation of the task, however, differed from the experiment with the French speakers due to lack of appropriate audio and visual equipment at the schools where the experiment with the Bulgarians took place. Thus, the Bulgarian speakers received the test in printed form with the relevant context story and picture immediately preceding each item (appendix 2). They were instructed to mark their answers on a separate answer sheet (appendix 3). The researcher had the advantage of speaking the participants' native language and was able to give occasional explanations or translations of unknown words to individual participants. The learners were encouraged to ask questions in either English or Bulgarian in case of lack of comprehension, and to avoid random guessing on the test. They were also instructed to follow their intuitions and not to go back and review their answers.

3.3.3. Results

Experiment 2 was successful in attesting *medial wh-* constructions in the English interlanguage of Bulgarian speakers. Similarly to the French data, the rate of occurrence was higher at the lower level of proficiency and decreased dramatically with the advanced group. Due to the lower number of participants in each group, it is difficult to provide statistically significant answers to some of the five research questions. Nonetheless, many of the patterns observed with the French groups were confirmed with the Bulgarian participants. Thus, the data as a whole forms a very strong body of evidence.

3.3.3.1 Group BG1 (high beginner)

Although it was not possible to administer a separate test to match the proficiency of the French participants against the Bulgarian participants, the level of group BG1 roughly corresponds to that of group F1. As illustrated in figure 21, BG1's overall accuracy rate is 63% which is equivalent to F1's 60% shown earlier in figure 12 (the difference is not significant).

Figure 21. Complex questions: Total responses by choice A, B, C & D
Group BG1, n=17
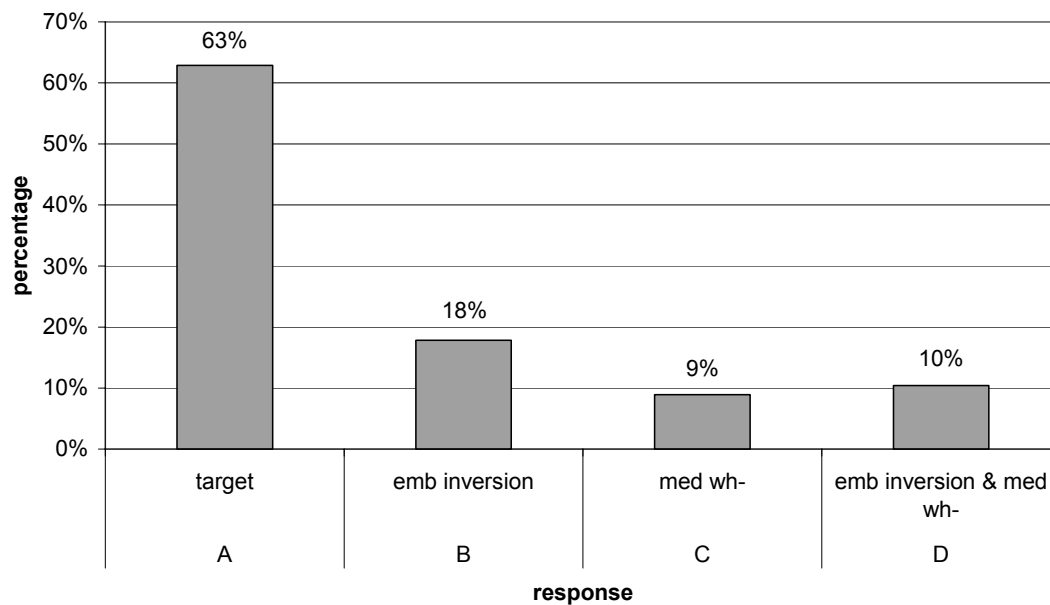
These data indicate that the Bulgarian participants at this level seem to accept both the *choice C* and *choice D medial wh-* representations, as was the case with group F1 (F1's rates on these choices were 16% and 10%, respectively).

With regards to **R2**, the *wh- copy* versus *wh- scope marking* asymmetry observed with the French population is replicated with this Bulgarian group, as shown in figure 22.

**Figure 22. Wh-copy vs. wh-scope marking
Group BG1, n=17**
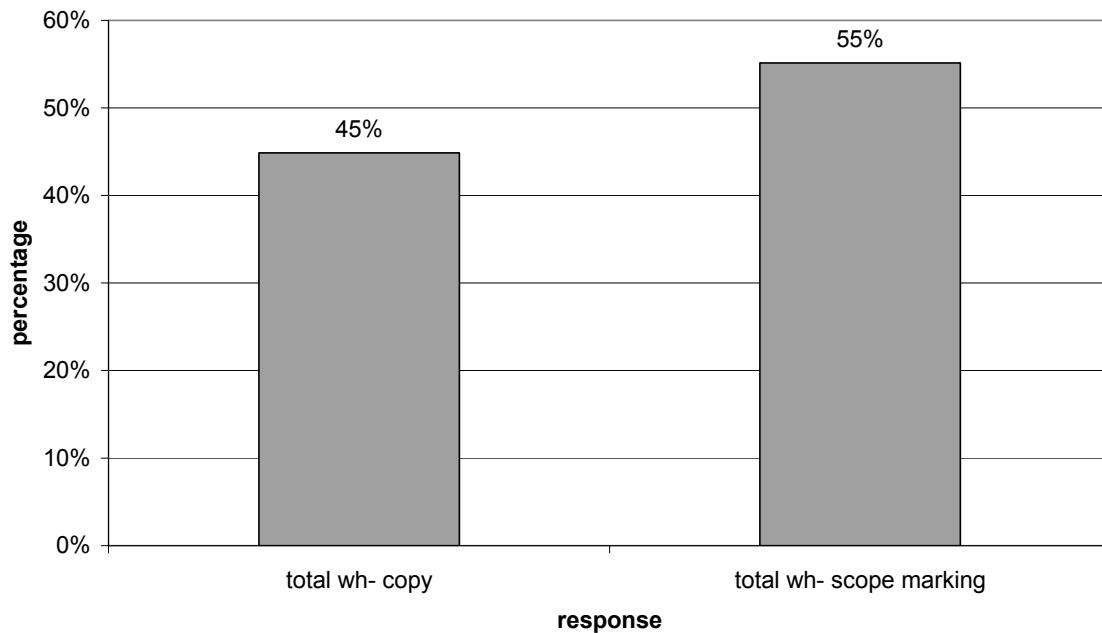
A logistic regression analysis showed that this difference is not significant; however, considering that the pattern has at this point been attested in several of the French groups, it should be viewed as uncontroversial.[46] Also note that the total *wh- copy* and *wh- scope marking* acceptance rates in figure 22 are calculated using the responses on both choices *C* and *D* of the test. If we extract only the *choice C* answers (i.e. if the *choice D* answers are excluded) the asymmetry becomes larger (33% *copy* vs. 67% *scope*) and reaches significance ($p< 0.03$).

In answering **R3**, the question regarding the nature of the dependency (*direct* versus *indirect*), the control procedure implemented with the French group F2-3 was repeated in order to identify the participants who have reliably acquired T-to-C movement in simple questions. That is, all participants who had achieved a score of 6 out of 8 (75%) or higher on the simple control questions were retained and the rest were excluded. This is illustrated in figure 23.

---

[46] That is, if the data with regards to *wh- copying* versus *wh- scope marking* from all relevant groups is taken together, the overall pattern would be significant.

Figure 23. Simple questions: Target responses (T-to-C movement)
Group BG1, n=17
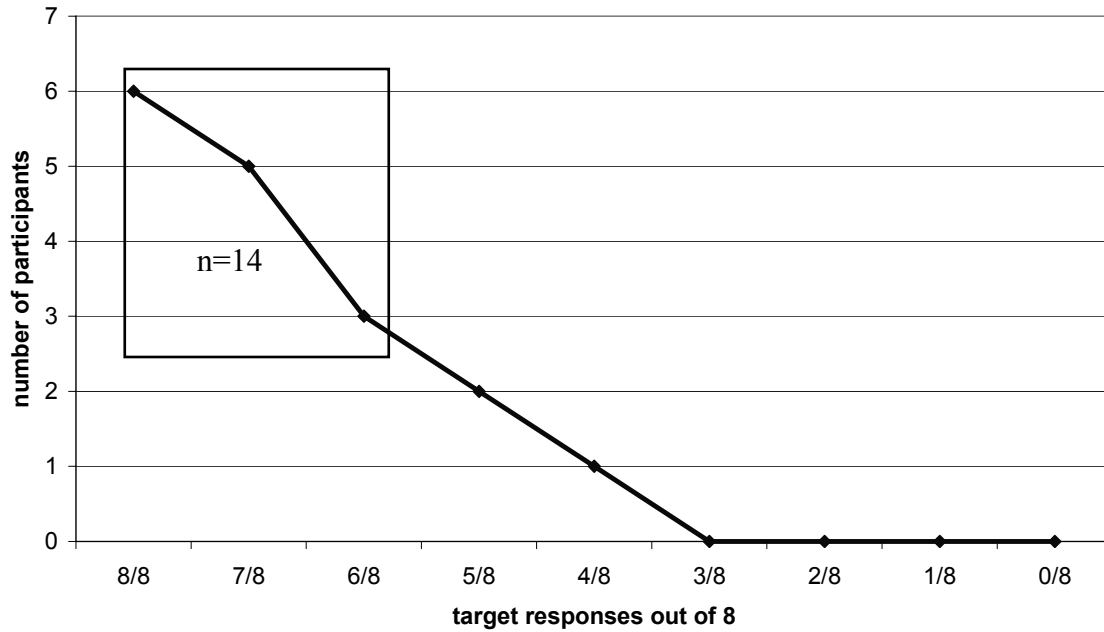
Due to the high scores achieved on the control questions, only 3 participants were excluded by this procedure and the remaining 14 were retained. The analysis of complex questions was re-run obtaining the results shown in figure 24.

Figure 24. Complex questions: total responses by choice A, B, C, & D.
Group BG1, n=14 (participants who have acquired T-to-C movement)
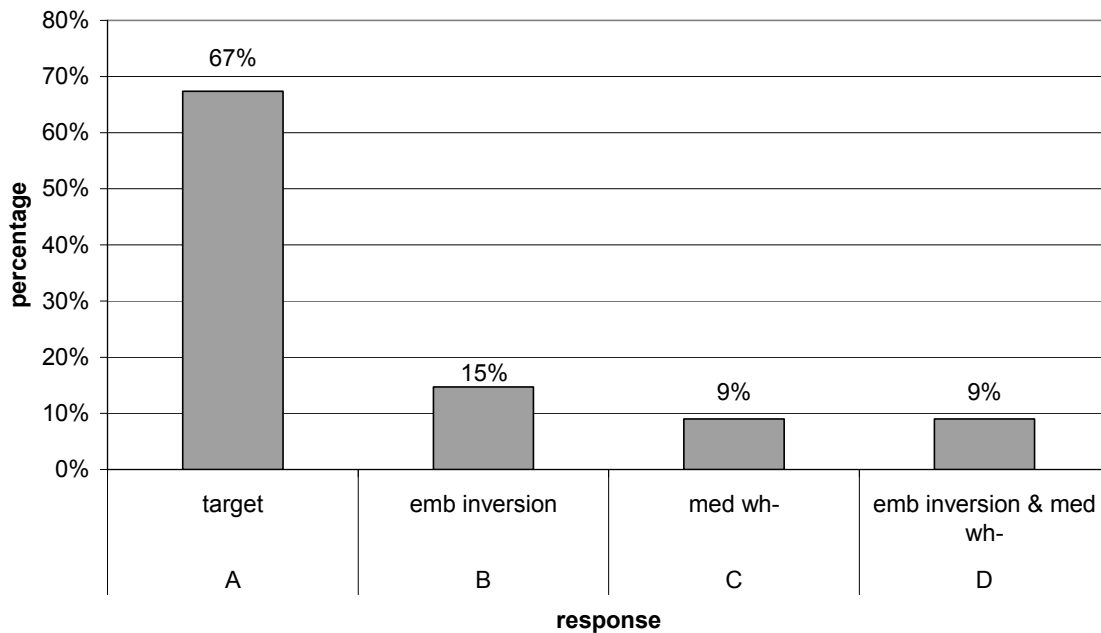
As these data indicate, there is no difference between the *direct* (*choice C*) and the *indirect* (*choice D*) *dependency* representations. As with group F1, one way of explaining this would be by proposing that the participants overgeneralize the English T-to-C movement rule and apply it to embedded clauses as well. Alternatively, it could be argued that both the *direct* and the *indirect medial wh-* dependencies co-exist at an equal level in the interlanguage grammar. Since most of the participants in this group achieved a high score on the simple questions, the T-to-C movement overgeneralization hypothesis seems very plausible. In order to pursue it, it is necessary to look at the individual results and compare the scores of matrix inversion with the scores of embedded inversion. These results are given in figure 25.

**Figure 25. Individual inversion rates in simple questions and
embedded clauses, Group BG1, n=17**
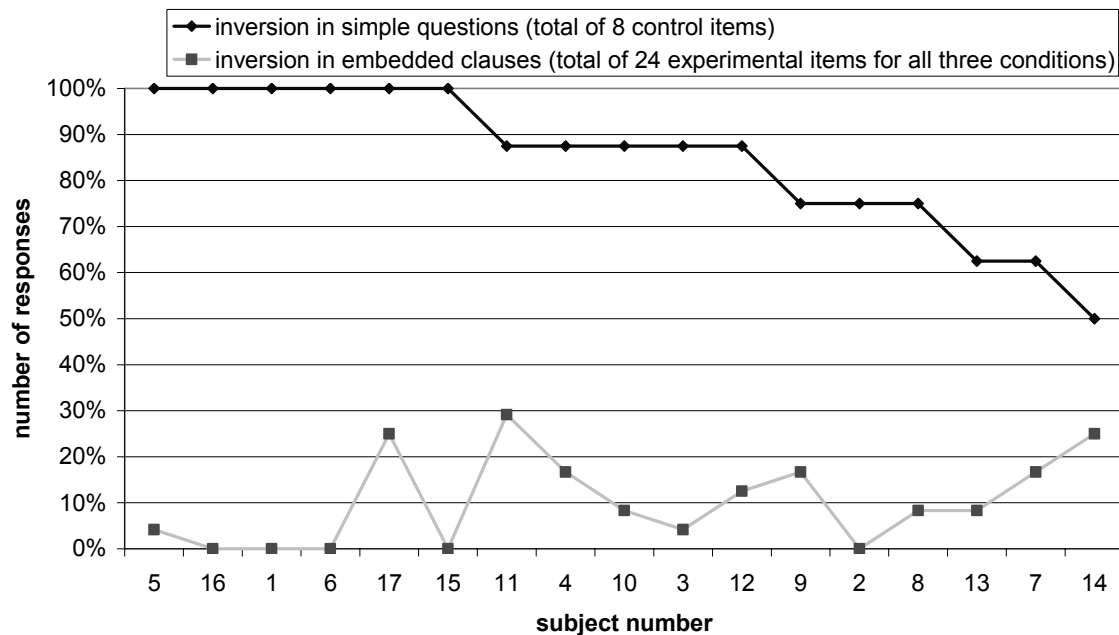
These data indicate that the overgeneralization hypothesis does not receive sufficient support. Almost all participants show inversion rates lower than 20% in embedded clauses. Only participants 17, 11 and 14 show embedded inversion between 20% and 30%, which is still relatively low. At the same time 14 out of the 17 participants invert 75% or more in simple questions, and the remaining three participants (13, 7, and 14) invert 50% or more, which is still a high score. Thus, the embedded inversion individual rates are decidedly lower than the inversion rates in simple questions, which makes it difficult to argue that the participants (except perhaps 17 and 11 to some extent) overgeneralize the inversion rule. Consequently, the explanation whereby the *direct* and *indirect dependency* are equally represented in the participants' L2 grammar at this level of proficiency seems most appropriate. Recall, that the same findings apply to the French speaking group F1, which had roughly equivalent proficiency in the L2.

Turning to **R4**, the question regarding co-occurrence of *medial wh-* and an overt complementizer, the data in which the different experimental conditions are juxtaposed is presented in figure 26.

**Fig. 26. Complex questions: Total responses for choices A, B, C & D
by condition, Group BG1, n=17**
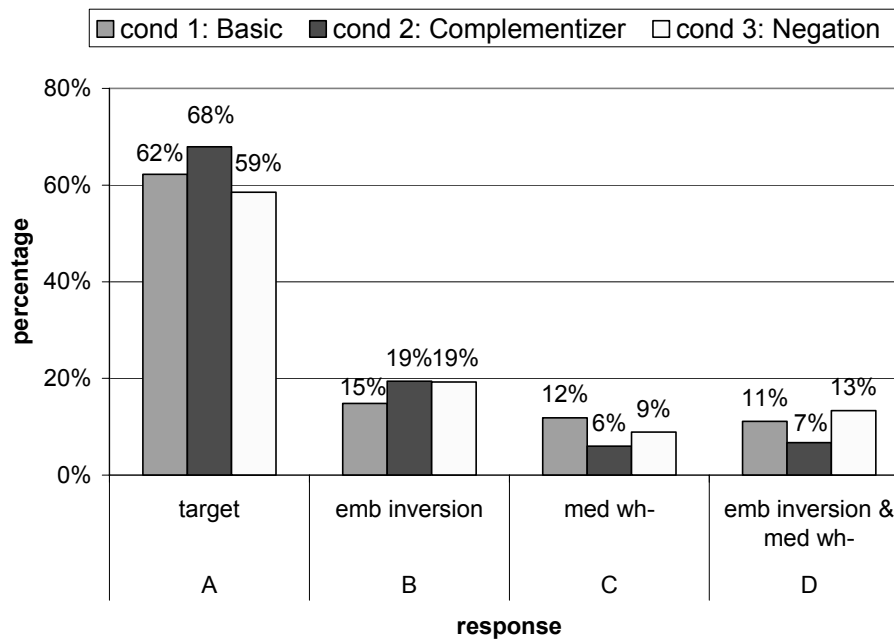
As with the French high beginner and low-intermediate groups, group BG1 shows a lower rate of *medial wh-* acceptance (*choice C*) in the *complementizer* condition (6%) compared with the *basic* condition (12%). This difference is marginally significant (p<0.095). *Medial wh-* and embedded inversion (*choice D*) is also tolerated to a lower extent in *condition 2* (7%) than in *condition 1* (11%), although this difference is not significant. Despite the lack of strong significance, the trend with regards to this research question mirrors the trend showed by the French groups, and thus receives the same answer: *medial wh-* is less likely to occur with an overt complementizer, but at the same time, complementizers do not have a full blocking effect. Turning to **R5**, the question regarding a possible negation effect, the data from group BG1 again pattern together with the equivalent French data (i.e. lower acceptance rate but no complete blocking), but do not reach significance.

The argument versus adjunct asymmetry observed with some of the French groups, receives some support from group BG1, as figure 27 indicates.

Figure 27. Complex questions: Arguments vs. adjuncts
Group BG1, n=17
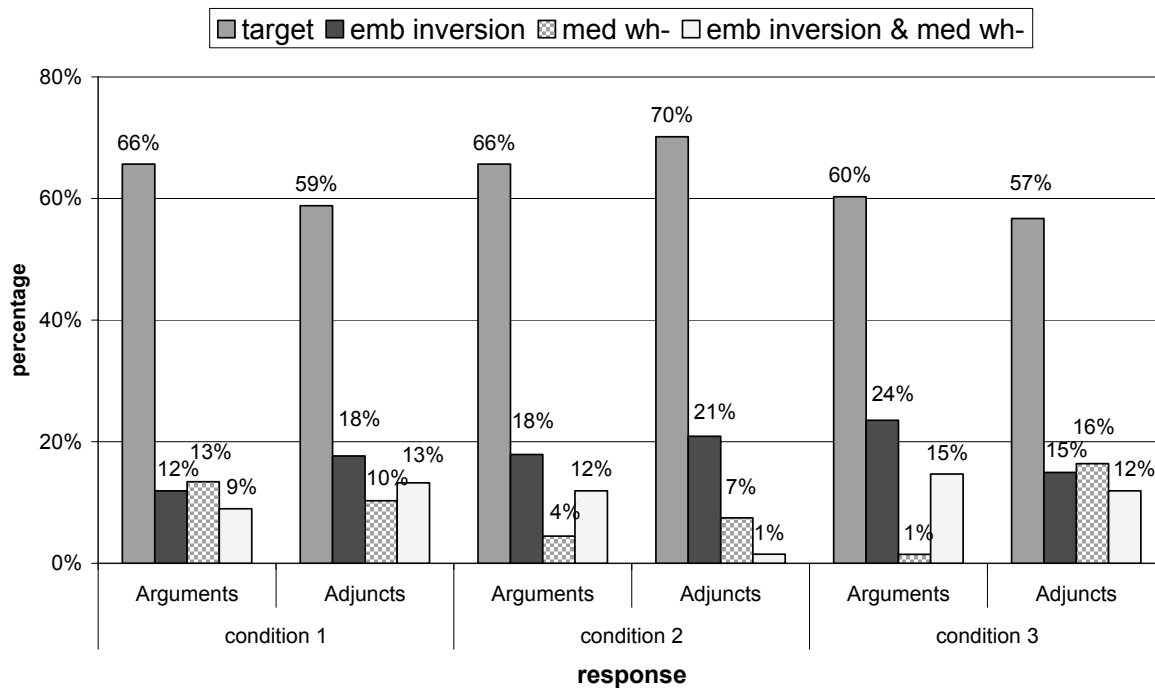
A logistic regression analysis showed that the difference between the 1% arguments and 16% adjuncts in *choice C* responses is significant (p<0.001) for *condition 3*. That is, *medial wh-* acceptance is higher in the case of adjunct wh- words than in the case of argument wh- words in this condition. The equivalent differences for *choice C* in conditions 1 and 2, however, are not significant.

Some individual results on matrix and embedded inversion rates were already presented above. In what follows, more individual data is given, in order to gain a more thorough understanding of the behaviour of group BG1. Table 10 shows the individual rates for *medial wh-* with and without embedded inversion.

Table 10. Group BG1: Individual medial wh- acceptance rates

Participant	CHOICE C [+]MED. WH-, [-]EMB. INV.		CHOICE D [+]MED. WH-, [+]EMB. INV.		SUM OF CHOICE C AND CHOICE D	
	number of responses out of 24	%	number of responses out of 24	%	number of responses out of 24	%
17	6	25%	6	25%	12	50%
11	4	17%	7	29%	11	46%
14	3	13%	6	25%	9	38%
9	4	17%	4	17%	8	33%
12	4	17%	3	13%	7	29%
8	3	13%	2	8%	5	21%
4	1	4%	4	17%	5	21%
7	1	4%	4	17%	5	21%
10	2	8%	2	8%	4	17%
13	2	8%	2	8%	4	17%
15	2	8%	0	0%	2	8%
16	2	8%	0	0%	2	8%
3	1	4%	1	4%	2	8%
5	1	4%	1	4%	2	8%
1	0	0%	0	0%	0	0%
2	0	0%	0	0%	0	0%
6	0	0%	0	0%	0	0%
mean	**2**	**9%**	**2**	**10%**	**5**	**19%**

While the average *medial wh-* acceptance rate (*choice C + D*) is about 19%, the responses are not distributed evenly among the participants. Participants 17, 11, 14, 9 and 12 have a total *medial wh-* acceptance rate of 29%-50% (choices *C + D*). Participants 15, 16, 3, 5, 1, 2 and 6, on the other hand, have a very low rate of acceptance of such constructions (0%-8%). Thus, as with the French participants, *medial wh-* in the Bulgarian population is not a pervasive phenomenon applying across the board. Rather, it is a competing representation only in some individual grammars, and even then it is highly outranked by the target construction.

As for the nature of the dependency, the group results already pointed out to the co-existence of the *direct* and *indirect* representations in the interlanguage grammar. The question arises whether these representations are equally distributed in the individual responses or whether certain participants favour exclusively one and others exclusively the other. Although the differences in rates within each individual are not very high, and therefore it is possible that some of them are accidental, the general pattern that emerges at first sight is for an *indirect dependency* preference. This can be seen in participants 11,

14, 9, 4, and 7. Note that participants 14 and 7 generally show similar rates of inversion in matrix and embedded clauses (they are placed furthest to the right in figure 25 above). This indicates that they do not seem to differentiate between matrix and embedded clauses very well with regards to T-to-C movement. Thus, their preference for *choice D* over *choice C* on the experimental task may reflect their ambivalence with regard to inversion, as opposed to a real preference for the *indirect dependency*. Participants 11, 14 and 9, on the other hand, have more disparate rates on inversion in simple questions versus embedded clauses, and thus show stronger evidence of a real *indirect dependency* preference in the L2 grammar. Finally, note that participant 17 shows an equivalent rate of *choice C* and *choice D*. Thus, it appears that the *direct* and the *indirect* dependencies have equal status in this participant's interlanguage grammar. However, it is important to point out that this participant, as well as the other participants who were shown to prefer the *indirect dependency* have some of the highest rates on embedded inversion (figure 25) in the entire group. Thus, even though the T-to-C overgeneralization hypothesis was rejected, it is possible that some of these participants choose the *indirect medial wh-dependency* representation on the test due to certain ambivalence between inversion in matrix and embedded contexts and not necessarily because their L2 grammars favour sequential questions.

On the whole, the results of group BG1 largely mirror those of group F1 (for summary, refer to section 3.3.4 below).

3.3.3.2 Group BG2 (advanced)

While matching the proficiency of the second Bulgarian group against the French speakers by using a separate test was again not possible, this group's level was estimated to be equivalent to group F6. Figure 28 shows the results for group BG2 by condition.

**Fig. 28. Complex questions: Total responses for choices A, B, C & D
by condition, Group BG2, n=14**
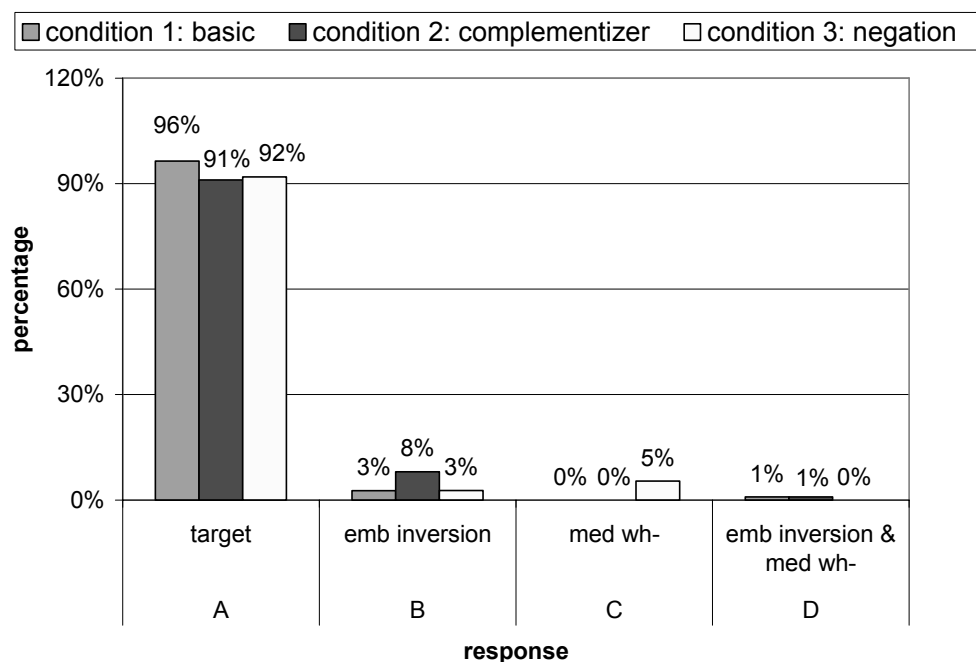
As with the advanced French group, *medial wh-* is virtually non-existent in the responses of group BG2. This re-enforces the previously established inverse correlation between the level of proficiency and the occurrence of *medial wh-* representations. In essence, this group proves that with enough experience and L2 input *medial wh-* disappears from the English interlanguage grammar of Bulgarian speakers. The overall results of this group closely approximate those of the native controls (i.e. no significant differences with regards to choices *C* and *D*[47]), as was the case with group F6.

3.3.4. Discussion

The purpose of the second experiment was to confirm that the findings of *experiment 1* are not limited to a population with a particular L1 background. Thus, showing that these findings apply to additional participants whose L1 is Bulgarian provides general support for the theory developed and proposals advanced in

---

[47] Although the logistic regression output comparing groups BG2 and N(ative) showed that the difference in *choice D* (1% vs. 0%, respectively, for conditions 1 and 2) is significant, this result is not meaningful. That is, the mathematical significance value is based on the difference between two tokens of this type of response for the entire group BG2 and zero tokens for group N, and as such cannot be considered valid. This is also confirmed by the logistic regression estimates table included in appendix 5.

section 3.2.4. Although the number of Bulgarian participants was smaller and the level of significance was often not as high as with some of the French groups, the fact that the overall patterns were very similar in both populations is a strong indication that *medial wh-* phenomena in L2 English apply similarly to learners with a variety of typologically different L1 backgrounds. That is, *medial wh-* was attested in both populations at similar rates in the similar proficiency levels; both the *indirect* and the *direct* dependencies were attested in the lowest proficiency groups; the learners showed sensitivity to an overt complementizer and to negation in the matrix clause; and a trend to accept *medial wh-* with adjuncts more often than with arguments was observed. Just as with *experiment 1*, that data from *experiment 2* also show that *medial wh-* is eliminated at the advanced stages of acquisition.

A clear disadvantage in the second experiment was that the two Bulgarian groups were placed at both ends of the spectrum and thus no intermediate proficiency level corresponding to the French group F2-3 was available. One of the most important findings with those French participants was that they strongly preferred the *direct medial wh- dependency*, while the *indirect* one was virtually not attested. The lower-level groups, BG1 and F1, on the other hand, showed evidence for both dependencies. The prediction with regards to a potential Bulgarian group at the intermediate level of proficiency is that the *indirect dependency* would be eliminated and an exclusive preference for the *direct* one would be attested, similarly to group F2-3. I leave this prediction to be tested by future research while acknowledging that before it is borne out, the juxtaposition of the two populations is somewhat incomplete. Nonetheless, the data presented in the two experiments constitute a very strong corpus of evidence pointing to a parallel in the behaviour of English L2 learners with typologically distinct L1 backgrounds. This finding supports the general L2 acquisition approach developed earlier in this chapter and shows that it can be pursued with other L2 learners of English who have typologically distinct L1s.

Since the results from experiments *1* and *2* show a strong parallel, the discussion of competing representations, co-occurrence of *medial wh-* with an overt complementizer, negation effects, syntactic status, arguments versus adjuncts, and learnability issues offered in section 3.2.4 applies in the same way to both speaker populations and will not

be repeated here. In what follows, I offer a summary and final remarks on the two written experiments.

3.4   Conclusion

The two experiments reported on in this chapter offer a significant contribution to the understanding of *medial wh-* phenomena in L2 English. An overwhelming preference for the *direct dependency* representation was attested with the (low)-intermediate level French speaking participants (group F2-3). The data from the less advanced French and Bulgarian speaking participants (groups F1 and BG1) also indicated that this particular *medial wh-* representation exists in their interlanguage. Such constructions are paradoxical in both learner populations because there is no obvious reason for their existence. To explain this learnability problem, an approach to second language acquisition whereby learners make hypotheses within the space provided to them by UG was developed; at the same time, the use of non-linguistic mechanisms such as probabilistic learning was also incorporated. The proposal is open to further modifications and while allotting an important role to the genetic language endowment, also allows for the integration of additional general cognitive learning mechanisms, insights from incremental processing considerations, and effects of explicit instruction. It can be argued that such mixed and open approaches have low predictive power; however, I believe that such proposals have the potential of looking at the overall picture in second language acquisition and of accounting for the multitude of diverse and complex factors at play. The challenge faced by approaches that combine views from conflicting frameworks is to specify to what extent each view plays a role in explaining a particular phenomenon. With regards to the phenomenon discussed in this dissertation, the proposal is quite simple and specific: UG is implicated in the existence of *direct dependency medial wh-* representations in the early interlanguage grammar; probabilistic learning is implicated in indirect negative input analysis (i.e. lack of *medial wh-* representations in the target grammar), which makes the more experienced learners abandon the non-target constructions.

Having recapitulated the overall views on L2 acquisition that this dissertation pursues, below is a summary of the specific findings of the two experiments. Table 11 shows the results on the attested non-target representations.

Table 11. Written experiments: summary of main findings (non-target representations)

Group	Total med. wh-	Wh-copy vs. scope	Direct dep. (subset of learners)	Indirect dep. (subset of learners)	Co-occurrence with comp. (choice *C*, basic vs. complementizer condition)	Co-occurrence with matrix neg. (choice *C*, basic vs. negation condition)	Arg. vs. adj.
F1 high beg.	26%	44% vs 56%**	8%	6%	[]	[]	√**
F2&3 (low) interm.	15%	32% vs 68%*	11%	3%	15% vs 10%*	15% vs 10%*	√*
F4 high interm.	6%	–	–	–	–	–	–
F5 low adv.	4.2%	–	–	–	–	–	–
F6 adv.	4.8%	–	–	–	–	–	–
BG1 high beg.	19%	45% vs 55%**	9%	9%	12% vs 6%**	12% vs 9%**	√*
BG2 adv.	3%	–	–	–	–	–	–

* significant
** close to significant or supported by a strong trend attested across groups
[] not significant
√ med wh- occurs with adjunct wh- words more often than with argument wh- words in one or more of the three conditions.[48]
– question not pursued or no data collected in the particular test version [49]

Table 12 shows a summary of the target representations attested in the two experiments.

---

[48] As shown in the results sections of the two experiments, in a few cases *medial wh-* arguments were actually accepted slightly more often than *medial wh-* adjuncts. However, these trends were never significant. As such, they are considered noise.

[49] The percentages listed in the *direct* and the *indirect dependency* columns do not add up to the percentage listed in the total *medial wh-* column. This is due to the fact that the percentages for the *direct* vs. the *indirect dependency* were calculated based on a subset of participants for each group (i.e. only those who had reliably acquired T-to-C movement).

Table 12. Written experiments: summary of main findings (target representations)

Group	Overall accuracy on complex questions	Overall accuracy on simple questions
F1 *high beginners*	60%	59%
F2-3 *low intermediate and intermediate*	74%	70%
F4 *high intermediate*	86%	81%
F5 *low advanced*	87%	91%
F6 *advanced*	95%	96%
BG1 *high beginners*	63%	85%
BG2 *advanced*	93%	99%

As these two tables indicate, the *medial wh-* error rates are higher with the lower levels of proficiency, and decrease drastically with the more advanced groups. At the same time, the level of accuracy on both complex and simple questions increases as the *medial wh-* representations decrease. The *wh- scope marking* construction generally outranks *wh-copying* in the groups where the *medial wh-* rate is high enough to warrant investigation of this issue. As for the *direct* versus the *indirect dependency*, it seems that there are three stages: in stage one, where the proficiency is lowest, both dependencies are attested; in stage two, at the (low)-intermediate level of proficiency, only the *direct dependency* is still attested; and in stage three, at the highest level of proficiency, *medial wh-* rates become virtually non-existent. The sensitivity to complementizers and negation in the matrix clause seems to be a phenomenon that occurs mostly in stage two (i.e. where the *medial wh-* errors are most systematic and the *direct dependency* is prevalent). The same applies to the argument-adjunct asymmetry.

It is important to point out that the learners' sensitivity to complementizers, negation, and arguments versus adjuncts should not necessarily be interpreted as independent evidence for access to UG in L2 learning. As mentioned earlier, the main evidence for this claim comes from the existence of the *direct dependency medial wh-* representation in the interlanguage grammar (i.e. a representation unattested in either the

L1 or the L2, but consistent with the UG hypothesis space). While the sensitivity to an overt complementizer, negation, and argument-adjunct distinctions shows that the learners' interlanguage is systematically constrained, it is possible that its source is L1 transfer at least in some cases. For example, the negation effects attested with regards to *medial wh-* in the L2 English of the French participants may be the result of abstract L1 transfer since French has a negation asymmetry in wh- *in situ* questions (recall examples in section 3.2.4.4). Furthermore, some argument-adjunct asymmetries have also been observed in French (see Rizzi 1990, among others), and presumably the learners' sensitivity to *medial wh-* arguments versus adjuncts in L2 English could also be due to an abstract L1 transfer effect in this population. Finally, it is also possible that the L2 results are to some extent influenced by semantic-pragmatic effects (i.e. different properties of arguments and adjuncts in this respect). Exploring such proposals in more detail is beyond the scope of this dissertation. The key point is that the results from both the French and the Bulgarian populations presented in this chapter show systematic patterns with regards to abstract linguistic properties attested in different natural languages of the world. As such, the data indicate that learner interlanguages are constrained in a similar way to natural languages. Furthermore, the results fit well within the general approach developed in this chapter whereby UG is implicated in L2 acquisition.

Another question that needs to be addressed is why the lower-level learners are in general equally inclined to accept both the *indirect* and the *direct medial wh- dependency* representations, whereas the learners closer to the intermediate level select exclusively the *direct* one. This finding is somewhat surprising considering that the *indirect dependency* representation $[_{CP} wh-_i... t_i] [_{CP} wh-_j... t_j]$ corresponds to sequential questions, which are available in both the L1 and the L2 of the two populations. As such, there is no obvious reason why this representation should be rejected by the intermediate, and even the advanced learners. One plausible explanation is that at the lower proficiency levels, the learners have more limited processing resources, which causes them to choose the *indirect dependency* response at a higher rate. This is based on the assumption that the long-distance wh- movement of a complex question is more difficult to process than the two independent local movements of sequential questions. This explanation will receive further support from the oral experiments discussed in next chapter, where some of the

participants resort almost exclusively to the *indirect dependency* in their spoken productions (i.e. the processing burden is higher in spoken production than in written grammaticality judgement, and thus the *indirect dependency* is attested to a higher degree).[50]

A second factor that may also play a role is explicit instruction and learning strategies. Simple questions are explicitly taught and frequently drilled in ESL/EFL classrooms, and thus it may be easier for lower proficiency learners to use a sequence of two simple questions than one complex question involving long-distance movement.

As a third factor, information structure and economy of the derivation may also play an important role. Jakubowicz and Strik (2008) point out that some authors have claimed that adults prefer dense information packaging, whereas children prefer information packaging that is closer to logical form (van Kampen 1997, van Kampen and Evers 1995). Other authors have made proposals along similar lines claiming that children prefer more economical derivations, that is, options with less/shorter movement (see Hulk and Zuckerman 2000, Zuckerman 2001, Hamann 2006, Rizzi 2000, Jakubowicz 2002, 2003, Jakubowicz and Nash 2001, and Soares 2003, 2006, among others). If we apply these ideas to L2 acquisition and assume that lower proficiency learners favour less dense information packaging or more economical derivations, then the fact that the lower proficiency groups F1 and BG1 were more inclined to choose sequential questions receives a natural explanation. That is, due to more limited language experience and higher processing load, the learners at the lower level of proficiency often opt for the less dense and more economical representation (*choice D*). Conversely, the higher proficiency groups have essentially abandoned these representations (despite the fact that they are available in both the L1 and the L2) because with increased experience and processing capacity the learners are inclined to choose the option with the more densely-packaged information (i.e. the representation preferred by native speakers).

This concludes the discussion of the written experiments on *medial wh-* structures in the L2 English of French and Bulgarian speakers. The next chapter will add a new perspective on some of the same issues by reporting on two oral production experiments.

---

[50] Recall that earlier in this chapter I argued that a processing explanation alone cannot account for the data, but that processing in combination with the theory and assumptions outlined in section 3.2.4 has an important role.

Chapter IV
Oral Experiments

- SUMMARY: THIS CHAPTER RERPORTS ON A SPOKEN STUDY OF COMPLEX WH- QUESTIONS WITH (CANADIAN) FRENCH AND BULGARIAN SPEAKING L2 LEARNERS OF ENGLISH; THE STUDY RAN PARALLEL TO THE WRITTEN EXPERIMENTS DESCRIBED IN THE PREVIOUS CHAPTER AND WAS DESIGNED TO ADD A NEW PERSPECTIVE TO THE FINDINGS THEY OFFERED.

4.1 Introduction and research questions

This chapter investigates the development of complex wh- questions in the L2 English of French and Bulgarian learners from the perspective of oral production. Under the view adopted in the previous chapter, whereby L2 acquisition is guided by both UG and non-linguistic learning mechanisms, *medial wh-* constructions should be expected in both written acceptability and in oral production experiments. At the same time, since *medial wh-* utterances are ungrammatical in both the two L1s and the L2 under consideration, their rate of occurrence in oral production may be more limited. That is, in the written task described in the previous chapter, *medial wh-* constructions (with and without embedded inversion) were explicitly included as an option in a multiple-choice test, and the learners were essentially asked if such representations were compatible with their interlanguage grammar. In an oral production experiment, where the variety of possible constructions that can potentially be attested is considerably higher, the question arises as to whether the participants would necessarily show evidence of this specific type of representation. As such, the focus of this chapter is broader: while *medial wh-* constructions remain a central issue, other productions related to the acquisition of complex wh- questions will also be examined (see **R4** below).

Due to this broader focus and the different data collection method in this chapter, the list of research questions from the written experiments was modified. The first three research questions were retained and are repeated below.

R1. Do *medial wh-* representations exist in the interlanguage of French and/or Bulgarian learners of L2 English?

R2. If *medial wh-* is attested, is it of the *wh- copying* or the *wh- scope marking* kind?

R3. If *medial wh-* is attested, does it constitute a *direct* or an *indirect dependency*?

It is important to acknowledge that **R3** cannot be addressed as reliably as with the written experiments because the oral task did not incorporate a strict control measure of the participants' knowledge of inversion in embedded and matrix contexts (for more details, see description of oral experimental task below).

The research questions regarding co-occurrence with complementizer *that* and negation in the matrix clause which were included in the previous chapter were not retained for the spoken experiments because it was deemed unlikely that such structures would occur in elicited production (i.e. it is very difficult to target such structures in an elicitation task). However, another, more general question was added in order to reflect the fact that a spoken task was likely to generate a wider variety of target and non-target constructions.

> **R4.** What other target and non-target constructions related to the acquisition of complex wh- questions exist in the productions of the two learner populations?

This research question anticipates the possibility of different individual LD wh-movement avoidance strategies yielding a variety of target and non-target constructions in spoken production. That is, less advanced learners may find it difficult to produce long-distance wh- movement in spoken production, as the processing burden is expected to be higher in speaking than in written comprehension or grammaticality judgment. Therefore, it was hypothesised that even if the learners already have the LD wh-movement mechanism in place, some of them may avoid using it by resorting to simple wh- questions, yes/no questions, sequential questions, or other strategies of shortening the dependency or simplifying the derivation.[1]

The prediction that such long-distance wh- movement avoidance strategies would be attested fits within the Derivational Complexity Hypothesis (Jakubowicz 2004, 2005, Jakubowicz and Strik 2008). Essentially, this hypothesis claims that more complex computational and representational structures are more difficult to acquire and/or process. One obvious question that arises is how to define and measure derivational complexity. An important concern in this respect is over-prediction. That is, if every single

---

[1] It should be pointed out that in the context of this dissertation the term *avoidance strategy* is not meant to imply intentional or conscious avoidance of LD wh- movement but rather spontaneous or unintentional use of an alternative structure, which is less complex from a derivational and processing point of view. For more on the nature of the avoidance strategies see the results and discussion sections below.

computational process is assigned some value, over-prediction is bound to occur. Jakubowicz (2005, quoted in Strik and Jakubowicz 2008) proposes a simple *derivational complexity metric* based on *Internal Merge* (IM), as illustrated in (1).

(1) Derivational Complexity Metric (DCM)

      A.   Merging α_i *n* times gives rise to a less complex derivation than merging α_i (n + 1) times.

      B.   Internal Merge of α gives rise to a less complex derivation than Internal Merge of $\alpha+\beta$.

                              (adapted from Jakubowicz and Strik 2008)

Essentially, with regards to wh- movement, the above metric predicts that short movement is less complex than long-distance movement. That is, the higher the number of times a wh- element must be moved in order to satisfy some computational requirement, the more complex the derivation (clause A).[2] Furthermore, moving one constituent is less complex than moving more than one constituent (clause B).

Adopting the DCH and its associated metric (DCM), and focusing strictly on wh-movement, structures such as simple questions and biclausal structures with only local or no wh- movement (e.g. embedded questions, yes/no questions, etc.) are predicted to occur as less complex alternatives to LD wh- movement. In addition, *indirect dependency medial wh-* structures (i.e. with embedded inversion) can be analysed as sequential questions with two local wh- movements, and as such they also constitute an LD avoidance strategy. *Direct dependency medial wh-* (i.e. without embedded inversion) can also be considered an avoidance strategy, although, it has a special status. *Wh- copying* involves LD movement and thus adhering to the DCM does not predict that it is less complex than the target LD structure with deletion of all intermediate copies. At the same time, by adding a further assumption that the intermediate copy fills a gap and as such alleviates the processing load (i.e. working memory) of the LD dependency, the *copying* structure can be considered less complex. As for *wh- scope marking*, if we adopt Cheng's (2000) account where a wh- feature separates from the wh- word in intermediate position and moves to the matrix CP, long-distance movement of the wh- feature is essentially assumed. In terms of derivational complexity, the question arises as to whether feature movement is less complex than moving the whole phrase. I will speculate that this is the

---

[2] Successive cyclicity is assumed.

case, and assume that *wh- scope marking* is less complex than *wh- copying*.[3] Note that the different constructions that are derivationally less complex than LD wh- movement are not ordered in a strict hierarchy. Thus the DCH is used to the extent that it provides a principled way of distinguishing between LD wh- movement and its alternatives (see section 4.2.5 for further discussion).

To summarise, avoidance strategies are expected to arise due to the difficulty associated with standard LD wh- movement with deletion of all intermediate copies. These strategies are defined in terms of short wh- movement, lack of wh- movement altogether, and filling intermediate gaps to ease processing. The predictions and theoretical assumptions associated with **R4** will be brought up again and evaluated in the results and discussion sections of this chapter.

Before I proceed with the description of the experiments, it should be noted that the approach adopted here is somewhat different from previous L2 studies on complex wh- questions. As already mentioned, Schulz (2006) uses a similar elicited production task as in the current study; however, in reporting the results she only lists the number of complex wh- questions and specifies how many of them contained *wh- scope marking*.[4] In other words, she does not report on any other productions that might have been attested in the task. Gutierrez (2005), also using a similar elicitation task, reports on a number of different LD constructions that her participants produced, but she still does not include the entire range of structures (e.g. simple questions) that were produced. The approach adopted in the present study aims to examine all productions that might occur in a context eliciting LD wh- movement and determine if and how their variation may be constrained. To put it differently, the idea is to examine what alternatives LD wh- movement might have, and whether these alternatives can provide additional insights into the L2 acquisition of the relevant target structure.

---

[3] As mentioned in the previous chapters, on some accounts *wh- scope marking* involves movement of the wh- phrase to the left periphery of the embedded clause and subsequent expletive insertion in the matrix CP. Under such views, *wh- scope marking* is clearly a less complex derivation as it involves only short wh- movement.

[4] Schulz's research does not investigate *wh- copying* constructions.

4.2   Experiment 3: L1 French; L2 English.

4.2.1. Participants

The initial number of participants in *experiment 3* was 36. As with the speakers for written *experiment 1*, the participants for *experiment 3* were recruited from the five-week immersion ESL program of the Centre Linguistique du Collège de Jonquière in Ottawa, Ontario (see chapter 3 for more details on the ESL program). It was not possible to administer an independent measure to determine the participants' level of proficiency, and thus, the placement test ran by the ESL program was used. The participants recruited were at the (lower)-intermediate level.

The learners were given the language background questionnaire used for the written experiments (see appendix 1; see also chapter 3 for details). Their age varied between 15 and 18 years old and they had had between 5 and 12 years of explicit ESL instruction prior to enrolment in the five-week immersion ESL program in Ottawa. Two of the participants reported Spanish as their L1 and were excluded. In addition, 8 French speaking participants were excluded at a later point because they were very comfortable and accurate asking complex wh- questions and thus the researcher switched to a different type of oral ESL practice appropriate for their higher level of proficiency.[5]

The final number of participants retained for the study was 26, as illustrated in table 13.

Table 13. Experiment 3: Number of participants
(French speaking learners of L2 English)

	Number	Excluded	Retained
L1 French	34	8	26
L1 Spanish	2	2	0
Total	36	10	**26**

In addition to these participants, 5 native controls were recruited. Since they behaved as expected, their results are not discussed in any detail. One important fact to point out is that occasionally the native controls would produce a simple question instead of a long-distance one (not more than one occurrence per person). After the first follow-up (see

---

[5] When recruiting the participants, the researcher promised that the experiment would provide them with a meaningful and fun L2 learning experience and thus did not proceed with the oral elicitation task if a participant seemed to find it too easy or uninteresting.

description of elicitation task below) the native speakers would easily re-phrase the question to make it a complex one.

4.2.2. Experimental task and procedures

Since long distance wh- questions are relatively infrequent in spontaneous production, most of the relevant L2 studies (Schulz 2006, Gutierrez 2005, Wakabayashi and Okawara 2003) have used a version of Thornton's (1990) elicited production task (see chapter 2 for details on the relevant L1 and L2 literature). The present study also employed an elicited production task (EPT) and, similarly to Thornton's original design, the participants were asked to play a guessing game; the task, however, was adapted to an adult L2 population. The props consisted of the following: a magnetic board with a hand-drawn picture of a two-storey house, which had a living-room, a kitchen and a bedroom; nine different magnets that could be placed at various locations in the house. The magnets were pictures of three characters (John, Silvia and Mary), two food items (an apple and a slice of pizza), two drinks (a glass of water and a glass of orange juice), and two kinds of reading material (a book and a newspaper). Photographs of the magnetic board with the house and the different magnets are provided in figures 29 and 30, respectively.[6]

Fig. 29 EPT: Magnetic Board   *Fig. 30 EPT: Magnets*

---

[6] The two additional characters (Joanne and Christine) drawn on the magnetic board were used in item/situation 10 of the elicited production task (see appendix 4).

The participants were invited to meet one-on-one with the researcher and play a guessing game in which they would need to ask different questions. After a brief warm-up, the researcher would introduce the magnetic board and ask the participant to describe what they saw in front of them. The magnets with the different characters and objects were always introduced and explicitly named by the researcher. The participant was asked to place the magnetic board on a stand in front of them so that the researcher could not see it. The game involved different situations in which the characters would need to go to a particular room and engage in different activities (eat, drink, sleep, watch TV, talk on the phone, kiss someone, etc.) For example, the participant would have to place one of the characters in any of the three rooms and make the researcher guess the location of the character or the activity in which the character was involved. An excerpt from the elicitation protocol is provided in (2). The full protocol is included in appendix 4.

(2) Elicitation protocol (excerpt)

Researcher:	Okay, now I'm going to be asking you to place the different people and the different objects somewhere. For example, can you place John in one of the rooms. But don't tell me where you put him. Okay?...... Now you want me to guess where you put him. So, you have to ask me a question. Okay?.... **So, ask me where I think John is.**
Participant:	Where do you think John is?
Researcher:	I think he's in the living-room.
Participant:	Yes.
Researcher:	Okay. So, because I guessed right, I get one point. (marks point). Okay. Now let's do the same with Mary. Put Mary anywhere in the house. But don't tell me where. Okay?.... So, you have to ask me where I think Mary is. Go ahead and ask.
Participant:	Where do you think she is?
Researcher:	I think she's in the kitchen…. cooking.
Participant:	No, she's in the living-room.
Researcher:	Ah, okay. So, because I didn't guess right, you get one point. (marks point; proceeds to next item).

While the above elicitation procedure targeted only complex questions, some participants, especially in the beginning of the game, would use simple questions (e.g. *Where is John?*

instead of *Where do you think John is?*). In those cases the researcher would provide up to three follow-up prompts in order to elicit a complex question. The protocol for the first follow-up is illustrated in (3).

(3) Elicitation protocol: first follow-up to a simple question.
Participant: Where is John?
Researcher: Well, I don't know for sure, so you have to make me guess. You have to ask me where *I THINK* he is.

In many cases, the participant would ask a complex question at this point. If they did not, however, the researcher would offer a second follow-up, as illustrated in (4).

(4) Elicitation protocol: second follow-up to a simple question.
Participant: Where is John?
Researcher: Okay, and if you have to use the word "THINK", how would you form the question? (Can you repeat the question using the word "THINK"?)

If the second elicitation follow-up failed as well, the researcher would use a third one, as in (5).

(5) Elicitation protocol: third follow-up to a simple question.
Participant: Where is John?
Researcher: Okay, and if you have to ask me "Where do you think dadadadada...?

Note that the first follow-up was used freely by the researcher; however, an effort was made to avoid the second and especially the third follow-up because they contained a strong clue as to the kind of structure the researcher was interested in.

It is also important to point out that the initial lead-in already provided a strong bias for a complex question. Essentially, it contained the targeted long-distance question, except it was in the form of an embedded question (see bold type in (2) above). The use of a biasing lead-in has been subject to some controversy in the literature. For example, Thornton (1990) used an elided question, as in (6) or (7), with child L1 English speakers.

(6) Now you can ask Cookie Monster where *he* thinks…
(7) Ask him where *he* thinks…   (my emphasis)

While Thornton's lead-ins were very successful in eliciting long-distance questions, they have been criticized for biasing the participant towards the *copying* construction (see Schulz 2006). That is, by providing the participants with the elided structure "where he thinks…", Thornton might have increased the chance of them copying the wh- word as in (8).

(8) *Where* do you think *where* Kermit hid the ball?   (my emphasis)

Alternatively, a lead-in using the word *what*, as in (9), would again arguably tempt the participants to complete the question with a *medial wh-* word, resulting in an increased chance of the *wh- scope marking* construction, as in (10).

(9) Ask him what he thinks…
(10) What do you think where Kermit hid the ball?

To avoid these problems, Schulz (2006) used what she called "neutral" lead-ins of the type "And now you ask X" or "And now X's opinion" for the majority of the questions; however, it should be noted that she resorted to an elided structure like "What do you think … dadadada….?" in the beginning of the elicitation protocol which also provides a strong bias. Gutierrez (2005), on the other hand, reasoned that a lead-in containing an elided question would be too difficult for the L2 participants to complete and would essentially fail to elicit the targeted constructions. Thus, she used full embedded question lead-ins, as in (2) above (bold type). She argued that if the participants were found to produce *medial wh-* utterances even after being given lead-ins essentially containing the correct English LD structure, then the evidence of access to structures unattested in the L1 or the L2, but within the hypothesis space of UG, would be even stronger. This reasoning was also followed in the present study, and thus, the full embedded question lead-in was adopted. However, in order to alleviate the strong biasing effect to the highest extent possible, these lead-ins were used only in the first three (or fewer) trials of the elicited production task. At that point, the lead-in was changed to a neutral one, as illustrated in bold type in (11).

(11) Elicitation protocol: neutral lead-in (introduced no later than the third trial).
Researcher: John is sitting on the couch (check comprehension, use word *sofa* if necessary) and he is kissing someone. **Go ahead and ask.**
Participant: Who do you think he is kissing?

Overall, while acknowledging that the lead-in has always been an area of contention in production tasks targeting *medial wh-* utterances, the present study tried to achieve the difficult balance between maximizing the likelihood of eliciting an LD question and minimizing the confound of a biasing lead-in. It is important to point out that even though different morphological and aspectual errors were not of interest in this study and will

therefore not be reported on in any detail, they provide some evidence that the participants did not engage in a mere repetition of the lead-in.[7] For example, a lead-in in the progressive, as in (12), would frequently generate a complex question in the habitual, as in (13), or with tense/aspect morphology entirely missing, as in (14).

(12) R: So, Silvia is eat**ing** something, right? ... Go ahead and ask.
(13) P: What do you think she eat**s**?   (participant FR 25)
(14) P: What do you think she eat$_{\emptyset}$?   (participant FR 2)

The fact that many participants failed to repeat the correct morphological information, provided to them immediately before the question they had to ask, is a strong indication that the chance of them simply repeating the researcher's lead-in was quite low.

The elicitation procedure targeted a total of 15 long-distance questions containing object and adjunct extractions as illustrated in table 14.

Table 14. Targeted long-distance structures

Targeted long-distance structure	Type of extraction	Wh- word
1) Where do you think John is?	adjunct	where
2) Where do you think Mary is?	adjunct	where
3) What do you think Silvia is eating?	object	what
4) What do you think John put in the fridge?	object	what
5) Who do you think Mary sent to buy eggs?	object	who
6) Who do you think Mary asked to buy butter?	object	who
7) What do you think he/she is drinking?	object	what
8) Where do you think he/she is watching TV?	adjunct	where
9) Where do you think he is sleeping?	adjunct	where
10) Who do you think she/he is calling?	object	who
11) Where do you think the newspaper is?	adjunct	where
12) What do you think she is reading?	object	what
13) Where do you think John put the book?	adjunct	where
14) Who do you think he is kissing?	object	who
15) Who do you think Silvia is talking to?	object	who

As with the written experiments, subject extractions were not targeted in the oral elicitation task due to the fact that inversion patterns in the embedded clause cannot be observed in such contexts (recall that inversion constitutes a key diagnostic with regards to the *direct* versus *indirect dependency* dichotomy).

---

[7] For discussion, see Gutierrez (2005) where some of the younger participants tended to mimic the researcher's lead-in of the type *Ask me where I think….* and use the same string of words in their question:
(i) *Where I think* the little baby fall asleep?   (participant 2)

The time the researcher spent with each participant was about 15 minutes. All interviews were recorded and all questions elicited by the guessing game were extracted and transcribed. False starts, hesitations, and pauses were transcribed impressionistically but are not represented in most of the examples below for the sake of simplicity. Self-corrections during production were also noted in the transcripts and during coding only the self-corrected version was taken into account (in cases where the participant completed a whole utterance and then self-corrected, both utterances were coded).

4.2.3. Group results

Experiment 3 elicited a total of 311 questions from the 26 participants (average of 12 questions per participant).[8] Of these, 260 (84%) were biclausal[9] and 51 (16%) were simple (8 fragments or incomprehensible utterances were excluded before tabulating the data). Recall that based on the Derivational Complexity Hypothesis, it was predicted that *medial wh-* constructions, as well as other structures shortening or eliminating the wh-dependency such as biclausal questions with short wh- movement (e.g. embedded questions) or no wh- movement (e.g. yes/no questions), simple questions, etc., would occur as less complex derivational alternatives to standard LD movement. As figure 31 illustrates, this prediction was borne out.

---

[8] A cut off of 5 relevant productions (i.e. questions) was set for a participant to be retained in the study. All 26 learners met this requirement. Recall, however, that 8 participants from the initial group of 33 were excluded because they seemed very comfortable with complex wh- questions and the researcher offered them another ESL exercise before they had completed 5 trials (see also footnote 5).

[9] The term biclausal is used here to incorporate both complex questions and sequential questions.

Figure 31. Distribution of target and non-target productions
Oral elicitation task, French speakers of L2 English, n=26
(total of 311 questions)
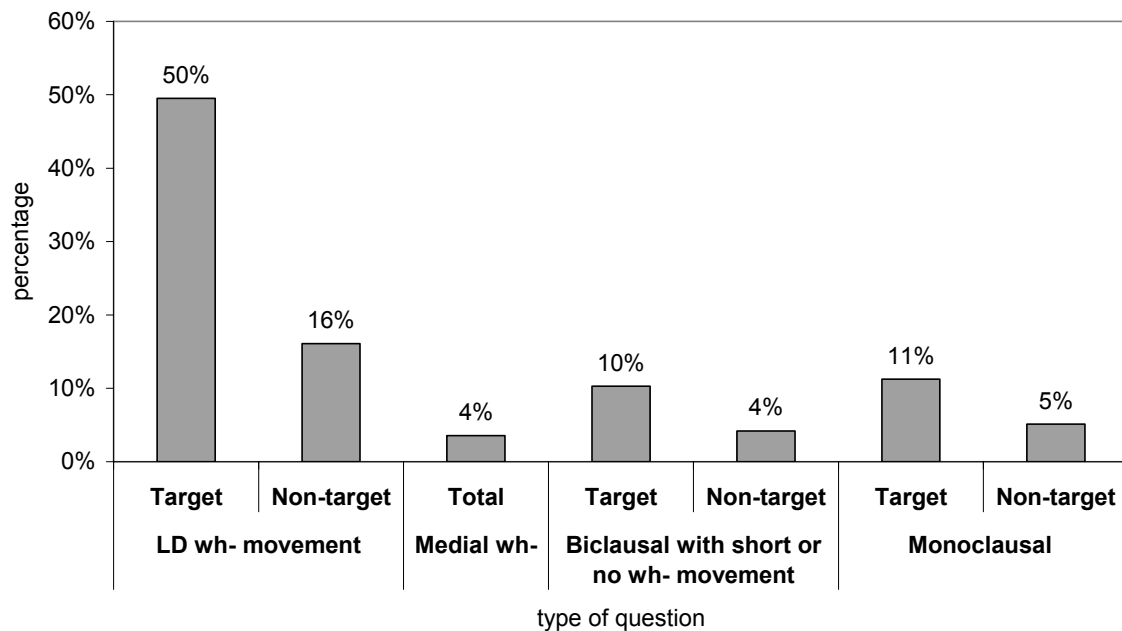
The productions in figure 31 add up to 100%, that is, they represent all 311 questions elicited. The data are grouped in four general categories: LD wh- movement, *medial wh-*, biclausal structures with short or no wh- movement, and monoclausal questions. The latter three categories are less complex computational and representational alternatives to LD wh- movement. As mentioned earlier, *medial wh-* fills an intermediate gap, which is expected to alleviate processing; biclausal structures with short or no wh- movement are easier than LD wh- movement because of the fewer Internal Merge applications they involve; finally, monoclausal questions also involve fewer applications of Internal Merge in comparison with LD wh- movement, and represent a less complex and usually shorter structure. Recall that the situations set up by the elicitation stories, the lead-ins, and the follow-up prompts were all designed to elicit LD wh- movement. The native speaker control group performed as expected and produced complex wh- questions virtually all the time (occasionally a simple question would be produced, but not more than one per person). The fact that the L2 learners produced wh- complex questions most of the time means that they were aware of the type of question most appropriate in the context of the elicited production task; at the same time, the fact that they also resorted to alternative,

derivationally simpler questions, which were generally inappropriate under the elicitation circumstances, is a strong indication that such questions were indeed used to avoid LD wh- movement (see also discussion below).

Each of the four general categories shown in figure 31 and the different subcategories which they contain will be discussed below. Starting with the category involving long-distance movement, both target (50%) and non-target (16%) productions were attested. Examples of target questions with LD wh- movement are listed in (15).

(15) Target LD   (50% of total productions)
    a. What do you think she's reading?   (participant FR 24)
    b. Who do you think is going to take the eggs?   (participant FR 6)
    c. What do you think he drinks?   (participant FR 10)
    d. Where do you think the newspaper are?   (participant FR 1)

The example in (15)*a* illustrates a wh- object extraction. Subject extractions, as in (15)*b*, were also attested, even though the experimental procedure did not aim to elicit such utterances (see table 14). Also note that some of the utterances coded as *Target LD* contained incorrect aspectual or agreement morphology, as in (15)*c&d*, respectively.[10] As mentioned earlier, morphological errors *per se* do not constitute the focus of this dissertation; furthermore, it has been claimed that they should be dissociated from abstract grammatical representations (see discussion of the Missing Surface Inflection Hypothesis in chapter 2). Thus, as long as morphological errors did not impinge upon determining the type of syntactic construction, they were ignored. This coding procedure was applied to all constructions listed in figure 31.

As for non-target long-distance structures, these included missing *do support* in the matrix clause, T-to-C movement in the embedded clause, and missing embedded auxiliary. Examples of such questions are given in (16)*a,b&c*, respectively. In addition, a limited number of *that-trace* violations (7 out of a total of 311 questions elicited), some of which included subject resumptive pronouns (3 of the 311 questions), were attested and are illustrated in (16)*d&e*.

---

[10] The sentence in (15)*c* is grammatically correct but pragmatically/aspectually infelicitous as the context in which the question was asked required the progressive.

(16) <u>Non-target LD</u> (16% of total productions)
  a. Where you think the book is?   (participant FR 3)
  b. Where do you think is John?   (participant FR 18)
  c. Who do you think Mary call?[11]   (participant FR 8)
  d. Who do you think that leaves the house to go take eggs? (participant FR 1)
  e. Who do you think that he go buy eggs?   (participant FR 4)

Since the above constructions involve a syntactic error (i.e. not purely morphological), they are differentiated from the target LD subcategory; however, it should be noted that these structures do not differ from the target with regards to wh- movement and cannot be considered avoidance strategies. As such, it can be assumed that learners who produce utterances as in (16) show evidence that the LD wh- movement mechanism is in place.

   The total rate of occurrence of LD wh- movement questions (target and non-target) was 66%, which makes this category the most highly attested one. The prevalence of such structures can be attributed to several factors. First, recall that LD wh- movement is a widely-attested typological option and is grammatical in both the native and the target languages of the participants. Thus, even in the early stages of acquisition, the appropriate LD mechanism may already be in place due to L1 transfer in combination with the L2 input[12]. Second, the elicitation lead-in described above might have given the participants a clue as to how to form complex wh- questions. Finally, recall that target LD wh- movement was also the most frequently chosen structure in the written experiments with both the French and the Bulgarian speaking populations at all levels (see chapter 3). Overall, this constitutes evidence that LD wh- movement is part of the interlanguage grammar even at the early stages of acquisition. It should be reiterated, however, that despite the high percentage of LD wh- movement constructions produced, the L2 learners did not perform like the native speakers. That is, LD wh- movement must pose a difficulty for these learners as 34% of the total productions involved derivationally less complex alternatives.

---

[11] This sentence is classified as non-target because of a missing embedded auxiliary. Alternatively, it could be considered target, as defined under (15), with missing 3rd person singular agreement morphology and infelicitous aspect. This is only a minor point, however, since the distinction between target and non-target LD utterances is not crucial for the purposes of this dissertation. The key fact is that in all these utterances long-distance wh- movement takes place.

[12] Although the occurrence of long-distance wh- movement questions is considered rare, the possibility that participants have "noticed" them in the target input cannot be disregarded.

Turning to the first general category of avoidance strategies in figure 31, *medial wh-* productions were attested at the rate of 4% of the total questions elicited. If the percentage of such structures is calculated based on the total number of complex wh-questions only (as was done in the previous L2 studies on medial), the rate increases to 6%. The number of tokens in this category was 11 (out of 311 questions elicited), and these included 5 without embedded inversion, as in (17), 5 with embedded inversion, as in (18), and 1 involving *silent scope marking*, as in (19). Note that the *silent scope marking* construction was included in the *medial wh-* category as it is essentially a variation of *wh- scope marking* where the scope marker in the matrix clause is null.[13] As discussed in chapter 2, such utterances are grammatical in some natural languages (e.g. Bahasa Indonesia, Quechua and Kitharaka) and have been attested in French child L1 acquisition (Oiry and Demirdache 2006) and in the L2 English of Japanese speakers (Okawara 2000; Wakabayashi and Okawara 2003). Furthermore, such utterances were also attested with the Bulgarian learners of L2 English at higher rates than with the French speaking population (see *experiment 4* below for details).

(17) *Medial wh-* without embedded inversion
    a. **Who** do you think **who** sent the butter?   (participant FR 10)
    b. **What** do you think the **what** John's drinking now.   (participant FR 19)
    c. **What** do you think about **where** Mary's sleeping?[14]   (participant FR 2)
    d. **What** do you think **where** he watch TV?[15]   (participant FR 2)
    e. **What** you think **what** Silvia eat?   (participant FR 20)

(18) *Medial wh-* and embedded inversion
    a. **What** do you think **where** is Mary?   (participant FR 18)
    b. **What** do you think **where** is the newspaper?   (participant FR 18)
    c. **What** do you think **who** is the person that John kiss?   (participant FR 18)
    d. **What** do you think **who** is talk to Silvia?   (participant FR 18)
    e. I can ask you **where** you think **where** is John?   (participant FR 2)

---

[13] Even though only one token of this structure was attested, it was included in the results because, as mentioned earlier, the approach taken in this dissertation is that all target and non-target productions should be reported and that even a small number of tokens can be informative (for further discussion, see section 4.2.5 below).

[14] This utterance can be considered grammatical assuming that the participant is asking a question along the lines of "What do you think about the particular place where Mary is sleeping?". Since the elicitation context did not call for a question with such a reading, the utterance above is (intuitively) considered to be an example of *medial wh-*. However, see also (20)*d* and the following discussion, as well as section 4.2.5.

[15] Note that (17)*d&e* are ambiguous structures as that they lack an auxiliary and/or verbal morphology showing presence or absence of embedded inversion. Nonetheless, they are included (intuitively) in this category because it is unnecessary to make a new category (*medial wh- ambiguous*) for two productions only. However, such utterances are more prevalent in the results of the next experiment and thus a separate category will be created for the Bulgarian speaking population (see (35) below).

(19) *Silent scope marking*

    Ø Do you think **where** is Mary?   (participant FR 19)

Since the number of *medial wh-* productions is limited, the answers to the first three research questions (repeated below for convenience) are only tentative.

R1. Do *medial wh-* representations exist in the interlanguage of French and/or Bulgarian learners of L2 English?

R2. If *medial wh-* is attested, is it of the *wh- copying* or the *wh- scope marking* kind?

R3. If *medial wh-* is attested, does it constitute a *direct* or an *indirect dependency*?

With regards to **R1,** it could be argued that even though *medial wh-* productions occurred at a very low rate, they were still attested in the interlanguage of a few French speaking learners of English and need to be accounted for. On the other hand, such low rates of occurrence may be viewed as noise. Considering that *medial wh-* was attested at much higher rates in the written experiment with a similar French-speaking population, I will adopt the former proposal and assume that the low rates of occurrence in the spoken experiment are due mainly to the small sample size. Recall that Gutierrez (2005) found *medial wh-* productions in a relatively limited subset (12%) of a very large participant population (i.e. 32 out of 260 participants produced *medial wh-*). Thus, it is possible that due to the smaller pool of participants in *experiment 3*, individuals who produce higher rates of such *medial wh-* utterances were simply not recruited. It should be acknowledged, however, that Schulz (2006) tested a relatively small English L2 population (28 participants with L1 Japanese) and attested higher occurrence of *medial wh-* utterances than the current experiment. Nonetheless, in her data there were also 12 individuals who never produced *medial wh-* and another 5 who produced it only once.[16]

Another import factor to take into account is that the French speaking L2 learners of English were in an ESL context (i.e. English is a dominant language in their country of residence), whereas the Japanese and the Spanish/Basque speakers reported on by Schulz and Gutierrez, respectively, were in an EFL context (i.e. English is a foreign language in the country of residence). As such, it is very likely that the French speakers of the current

---

[16] As an additional point, the fact that Japanese is a wh- *in situ* language may have an influence on the rate of *medial wh-* productions. It could be that learners whose L1 is a wh- movement language produce lower rates of *medial wh-*, whereas learners whose L1 is an *in situ* language produce higher rates of *medial wh-*.

study had a much higher degree of exposure to and experience with spoken English, resulting in a higher spoken proficiency than that of the Japanese and Spanish/Basque EFL learners.

Taking into account these factors, I assume that even though the oral production data did not elicit a sufficient number of tokens to provide a reliable answer to **R1**, future research would be able to attest more robust rates of *medial wh-* production in a French speaking population.[17] Further discussion of the reasons why *medial wh-* constructions were not attested at higher levels and why they should not be viewed as noise in this experiment will be provided in section 4.2.5 below. Turning to **R2**, 4 out of the 11 *medial wh-* utterances involve *wh- copying* and the remaining 7 are *wh- scope marking* structures.[18] This mirrors the pattern established in the written experiments. Regarding **R3**, 5 of the 11 utterances represent *direct dependency* (i.e. no embedded inversion) and the other 6 *indirect dependency* (i.e. presence of embedded inversion).[19] Once again, the answers to **R1-R3** are only tentative and cannot be pursued any further.

The data regarding the next research question, **R4** (repeated below for convenience), however, is much more conclusive.

> **R4.** What other target and non-target constructions related to the acquisition of complex wh- questions exist in the productions of the two learner populations?

Recall that **R4** was formulated in anticipation that even if the LD wh- movement mechanism is in place in the interlanguage, some learners might find such structures difficult to produce due their derivational complexity and associated processing pressures (e.g. working memory constraints); thus it was predicted that various strategies (other than *medial wh-*) of shortening or eliminating the wh- dependency would be attested. As indicated in figure 31, both target and non-target biclausal and monoclausal structures were produced as alternatives to LD wh- movement. The target structures in the biclausal category accounted for 10% of the total utterances and include embedded wh- questions,

---

[17] For future research on this issue it may be better to recruit French speakers who are acquiring English in an EFL setting (i.e. learners who do not reside in Canada or in another country where they would be immersed in the L2).

[18] As with the written experiments, questions involving *what* in both the matrix and the embedded clauses are considered *wh- copying* to reflect the fact that the wh- element in both clauses is phonologically the same. However, such constructions can also be analysed as *wh- scope marking*.

[19] See footnote 15.

complex yes/no questions, wh- clauses followed by an *if* clause (*wh-…if…*), and wh-clauses followed by a prepositional *about/of* clause (*wh-…about/of…*). These are illustrated in (20)*a,b,c&d*, respectively.

(20) Target biclausal structures with short wh- movement or no wh- movement
(10% of total productions)
    a. Do you know who buy eggs?   (participant FR 2)
    b. Do you think John's watching TV in the bedroom?   (participant FR 19)
    c. What do you think if John is tired?   (participant FR 13)
    d. What do you think about the place I put the newspaper? (participant FR 2)

The embedded question example in (20)*a* involves only short wh- movement in the second clause, while in the yes/no question in (20)*b* wh- movement is avoided altogether. The questions with *if* and *about* clauses, as in (20)*c&d*, respectively, are similar to the embedded question in the sense that LD wh- movement is avoided and only local wh-movement applies in the first clause. However, what is particularly interesting about the utterances with *if* and *about* clauses is that they are pragmatically completely infelicitous under the elicitation circumstances. Recall that the elicitation production task set up a context for the participants to ask what the researcher thinks about a specific activity or location of an object or a character. In (20)*c*, the target LD wh- structure would be *What do you think John is drinking?* and in (20)*d* it would be *Where do you think I put the newspaper?* Instead of producing such utterances, the participants use the *if* and *about* clauses which result in open-ended questions inviting the researcher to volunteer the answers, without being specifically asked for them. It seems that the participants producing these questions know what they want to ask, but since LD wh- movement poses difficulty for them, they formulate the question in a vague, pragmatically infelicitous way, using a less complex structure. In other words, the structure employed is grammatically well-formed, but contextually inappropriate. [20] This introduces an interesting new distinction in categorizing the different productions. So far, the different avoidance strategies were divided into target and non-target based on their grammaticality (abstracting from morphological errors, etc.). However, a further distinction needs to be introduced based on the contextual appropriateness of the

---

[20] It could be argued that (20)*c* is ungrammatical. However, if we closely consider just the structure of the sentence, it becomes obvious that it is licit. That is, if we replace *do* by *would* and *is* by *were,* the sentence becomes absolutely grammatical, showing that the ungrammaticality was due only to modal/aspectual choice, which are factors that I abstract away from in this discussion.

productions. Thus, the examples in (20)*a&b* are both grammatically well-formed and contextually appropriate, while the examples in (20)*c&d* are grammatically well-formed but contextually inappropriate. More details on the implications of this distinction will be provided in the discussion section below.

Apart from the target biclausal structures used as LD wh- movement avoidance, there were also non-target ones which accounted for 4% of the total utterances. These include sequential questions without a *wh-* word in the second clause (to be distinguished from the *medial wh-* sequential representation where a wh- word appears in both clauses), clefts, and wh- *in situ*, and are illustrated in (21)*a,b,&c*, respectively.

(21) Non-target biclausal structures with short wh- movement or no wh- movement
    (4% of total productions)
    a. What do you think? Is Joanne or Christine Silvia call?   (participant FR 18)
    b. You think it's who go to the search butter?   (participant FR 3)
    c. You watch TV and you drink what?   (participant FR 13)

The example in (21)*a* involves two sequential interrogative clauses where wh- movement applies only in the first one. Note the use of disjunction in the second clause, which allows for the avoidance of a wh- word altogether. That is, by using disjunction, the participant obviates the need of using the wh- word *who* as in *Who is Silvia calling?* While this example is grammatically ill-formed because of the incorrect word order in the second clause, note that use of sequential questions is contextually appropriate (i.e. it is a felicitous substitution of the target LD wh- movement utterance *Who do you think Silvia is calling?* The cleft structure in (21)*b* is also contextually appropriate, although ungrammatical in English. Such structures in the L2 of this learner population can be attributed entirely to L1 transfer because they are a licensed option in the L1, as illustrated earlier and repeated in (22) for convenience.

(22) a. *Tu   dis   que **c'est**   où   que Marie va?*
         you   say   that it is   where   that Marie goes?
         'Where is it that you say Mary is going?'

Finally, the non-target production in (21)*c* is also attributed to L1 transfer because, as mentioned earlier, questions involving wh- *in situ* are licensed in French as well.

Turning to the last category of avoidance strategies, monoclausal questions (target and non-target) were attested in 16% of the total productions. Monoclausal wh- questions

involve only local movement and as such are derivationally less complex than the target LD wh- movement. Target simple questions, illustrated in (23), involve T-to-C movement, while their non-target counterparts, as in (24), lack T-to-C movement.[21]

(23) Target simple questions (11 % of total productions)
    a. What food does she eat?   (participant FR 3)
    b. According to you, who is outdoor the house?   (participant FR 5)
(24) Non-target simple questions (5% of total productions)
    a. Who John is kissing?   (participant FR 17)

With regards to the pragmatic context, such questions are generally inappropriate because they ask only about the location/activity of the character/object but not about the researcher's opinion. In other words, the "*What do you think…*" clause of the question is missing in these simple structures. One exception to this were simple questions preceded by *according to you*, as in (23)*b*, which are both grammatically well-formed and contextually appropriate. A total of 8 out of the 311 questions elicited were of this type. Since such questions were never produced by the native speakers, however, they were classified as a wh- movement avoidance strategy.

      To summarize the group results of the elicited production task with the French-speaking learners of L2 English, a wide variety of target and non-target constructions were attested. The target LD wh- movement productions accounted for 50% of the data; another 16% were also LD wh- movement structures which involved syntactic errors such as lack of *do support*, T-to-C movement in the embedded clause, missing embedded auxiliary, *that-trace* violations and subject resumptive pronouns. *Medial wh-* constructions accounted for 4% of the total productions (6% of the total LD wh- movement constructions) and included *medial wh-* with and without embedded inversion and *silent scope marking*. The third general category included biclausal questions with short or no wh- movement. The target (grammatical) utterances in this category accounted for 10% of the total productions and included embedded wh- questions, yes/no questions, wh- clauses followed by an *if* clause, and wh- clauses followed by a prepositional *about/of* clause. It was observed that even though the questions with *if* and *about/of* clauses were grammatically well-formed, they were contextually completely

---

[21] Note that the example in (23)*b* has a lexical choice error (*outdoor* instead of *outside*) but such errors, similarly to morphological inaccuracies and/or omissions, were ignored.

inappropriate. The non-target (ungrammatical) structures in the biclausal category with short or no wh- movement represented 4% of the total productions and included sequential questions[22], clefts, and wh- *in situ*. Finally, the monoclausal question category included target simple constructions with T-to-C movement (11%) and non-target simple constructions without T-to-C movement (5%); both of these types of questions are infelicitous in the context of the elicitation task as they lack the part of the question that asks about the researcher's opinion.

What is noteworthy about the group results of the L2 learners is the large variety of different constructions attested. By contrast, the native controls' productions were almost exclusively restricted to LD wh- movement. At first sight, this constitutes somewhat of a puzzle because the native speakers are expected to have a larger array of structures at their disposal. However, considering that LD wh- movement is associated with a high degree of processing and representational difficulty, and following the predictions of the Derivational Complexity Hypothesis discussed earlier, it is not surprising that the L2 learners resort to a wide variety of structures that are less complex than LD wh- movement (a total of 44%); the native speakers, on the other hand, do not seem to be affected by the higher degree of difficulty of LD wh- movement, and produce such questions consistently, because they are grammatically well-formed and contextually most appropriate for the particular discourse circumstances created by the oral elicitation task. Further consideration of the results with regards to the Derivational Complexity Hypothesis will be provided in the discussion section below. Before that, however, I will report on the individual results.

4.2.4. *Individual results*

The general question to ask with regards to the individual results is how the different target and non-target productions listed in figure 31 above are distributed among the participants. Are there learners who have acquired fully the LD wh- movement mechanism and therefore produce such utterances all the time? And on the other hand, are there learners for whom LD wh- movement represents a significant obstacle and thus

---

[22] Recall that these have a wh- element only in one of the two clauses and are to be distinguished from the *medial wh-* with embedded inversion (sequential question) representation which has a wh- element in both clauses.

they consistently resort to avoidance strategies? Note that such strategies could include both target (grammatical) and non-target (ungrammatical) structures. For example, simple questions, embedded questions, complex yes/no questions, etc., can be completely grammatical. Nonetheless, if used consistently instead of LD wh- movement, they represent avoidance mechanisms in a speaker's interlanguage.

The distribution of the individual responses is given in table 15.

Table 15. Distribution of individual responses: FR learners of L2 English (experiment 3)

	Participant	1	2	3	4	5	6	7	8	9	10	11	12	13
LD wh- movement	Target LD	13	3	1	13	3	13	11	5	6	11	10	5	1
	Missing do support	0	1	9	0	0	0	1	1	0	0	0	0	0
	Emb. Inversion	0	0	1	0	3	0	0	2	0	0	0	0	1
	Missing emb. aux.	0	1	1	0	0	0	1	5	0	3	3	2	0
	That-trace violation	2	0	0	2*	0	0	1	2	0	0	0	0	0
	Resumptive pronoun	0	0	0	2	0	0	0	0	0	0	0	1[†]	0
LD wh- movement avoidance	Medial wh-	0	2	0	0	0	0	0	0	0	1	0	0	0
	Medial wh- & emb. inv.	0	1	0	0	0	0	0	0	0	0	0	0	0
	Silent scp marking	0	0	0	0	0	0	0	0	0	0	0	0	0
	Emb. wh- question	0	2	0	0	0	0	0	0	0	0	0	0	0
	Yes/No question	0	0	0	0	0	0	0	0	0	0	0	3	0
	Wh-…if…	0	0	0	0	0	0	0	0	0	0	0	0	5
	Wh-… about / of…	0	8	0	0	0	0	0	0	0	0	0	0	0
	Sequential no wh- in 2nd clause	0	0	0	0	0	0	0	0	0	0	0	0	0
	Cleft	0	0	2	0	0	0	0	0	0	0	0	0	0
	Wh- in situ	0	0	0	0	0	0	0	0	0	0	1	2	2
	Simple target	2	0	4	0	9	2	0	4	0	1	2	3	0
	Simple without inv.	0	0	1	0	2	1	0	4	0	0	2	3	0
	Excluded	1	0	1	0	0	0	0	1	0	0	0	0	2
	Total	18	18	20	15	17	16	14	24	6	16	18	18	11

Table 15. Continued

	Participant	14	15	16	17	18	19	20	21	22	23	24	25	26
LD wh- movement	Target LD	4	5	2	3	1	11	6	5	5	5	3	4	6
	Missing do support	0	0	0	0	0	0	1	0	0	0	0	1	0
	Emb. Inversion	0	0	0	0	1	1	0	0	0	0	0	0	0
	Missing emb. aux.	0	0	0	1	0	1	0	0	1	0	0	0	0
	That-trace violation	0	0	0	0	0	0	0	0	0	0	0	0	0
	Resumptive pronoun	0	0	0	0	0	0	0	0	0	0	0	0	0
LD wh- movement avoidance	Medial wh-	0	0	0	0	0	1	1	0	0	0	0	0	0
	Medial wh- & emb. inv.	0	0	0	0	4	0	0	0	0	0	0	0	0
	Silent scp marking	0	0	0	0	0	1	0	0	0	0	0	0	0
	Emb. wh- question	0	0	0	0	0	0	0	0	0	0	0	0	0
	Yes/No question	0	0	3	0	0	2	0	0	1	0	0	0	0
	Wh-…if…	0	0	0	0	0	0	0	0	0	0	0	0	0
	Wh-… about / of…	0	0	0	0	3	0	0	0	0	0	6	0	0
	Seq. no wh- in 2nd clause	0	0	0	0	6	0	0	0	0	0	0	0	0
	Cleft	0	0	0	0	0	0	0	0	0	0	0	0	0
	Wh- in situ	0	0	0	0	0	0	0	0	0	0	0	0	0
	Simple target	1	1	1	1	2	1	0	0	0	0	0	0	0
	Simple without inv.	0	0	0	1	0	0	0	1	0	0	0	1	0
	Excluded	0	0	0	1	1	0	0	0	1	0	0	0	0
	Total	5	6	6	7	18	18	8	6	8	5	9	6	6

*   utterance also coded as resumptive pronoun
†   utterance also coded as a complex question without wh-

As the data in table 15 indicate, there were a number of learners who were very comfortable producing the long-distance wh- dependency deleting all intermediate wh-copies. Five of the 26 participants (4, 7, 9, 23 and 26) showed exclusively LD wh-patterns. Note that apart from the target LD construction, some of these participants'

productions included a few of the non-target LD utterances mentioned earlier (i.e. missing do support, embedded inversion, missing embedded auxiliary, *that-trace* violations and resumptive pronouns). Nonetheless these constructions cannot be considered avoidance strategies, as successive cyclical movement of the wh- phrase has taken place.

On the other end of the spectrum are 7 participants (2, 5, 12, 13, 16, 18 and 24) who employed avoidance strategies involving a wide variety of target (grammatical) and non-target (ungrammatical) constructions more frequently than complex questions with LD wh- movement. For example, participant 2 produced a total of 18 questions, 8 of which were the *wh-… about…* structure illustrated in (20)*d*, 2 were embedded wh- questions, as in (20)*a*, 2 were *medial wh-* without embedded inversion, 1 was *medial wh-* with embedded inversion, and 5 were LD wh- movement structures (target and non-target). As such, only 27% (5/18) of this participant's utterances involved LD wh- movement and the rest were various structures avoiding it.[23] Another example of use of a wide variety of avoidance strategies is provided by participant 12. Of 18 questions elicited, this participant used 2 in-situ structures, 3 yes/no questions, 6 simple questions (3 with and 3 without inversion) and 7 long-distance wh- movement structures (39% LD wh- movement vs. 61% avoidance strategies). A third example of use of several different avoidance strategies is provided by participant 18. Of 18 productions, only 2 involved long-distance wh- movement (1 target and 1 with embedded inversion); the rest included 4 *medial wh-* structures with embedded inversion (i.e. sequential questions), 6 *wh-…is…* structures, as in (21)*a* (i.e. sequential questions without a wh- word in the second clause), 3 *wh-…about…* structures, as in (20)*d*, 2 simple target questions, and 1 fragment (excluded).

The remaining 3 participants (5, 24, and 13) who strongly relied on strategies avoiding long-distance wh- movement did not resort to a wide variety of structures; rather, they generally focused on one particular strategy and used it consistently. Participant 5 produced a total of 17 questions, 11 of which were simple (9 target and 2 without inversion), and 6 were complex (3 target and 3 with embedded inversion). Considering

---

[23] Recall that *medial wh-* constructions are included as a long-distance avoidance strategy because in such structures the dependency is "shortened" or "bridged" by the extra wh- word filling the intermediate gap.

that the elicitation task targeted complex questions, and that special follow-up prompts were included in the design to encourage participants to switch from simple to complex questions, this participant showed very strong resistance to LD wh- movement. Participant 24 produced a total of 9 questions, 6 of which were of the *wh-… about…* type, as in (20)*d*, and 3 were target long-distance wh- movement. Finally, participant 13 had a total of 11 productions, 2 of which constituted LD wh- movement (1 target and 1 with embedded inversion), 2 were wh- *in situ*, 5 were the *wh-…if...* structure illustrated in (20)*c,* and 2 were excluded.

Apart from the two ends of the spectrum, there were also participants who used a mixture of long-distance wh- movement and avoidance strategies (3, 8, 17, 11, 19, 14, 6, 15, 21, 25, 22, 10, 20, and 1). The general picture that emerges from these data is that there are three distinct groups of participants. The first group (*group A*) includes learners who have not acquired LD wh- movement sufficiently well or can rarely produce such structures due to a processing deficit at this point of acquisition; this group shows a higher number of avoidance strategies than LD wh- movement. The participants who fall in the second group (*group B*) still use avoidance strategies, but are already showing a preference for LD wh- movement; that is, the LD wh- movement is used more frequently than avoidance strategies. The third group (*group C*) contains the participants who have already acquired LD wh- movement sufficiently well to apply it consistently in oral productions; the learners in this group produce exclusively questions with LD wh- movement. A relative distribution of the use of avoidance strategies versus target LD wh- movement and the division of the participants in these three groups is illustrated in figure 32.[24]

---

[24] The productions for each participant in this figure add up to 100%. Thus, participant 18, for example, uses 12% LD wh- movement and 88% LD wh- movement avoidance structures, participant 13 uses 22% LD wh- movement versus 78% avoidance strategies, and so on.

**Figure 32. Individual distribution of LD wh- movement versus
avoidance structures. Oral elicitation task, French speakers of L2
English, n=26**
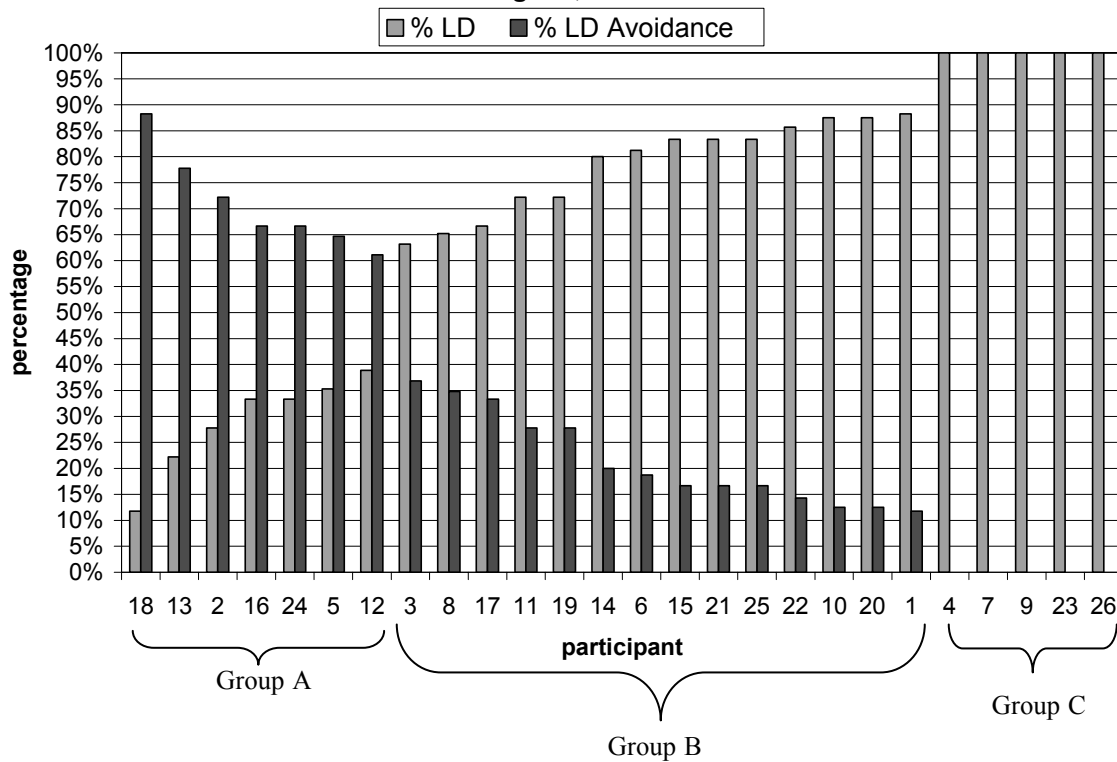
The behaviour of the participants in *groups A* and *C* is straightforward. While the former clearly need more experience with the L2 in order to be able to produce LD wh- structures movement more frequently, the latter seem to be completely comfortable using this mechanism. The participants in *group B*, on the other hand, show a pattern which can be interpreted in different ways. It can be argued that they still have not acquired LD wh- movement sufficiently well and thus resort to some use of avoidance strategies. Alternatively, it could be that the LD wh- mechanism is already in place, but due to higher processing deficits at the earlier stages of acquisition, these participants are often unable to produce such structures. Note that 9 of these participants (14, 6, 15, 21, 25, 22, 10, 20 and 1) use avoidance strategies only 20% of the time or less, and produce LD wh- movement 80% of the time or more. Thus, it can be assumed they have acquired LD wh- movement in the L2 sufficiently well, but occasionally produce structures without LD wh- movement for the sake of variety, or for some other reasons (recall that some of the native speakers mentioned earlier in this chapter also occasionally produced simple questions). The rest of the participants in group B, who have higher rates of avoidance

structures and lower rates of LD wh- movement, are probably still somewhat uncomfortable with the latter, and need more experience in order to strengthen their ability to use this mechanism.

4.2.5. Discussion

The results of the oral experiment with French speaking learners of L2 English proved to be different from the ones of the written experiment with an equivalent population; furthermore, the results differed from other L2 studies which employed similar oral elicitation procedures. Essentially, while *medial wh-* productions were attested in the current production data, their rate was limited. Possible explanations in terms of the proficiency level, the sample size and the acquisition setting was already proposed in section 4.2.3. In addition, it is important to point out that *medial wh-* was just one way of avoiding long-distance wh- movement, and a few other avoidance strategies were attested at similar or even lower rates. For example, the cleft constructions and the wh- *in situ* utterances were even more limited than the use of *medial wh-*; nonetheless, clefts and wh- *in situ* can not be discounted as accidental because they are licensed in the L1 of the participants and represent clear cases of transfer. As such, the low rates of *medial wh-* constructions should not be considered noise. Overall, the picture emerging from *experiment 3* is one in which multiple productions (sometimes occurring at low rates) are spread across the different participants. Interestingly, despite the variation across learners, what unifies 12 of the 18 types of productions attested (see table 15) is that they represent different ways of avoiding movement of a wh- phrase from one clause to another (i.e. LD wh- movement). The immediate question that this finding raises is what motivates the existence of such avoidance structures.

As already mentioned in section 4.1, it is possible to seek an explanation for such facts in terms of the derivational and processing status of the utterances. Different versions of the idea that the complexity of the derivation influences sentence processing and/or acquisition have existed in the literature since the 1960s. In the early stages of generative theory (after Chomsky 1957), it was proposed that the number of operations applying to a particular derivation can be correlated with behavioural data such as reaction times on different psycholinguistic measures. The idea was that the more

grammatical operations applied to a particular sentence, the more difficult (i.e. slower) its processing would be. While it is often argued that this early proposal known as the *Derivational Theory of Complexity* (DTC) did not generate enough empirical support, some of its insights have been preserved over the years (for discussion of the DTC see Duffield 2006 and references therein; see also Berwick and Weinberg 1984; Phillips 1996; Marantz 2005, and references therein). In Jakubowicz and Stirk (2008; see also Jakubowicz 2004, 2005), a modern version of this theory called the *Derivational Complexity Hypothesis* (DCH) is applied to French and Dutch speaking children's L1 acquisition, and more specifically to *medial wh-* phenomena. In essence, the claim is that less complex derivations are "input convergent (i.e., correctly spelled out and 'pronounced' at the interfaces) before more complex ones" (Jakubowicz and Strik 2008, p. 106). In other words, all other things being equal, derivations that are less complex appear earlier in language development than ones that are more complex. As explained in section 4.1, Jakubowicz (2005) proposed a *derivational complexity metric* (DCM) whereby short movement is considered less complex than long-distance movement, and moving one constituent is less complex than moving multiple constituents. As Jakubowicz and Strik (2008) claim, different types of *medial wh-* structures are expected to precede structures involving LD wh- movement, in which all intermediate copies have been deleted. Under their analysis, the *silent scope marking* and *wh- scope marking* strategies are less complex than LD wh- movement because they involve only partial movement of the wh- phrase to the left periphery of the embedded clause. The *wh-copying* construction, on the other hand, is closer in derivational complexity to the standard LD wh- mechanism. However, even if we assume that the two are essentially the same structure, the version with the overtly pronounced intermediate copy is easier to process; that is, the intermediate copy alleviates the working memory (WM) load. Following such logic, Jakubowicz and Strik (2008) discuss French and Dutch L1 children's wh- productions and argue that they generally fit within the predictions made by the DCH.

If we adopt similar reasoning in L2 acquisition, the facts from *experiment 3* receive an explanation. Due to the higher grammatical and processing load imposed by LD wh- movement structures, the French speaking learners of L2 English resort to a

variety of avoidance strategies. *Wh- copying*, *wh- scope marking* and *silent scope marking* are among these strategies, but the learners' productions need not be limited to them. Biclausal structures such as embedded questions, yes/no questions, various sequential questions, *wh-… if…* structures, and *wh-… of/about…* structures are all less complex derivational alternatives to LD wh- movement; in other words, such constructions involve only local wh- movement or absence of a wh- phrase altogether. Simple (target and non-target) wh- questions are also derivationally less complex than LD wh- movement not only because they involve local wh- movement, but also because they are comprised of just one clause. With such a variety of options available to the L2 learners, it is not surprising that their utterances aimed to avoid the complexity of LD wh- movement are not limited to *medial wh-* constructions.

As mentioned in section 4.1 of this chapter, I espouse a weak version of the DCH. That is, the DCH provides an explanation for the fact that long-distance wh- movement poses a higher degree of difficulty compared to *medial wh-*, to a variety of biclausal structures with short or no wh- movement, and to monoclausal questions. However, I do not propose a specific hierarchy within the attested avoidance strategies. It could be argued, for example, that in terms of complexity – and therefore order of acquisition – monoclausal structures with short wh- movement precede biclausal structures with short movement, which in turn precede long-distance wh- movement. Such a strict order, however, was not observed in the data of the present experiment where the target LD wh- movement option often co-existed with different avoidance strategies. Indeed, such strong predictions are unlikely to be borne out by acquisition data in general for several reasons. First, complexity can be measured in terms of the number and length of wh- movement applications but other derivational factors such as non-wh- movement, agreement processes, null categories, etc. may have additional effects. It is very difficult to determine how to weigh such factors and how they really affect derivational complexity and acquisition orders. Second, it is very likely that L1 transfer (i.e. the fact that the LD wh- movement mechanism is part of the L1) plays a strong role and thus many of the participants showed evidence of co-existence of LD wh- movement constructions with avoidance strategies rather than a strict acquisition hierarchy. Finally, the relative frequency of different LD and short movement wh- constructions (both in the

L1 and the L2) in combination with the discourse circumstances (i.e. the context set up by the elicitation task) also have an effect on the types of productions attested. Due to the difficulty in predicting and measuring the precise effects of all these factors, a strict derivational and acquisition hierarchy of avoidance strategies is not pursued in this dissertation. Instead, a more realistic view of the L2 acquisition process where a multitude of factors are at play is adopted. Essentially, the higher degree of difficulty LD wh- movement poses due to its derivational complexity and associated processing load is uncontroversial; however, a more fine-grained definition and implementation of the DCH is left to future research.

Having addressed the question regarding the nature of the LD wh- movement avoidance strategies and the underlying reasons for their existence, it is also important to discuss the issue of their variation across the participants. The results presented in figure 32 show a gradual shift of preference from avoidance strategies to LD wh- movement. These results can be viewed as stages in the process of acquisition of LD wh- movement. Note that despite the fact that all the participants were at the (lower)-intermediate level, some variation with regards to their proficiency is possible as the placement test did not include a spoken component. Thus, it could be that some speakers had a higher level of spoken proficiency than others and used fewer or no avoidance strategies. Alternatively, it could be that even if the participants did have similar levels of spoken proficiency, some of them had a lower working memory (WM) capacity and thus had to resort to LD avoidance strategies more often. This is based on the assumption that the standard LD wh- movement option poses a high processing burden because it requires intermediate copies of the wh- word to be kept active in the working memory. Thus, participants with more limited WM capacity, especially during the early stages of acquisition, are expected to resort more often to alternative structures where only short or no wh- movement applies. This idea fits well within a proposal developed by Zukowski (2001; to appear) with regards to relative clause formation by speakers with Williams Syndrome (WS). Essentially, Zukowski argues that people with WS show evidence of intact grammatical representations but due to their more limited processing resources, compared to typically developing children and adults, they produce certain types of RCs less frequently. By the same token, availability of processing resources may have especially strong effects in the

early stages of L2 acquisition, and as such, learners with lower WM capacity may use the target grammatical LD wh- movement option less frequently. Presumably, at the more advanced stages of acquisition individual differences in WM capacity would not play such an important role and all learners would be able to converge with the target, showing consistent use of LD wh- movement. Of course, since psycholinguistic measures such as WM capacity tests were not included in this study, a definitive answer to these issues is left open to future research. Overall, however, the results presented above remain a very strong indication that because of the high degree of derivational complexity and associated processing pressures, long-distance wh- movement constitutes a difficulty in early L2 development. Thus, before the learners become completely comfortable using this structure, they resort to a variety of avoidance strategies. Further research incorporating psycholinguistic techniques may be able to determine whether non-linguistic factors such as working memory capacity can be linked to the amount of LD avoidance used.

A more general question that needs to be addressed in this discussion is what the results of *experiment 3* offer in terms of our understanding of L2 acquisition processes and SLA theory. Note that the idea that acquisition is constrained by Universal Grammar, which was adopted in the previous chapter can be applied successfully to the oral production data as well. The avoidance strategies used by the learners are mostly options consistent with the UG hypothesis space: embedded questions, complex yes/no questions, sequential questions with or without a wh- word, and simple wh- questions are all licensed options in the target language; clefts and wh- *in situ*, on the other hand, are structures that are clearly coming from the L1; finally, *medial wh-* structures (including *silent scope marking*), are options that are unattested in both the L1 and the L2, but attested in other languages of which the participants report no knowledge. The probabilistic learning mechanism adopted in the previous chapter comes into play again in explaining the lower rates of *medial wh-* productions: such structures are unattested in the L1 and L2 and thus their probability of occurrence would be lower than the probabilities associated with the wide variety of other avoidance strategies supported by the native grammar and/or the target input. As such, the high degree of variability of utterances attested in this experiment proves to be constrained by both UG and domain-

general probabilistic mechanisms (for further discussion of the issue of variability see next chapter).

To return to the issue of L2 access to UG, one of the challenges for the supporters of this view is to show that L2 learners' utterances never fall outside of the limits of a pre-defined hypothesis space. That is, if structures unattested in the languages of the world appear in the interlanguage of an L2 learner population, this would provide evidence against UG access in L2 acquisition. Crucially, all productions elicited in *experiment 3* fall within the UG hypothesis space. With regards to *medial wh-*, it has been noted in earlier literature that children acquiring English as their first language, never produce *medial wh-* words in infinitival clauses, as in (25), and that such constructions are ungrammatical crosslinguistically (i.e. even in the languages that license *medial wh-* in the adult grammar).

(25) *Who do you want who to eat the pizza?   (Thornton, 1990:213)

The fact that the participants in *experiment 3* did not produce such utterances, even though many of the elicitation contexts were compatible with them (see appendix 4), provides strong support for the claim that interlanguage grammars are UG-constrained. Of course, it could be argued that the total number of *medial wh-* productions was limited and thus it is difficult to claim that *medial wh-* in infinitival contexts is impossible in the participants' L2 grammar. I assume that this is unlikely to be the case, considering that none of the other productions in *experiment 3* included structures that are not licensed UG options. The non-target LD productions (i.e. ungrammatical in the L2) involved the following: morphological and lexical errors, which were largely ignored (see section 4.2.3 for justification); inversion in the embedded clause, which is licensed in different languages of the world and also in certain dialects of English (e.g. Irish English); and *that-trace* violations and resumptive pronouns, which are also licensed crosslinguistically. As such, the data does support the idea that L2 learners' non-target productions fall within the limits of a pre-defined hypothesis space. Even the two types of production attested in this experiment that seemed most inappropriate from a pragmatic point of view, the *wh-... if...* and the *wh- ... about / of...* structures, fall within the realm of UG. Examples of these two productions are repeated in (26) and (27), respectively.

(26) <u>Wh-… if…</u>
    a. How do you think if Mary like to read?   (participant FR 13)
    b. What do you think if John is very tired?   (participant FR 13)
(27) <u>Wh-… about/of…</u>
    a. What do you think about the drink of John?   (participant FR 2)
    b. What do you think about John's put something in the fridge? (FR 2)
    c. What do you think about the person who buy butter?   (FR 2)

Note that in (26)*a*, for example, the researcher has told the participant that Mary would like to read something (either a book or a newspaper); thus, the participant is expected to make the researcher guess what the reading material is. In this context, the participant's question is infelicitous. Similarly, the rest of the utterances in (26) and (27) are inappropriate for the particular elicitation contexts in which they occur (see appendix 4 for the elicitation contexts). Two possible accounts for these utterances can be pursued. On the one hand, it could be that the participants did not understand the particular story/context provided to them by the researcher before they asked the question. Alternatively, it could be that they understood the story/instructions well and formed the appropriate semantic representation of the question, but simply framed it in an inappropriate structure. Of course, it is impossible to determine with certainty which of these two scenarios applies to each of the above utterances. What is remarkable, however, is that regardless of their comprehension of the situations or their inability to ask about the pragmatically salient information, these participants still opted for structures licensed by UG, and in fact by the target language. Essentially, the *if…* clauses and the *about…* prepositional phrases/clauses are contextually inappropriate but grammatically well-formed constructions. As Zukowski (to appear; see also references therein) points out, cases of contextually inappropriate but grammatically well-formed language errors have been reported in typically developing children acquiring their L1, in people with Williams Syndrome (WS), and in L2 acquisition. Zukowski argues that people with WS thus show evidence of unimpaired grammatical competence despite their atypical brain development and mild language deviations. That is, such speakers are able to acquire normal grammars, even though their productions are at times somewhat deviant from those of normally developing individuals (i.e. higher rate of certain types of errors) due to different processing and/or cognitive deficits. In parallel to Zukowski's claim, the theoretical account developed in this dissertation assumes that L2 learners are also able to

acquire target grammars, despite the fact that their productions are often not target-like and differ from those of native speakers along certain dimensions. The fact that some participants in *experiment 3* opted for well-formed English structures, even if they did not have the necessarily comprehension and/or production facility to ask contextually appropriate questions shows strong evidence that their grammars are unimpaired and constrained by UG. That is, their language instinct helps them provide target-convergent structures even in situations of comprehension and/or production deficits.

The final issue this discussion aims to address pertains to the definition of the proposed avoidance strategies. Essentially, all types of productions (both target and non-target) attested in the experiment, and illustrated in figure 31, that did not involve LD wh-movement were classified as avoidance strategies. This was based on the reasoning that such structures involve an intermediate wh- element that alleviates processing, shorten the distance of moving the wh- word, or eliminate wh- movement altogether. One potential criticism regarding this claim could be that these productions (target or non-target) might not necessarily be related to LD wh- movement. Ungrammatical utterances that do not involve LD wh- movement, for example, could presumably result from independent reasons, which are unrelated to LD avoidance. Grammatical utterances that do not involve LD wh- movement, on the other hand, could be just alternative ways of expression preferred by certain speakers, and again, not necessarily avoidance strategies. That is, speaker variation (both within and across individuals) can be conditioned by a number of linguistic, social, situational, and other factors.

While the above potential scenarios may seem plausible, several factors in terms of both the experimental design and the results obtained point against them. First, recall that the native speaker control group consistently produced target complex wh- questions involving long-distance wh- movement and deletion of all intermediate copies. The native speakers did produce some simple wh- questions but these were limited to not more than one per person (i.e. the task clearly called for LD wh- movement). By contrast, the L2 learners produced a wide variety of target and non-target structures which lacked LD wh- movement. Second, as discussed in section 4.2.2, each context/story of the elicitation protocol was specifically designed to elicit complex wh- questions and the initial lead-in itself contained LD wh- movement; furthermore, up to three different

follow-up prompts were provided to help the learners produce a complex wh- question. Considering that all these factors combined constituted a strong clue as to the type of construction required, analysing the productions without LD wh- movement as avoidance strategies seems well motivated. Finally, some additional evidence can be drawn from the prosodic properties of the L2 learners' productions. As mentioned earlier, hesitations, false starts, pauses, fillers and paraphrases were transcribed only impressionistically and since their formal analysis remains beyond the scope of this dissertation, they were not included in the examples given so far. Nonetheless, it is worth mentioning that such factors also indicate that the utterances produced as alternatives to LD wh- movement were indeed avoidance strategies. Consider the examples in (28)-(32), which include more detailed transcription.

(28) a. What do you think there… again… like… Do you think there's some TV in the… the room?   (participant FR 16)
b. What do you think eh…no... eh…Do you think there's some… do you think there's a bed or a sofa?   (participant FR 17)

(29) a. What do you… no… no no no…. Ah, my God, again. What do you think… is Silvia or John do… go outside?
b. What do you think uh.. where is the newspaper?   (participant FR 18)

(30) Uhm… do you think… do you think or think… do you think where is Mary?
                                    (participant FR 18)

(31) a. Ah…. where do you think ah…. Maria ah… where?
b. Where you think the newspaper… where?   (participant FR 12)

(32) a. What do you think uh… if you have hungry, you eat what?   (prt FR 13)

In (28)*a&b* the participants attempt to produce LD wh- movement structures following the researcher's lead-in but fail to complete them and thus rephrase the utterances using complex yes/no questions (i.e. wh- movement is eliminated). In (29)*a*, the participant starts with a wh- structure but pauses (likely because it is difficult to establish the LD wh-dependency) and then proceeds to ask sequential questions (one with local wh-movement and one without wh-). In (29)*b*, the same participant again pauses after the first clause producing the filler *uh*, and then instead of continuing with the LD wh-movement structure produces a *medial wh-* (sequential) question. In (30) the speaker attempts to ask the question two times failing to include an overt wh- word in the matrix clause; in the third attempt, not being able to complete the LD wh- movement, the learner pronounces the wh- word in medial position producing a *silent scope marking*

construction. In (31)*a* the participant pauses between the two clauses of the sentence and then after another pause in the embedded clause produces a wh- word *in situ* doubling the one in the matrix clause. The same speaker produces an equivalent utterance in (31)*b*. Finally, in (32) the participant is unable to complete a LD wh- question and introduces a pause after the first clause; then the speaker introduces an if clause (could be a separate biclausal question) which ends with the wh- word *in situ*. The point that all these examples illustrate is that the learners are indeed struggling in trying to compute LD wh- movement structures in their productions, most likely because of processing limitations; thus they resort to alternative structures which shorten or eliminate the long-distance wh- dependency. The attested pauses, hesitations, and paraphrases provide supplementary evidence that the structures in question are indeed the result of avoidance mechanisms, and are not due to other reasons such as speaker variation or accidental L2 errors.

It should be noted that apart from the above arguments for the validity of what I identify as avoidance strategies, independent support for my proposal comes from a recent study of LD wh- movement in the L2 English of Japanese speakers by Radford and Yokota (2006). These authors also argue that production of short movement in contexts where native speakers use long-distance questions is a direct consequence of the difficulty posed by LD wh- movement. In addition, Radford and Yokota also found that all productions attested (target and non-target) were within the UG hypothesis space and thus the learners' interlanguage is constrained by UG principles.

This concludes the discussion of the results of *experiment 3*. In what follows, I present the data from an equivalent oral experiment conducted with Bulgarian learners of L2 English.

4.3   Experiment 4: L1 Bulgarian; L2 English.

The purpose of *experiment 4* was twofold: on the one hand, it was designed to supplement the results from the written experiment with Bulgarian learners of L2 English reported on in the previous chapter; on the one hand, it was intended to provide a comparison base for the findings of the oral experiment with the French speaking learners of L2 English. The research questions and experimental methodology (background questionnaire and elicited production task) were the same as in *experiment 3*. What

differed was the population (i.e. Bulgarian L1 speakers) and the acquisition setting (i.e. EFL and not ESL). As mentioned in the previous chapter, a valid comparison with regards to the proficiency of the Bulgarian versus the French learners of L2 English is not possible due to the fact that they were not given the same proficiency tests; in fact, even if they were given the same tests, the comparison would still be to some extent inaccurate due to the different acquisition settings. Because of this, the general assumption adopted was that the target population for the *experiment 4* would be at the (lower)-intermediate level, and that this EFL population would likely perform at a lower level in spoken English than an equivalent ESL population (immersed in the target language). As mentioned earlier, despite the fact that it was not possible to administer a common proficiency test to both populations, a comparison between the French and the Bulgarian results can be drawn due to the nature of the research questions and the phenomena investigated. That is, the purpose of the experiment was to determine whether *medial wh-* constructions can be attested and to investigate possible LD wh- movement avoidance strategies. Thus, comparisons of these phenomena (as long as they are attested) can be drawn between the two populations while acknowledging the potential differences in proficiency levels.

4.3.1. Participants

The participants for *experiment 4* were recruited at the Department for Language Teaching and International Students at Sofia University. Some of them were university students while others were working professionals enrolled in an EFL course at the university. As such, the age and backgrounds of the participants varied more in comparison with *experiments 1,2,* and *3*. The total number of participants was 31 and they were between 18 and 40 years old. The level of proficiency was between low-intermediate and intermediate, as established by their EFL program's placement test. One participant was excluded as he was very comfortable with complex wh- questions and did not seem interested in the oral elicitation task (i.e., the researcher switched to a different type of oral ESL activity appropriate for the participant's higher level of proficiency). The remaining 30 participants were retained, as illustrated in table 16.

Table 16. Experiment 4: Number of participants
(Bulgarian learners of L2 English)

	No. of participants
Total	31
Excluded	1
Retained	**30**

4.3.2. Group results

Experiment 4 elicited a total of 416 questions from the 30 participants (the average number of questions per participant was about 14). Of these, 331 (79%) were biclausal questions[25] and 85 (21%) were monoclausal (9 fragments or incomprehensible productions were excluded before tabulating the data). The overall number of *medial wh-* productions was higher in comparison with the French group. Figure 33 shows the distribution of the different target and non-target productions across all participants.

Figure 33. Distribution of target and non-target productions
Oral elicitation task, Bulgarian speakers of L2 English, n=26
(total of 416 questions)

---
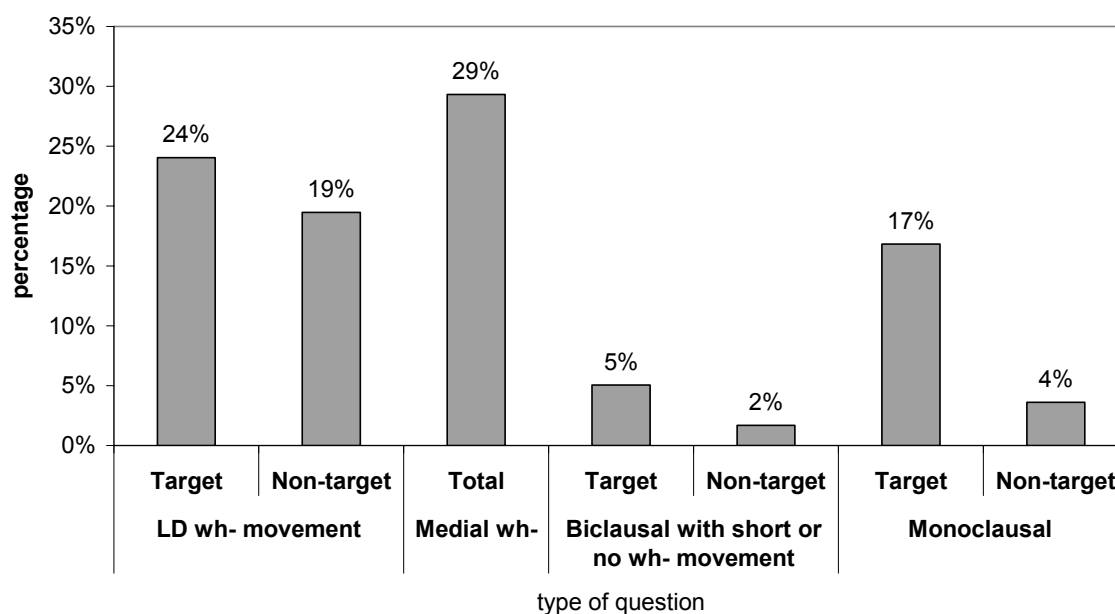
[25] As with *experiment 3*, the term biclausal here incorporates both complex and sequential questions.

As these data indicate, the target LD wh- movement structure featuring deletion of all intermediate wh- copies had the second highest rate (24%) of all productions.[26,27] Thus, the Bulgarian participants have a lower level of accuracy in comparison with the French ones whose target LD wh- movement structures had the highest rate of all productions (50%; refer to figure 31 in section 4.2.3 for more detailed comparisons). However, similarly to the French speaking participants, the Bulgarian learners of L2 English resorted to a wide variety of LD wh- movement avoidance strategies; the overwhelming majority of these were of the same types that were attested with the French group. What is different is the ratio between target LD wh- movement and avoidance strategies in the two groups. On the one hand, the French group had a higher number of LD wh- movement (target and non-target) than the Bulgarian participants (66% vs. 43%, respectively). The Bulgarian group, on the one hand, had a higher number of avoidance structures than the French group (57% vs. 34%, respectively). This result was predicted since the Bulgarian learners acquire English in an EFL setting, and as such have less experience with the L2, especially with spoken English. Furthermore, since the two different populations were given different placement tests, the proficiency label (lower)-intermediate used for both is only a relative one. Nonetheless, the fact that some of the most frequently occurring structures in the productions of the Bulgarian participants involved target LD wh- movement shows that acquisition is going in the same (target) direction. That is, the French and the Bulgarian learners seem to be following the same path of acquisition, the former being more proficient in speaking (i.e. showing higher accuracy) due to their higher level of exposure to the TL.[28]

Since most of the production types shown in figure 33 overlapped with the ones attested with the French speaking participants, they will not be discussed in detail. The

---

[26] As in the experiment with French speaking participants, morphological and aspectual errors and omissions were ignored, as long as they did not interfere with determining the type of syntactic structure. In addition, impressionistically speaking, the Bulgarian population was less accurate with prepositions, lexical choice and other not necessarily syntactic aspects of the L2. Such errors and omissions were also ignored.

[27] Also in parallel with *experiment 3*, subject extraction questions were attested despite the fact that they were not specifically targeted (recall table 14).

[28] Another interesting difference between the French and the Bulgarian populations was that the productions of the latter involved pauses, hesitations, false starts and paraphrases more often than those of the former. As mentioned earlier, these were noted only impressionistically in the transcripts and a quantitative analysis is beyond the scope of this dissertation. It is nonetheless worth pointing out that this observation provides additional support for the proposed differences due to the EFL versus the ESL setting of acquisition.

non-target LD wh- movement category (19% of the total productions) in figure 33 included utterances lacking *do support* in the matrix clause, T-to-C movement in the embedded clause, and missing embedded auxiliary. These non-target productions were also attested in *experiment 3* (for examples of such utterances see (16) above). Also in parallel with the data from *experiment 3*, a limited number of *that-trace* violations was attested (a total of 3); unlike *experiment 3*, no resumptive pronouns were found. As mentioned earlier, the non-target LD wh- movement productions involve syntactic errors but cannot be classified as avoidance strategies, unlike the rest of the categories in figure 33.

One of the most interesting findings of the oral experiment with Bulgarians learners of L2 English is that *medial wh-* constructions were attested at a much higher rate in comparison with *experiment 3*. A total of 122 *medial wh-* tokens were produced (29% of the total productions). Of these, 18 tokens were of the *[+]medial wh-[–]embedded inversion* type (8% of all complex questions[29]) and 76 tokens were of the *[+]medial wh- [+]embedded inversion* type (32% of all complex questions). Another 16 tokens (7% of all complex questions) were ambiguous between the two types because insufficient morphology or lack of an auxiliary masked the absence or presence of inversion in the second clause. Examples of these three types of productions are given in (33), (34), and (35), respectively.

(33) *Medial wh-* without embedded inversion (8% of all complex questions)
    a. What do you think what he's doing now?   (participant BG 1)
    b. Where do you think where John is watching TV?   (participant BG 18)
    c. What do you think who John kisses?   (participant BG 24)
    d. What do you think where Mary is sleeping?   (participant BG 20)

(34) *Medial wh-* with embedded inversion (32% of all complex questions)
    a. What do you think who did Mary sent to buy some eggs? (participant BG 4)
    b. What do you think where is John now?   (participant BG 12)
    c. What do you think who is John kissing at moment?   (participant BG 17)
    d. What do you think what is the John drinking?   (participant BG 28)

---

[29] As mentioned earlier, sequential questions without wh- are excluded from the count of complex questions but included in the count of biclausal utterances. The total number of complex questions was 231.

(35) *Medial wh-* (ambiguous)   (7% of all complex questions)
    a. What do you think who John kissing?   (participant BG 22)
    b. What do you think who read newspapers now?   (participant BG 1)
    c. What do you think what read Silvia?   (participant BG 10)
    d. What do you think who John kiss?   (participant BG 15)

In addition, 12 tokens (5% of all complex questions) of the *silent scope marking* construction, as illustrated in (36), were attested.

(36) *Silent scope marking* (5% of all complex questions)
    a. Do you think who Mary sent to buy eggs?   (participant BG 16)
    b. Do you think who Silvia's calling?   (participant BG 15)
    c. Do you think who is go out for the shopping? (participant BG 6)

As discussed earlier, *silent scope marking* constructions were included in the *medial wh-* category, as they can be analysed in the same way as the overt *scope marking* structures illustrated in (33)*c&d*. The only difference is that in the *silent scope marking* construction, the element appearing in the left periphery of the matrix clause to mark scope and type the structure as interrogative is covert. As discussed in chapter 2, *silent scope marking* constructions have been attested in languages such as Bahasa Indonesia, Quechua and Kitharaka.

      Recall that the first three research questions formulated for the oral experiments aimed to find out if *medial wh-* would be attested in the L2 learners' spoken productions, and to examine some of its properties. These questions are repeated below.

    R1. Do *medial wh-* representations exist in the interlanguage of French and/or Bulgarian learners of L2 English?
    R2. If *medial wh-* is attested, is it of the *wh- copying* or the *wh- scope marking* kind?
    R3. If *medial wh-* is attested, does it constitute a *direct* or an *indirect dependency*?

While the occurrences of *medial wh-* in *experiment 3* were too limited to warrant discussion of the findings with regards to these three questions, the rates in *experiment 4* are much higher and make it possible to examine these issues. **R1** was already addressed in the discussion of the percentages of *medial wh-* as in (33)-(36) above; that is, *medial wh-* productions accounted for 29% (122/416) of the total productions and 53% (122/231) of all complex questions elicited. As for **R2**, *wh- scope marking* (including *silent scope marking*) constituted 76% (93/122) of the total *medial wh-* productions versus 24% (29/122) for the *wh- copying* construction. Note that productions involving *what* in both

clauses were classified as *wh- copying*; however, it is equally possible that such utterances are in fact of the *wh- scope marking* type, in which case the asymmetry would be even greater (see chapter 3 for similar patterns and reasoning). **R3**, the question regarding the nature of the dependency can be answered only tentatively; that is, because the experiment did not incorporate independent measures of the participants' knowledge of inversion in matrix and embedded clauses, conclusive evidence from inversion patterns cannot be drawn. Nonetheless, it is important to note that the overwhelming preference of the L2 learners was for the *[+]medial wh- [+]embedded inversion* option (i.e. *indirect dependency*). The tokens of this type represent 81% (76/94) of the total *medial wh-* utterances versus 19% (18/94) for the *[+]medial wh- [–]embedded inversion* type (*direct dependency*).[30] As such, the *indirect dependency* (sequential wh- questions) *medial wh-* representation was attested to a much higher degree than the *direct dependency*. Recall that the former does not pose a learnability problem (i.e. sequential questions are attested in the learners' L1 and L2) while the latter does (i.e. it is not attested in either the L1 or the L2 of the participants). This asymmetry and its relation to the results from the written experiments will be addressed further in the discussion section below.

In discussing the data obtained in *experiment 3*, the Derivational Complexity Hypothesis (Jakubowicz and Strik 2008) and its predictions about *medial wh-* structures were adopted. The data from the Bulgarian speakers of L2 English in *experiment 4* are also consistent with the DCH. As discussed earlier, Jakubowicz and Strik proposed that because the *wh- scope marking* construction involves only partial wh- movement (i.e. to the embedded clause), it is derivationally less complex than its counterpart with LD wh- movement. *Wh- copying*, on the other hand, involves the same derivational steps in terms of movement. However, since it is reasonable to assume that the pronounced *wh- copy* in medial position is a source of alleviating the working memory load involved in the computation, the *copying* construction is also considered easier to derive and process than

---

[30] With regards to the *silent scope marking* construction, no specific predictions about the relationship between embedded inversion and the nature of the dependency were postulated. Thus, such constructions are excluded from the count here. However, since it seems unreasonable to analyse *silent scope marking* utterances as well-formed sequential questions even when inversion is present in the second clause (e.g. **Do you think… where is he going?*), they should perhaps be considered *direct dependency* in all cases. Following this reasoning, the total percentage of *direct dependency medial wh-* would be higher.

the target LD wh- movement structure. With regards to the *direct* versus *indirect dependency* dichotomy, the latter representation is less complex from a derivational and processing perspective because it involves two independent local wh- movements. Overall, these predictions are well supported by the results of *experiment 4*. The *indirect dependency* (sequential questions) was attested at a much higher rate than the *direct* one (32% vs. 8% of the total complex questions produced, respectively). Furthermore, the *wh- scope marking* construction was also attested at a much higher rate than *wh- copying* (76% vs. 24%). As such, the formulation of the DCH adopted in this chapter proved to have strong explanatory power with regards to the *medial wh-* productions in *experiment 4*.

The next research question, **R4**, addresses other types of LD avoidance strategies, and is repeated below for convenience.

> **R4.** What other target and non-target constructions related to the acquisition of complex wh- questions exist in the productions of the two learner populations?

As already mentioned, *medial wh-* structures constituted only one way of LD wh- movement avoidance in the experiment with the French speaking population. In addition to these, a number of other biclausal and monoclausal target and non-target structures that involved shortening or eliminating the wh- dependency were attested. This was also the case with the Bulgarian learners of L2 English. With a few minor exceptions, the Bulgarian population used the same avoidance strategies as the French participants. With regards to the target biclausal category with short or no wh- movement (5% of the total productions) the same structures as in the previous experiment were attested: embedded wh- questions, complex yes/no questions, sequential questions without wh- in the second clause[31], and *wh-… if…* structures. Examples of such productions for the French speakers were provided in (20). Equivalents from the Bulgarian population are listed in (37).

(37) Target biclausal structures with short wh- movement or no wh- movement
<div style="text-align:right">(5% of total productions)</div>

      a. Do you know where John is now?   (participant BG 4)
      b. Do you think John is at the kitchen?   (participant BG 11)
      c. What do you think does John like Silvia?   (participant BG 1)
      d. What do you think if John put a glass of water in the fridge? (prt BG 12)

---

[31] To be distinguished from *medial wh- indirect dependency* where both clauses contain a wh- word.

While the utterances in (37)*a,b* and *c* are not as contextually appropriate in the situations created by the elicited production task as the equivalent target LD wh- movement structure, they may be considered acceptable, at least to some degree. The *wh-…if…* structure in (37)*d* (1 token only), on the other hand, is completely infelicitous. As discussed in *experiment 3*, it could be that in such cases the participant has a poor understanding of the situation described in the elicitation task; alternatively, the participant may fully comprehend the context but be unable to select a numeration and a structure that allows for the appropriate semantic representation. In either case, what is noteworthy is that even when participants ask completely infelicitous questions, they are able to resort to a grammatically well-formed structure (recall discussion of Zukowski 2001, to appear, in section 4.2.5).

The next category in figure 33 involves non-target biclausal structures with short or no wh- movement, which constitute only 2% of the total productions. This low rate parallels the results from the French speaking population where equivalent utterances were attested 4% of the time. In the results of the French participants these included sequential questions (without a wh- word in the second clause), clefts and wh- *in situ*. The Bulgarian participants produced non-target embedded questions (with inversion in the embedded clause), and wh- *in situ* (1 token only), as illustrated in (38)*a&b*, respectively.

(38) <u>Non-target biclausal structures with short wh- movement or no wh- movement</u>
                                                    (2% of total productions)
    a. Do you know where is the newspaper?   (participant BG 3)
    b. What do you think John is kissing which?   (participant BG 18)

Turning to the last category of avoidance strategies in *experiment 4*, monoclausal questions were attested at a total rate of 21%, as indicated in figure 33. The target utterances in this category accounted for 17% of the total productions. As with the French speaking participants, these were simple wh- questions with target T-to-C movement, as in (39)*a&b*; in addition, one token of a simple yes/no question was attested, as illustrated in (39)*c*.

(39) <u>Target simple target questions</u>　　　　(17% of the total productions)
  a. Where is he?　　　　　　　　　　　　(participant BG 24)
  b. Who did you go to the shop?　　　　　(participant BG 21)
  c. Does Silvia like pizza?　　　　　　　(participant BG 7)

As with all other productions, morphological and aspectual errors and omissions, as well as lexical choice, preposition use, etc., were ignored as long as the target wh- short movement and subject-auxiliary inversion had taken place. Thus, the utterance in (39)*b*, for example, was coded as a simple target question, despite the fact that the participant incorrectly used the verb *go* instead of *send*.

Finally, the non-target simple questions represented 4% of the total productions. These utterances lacked T-to-C movement or an auxiliary, as illustrated in (40)*a&b*, respectively.

(40) <u>Non-target simple questions</u>　　　　　(4% of the total productions)
  a. Who John is kissing?　　　　　　　　(participant BG 22)
  b. What Silvia read?　　　　　　　　　(participant BG 16)

To summarize the group results of *experiment 4*, the types of productions elicited from the Bulgarian learners of L2 English were very similar to the ones elicited from the French speakers. The distribution of the LD wh- movement versus avoidance strategies was also essentially equivalent. The overall lower accuracy of the Bulgarian speakers can be explained by the different setting of acquisition and by the potentially lower level of spoken proficiency. The same explanation applies to the higher error rate in general, and in particular to the overall very high occurrence of *medial wh-* (*indirect dependency*) constructions.

4.3.3. Individual results

As with *experiment 3*, the questions with regards to the individual results in *experiment 4* are how the different target and non-target productions are distributed among the participants and whether a particular learner tends to focus on one type of production (e.g. a particular avoidance strategy) or uses a wide variety of structures. The individual results are provided in table 17.

Table 17. Distribution of individual responses: BG learners of L2 English (experiment 4)

	Participant	1	2	3	4	5	6	7	8	9	10	11	12	13	14
LD wh- movement	Target LD	1	7	0	1	8	4	2	2	5	4	5	1	5	5
	Missing do support	0	0	0	0	0	1	0	0	0	0	0	0	0	1
	Emb. Inversion	0	0	0	1	0	2	1	5	0	0	3	0	0	1
	Missing embedded auxiliary	0	0	0	0	0	1	0	0	0	1	2	0	0	1
	That-trace violation	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LD wh- movement avoidance	Medial wh-	2	1	0	0	1	1	0	0	0	2	0	0	0	0
	Medial wh- & emb. inv.	9	0	7	9	0	1	0	0	0	6	1	7	0	0
	Medial wh- ambiguous	3	0	0	0	0	0	0	0	0	2	0	2	0	0
	Silent scope marking	0	0	0	0	1	1	0	0	0	0	0	0	0	0
	Target emb. wh- question	0	0	0	1	1	0	1	0	0	0	0	0	0	0
	Yes/No question	0	0	0	0	0	0	0	0	0	0	1	0	0	0
	Wh-…if…	0	0	0	0	0	0	0	0	0	0	0	1	0	0
	Sequential no wh- in 2nd clause	1	0	0	0	0	0	0	0	0	0	0	3	0	0
	Non-target emb. wh- question	0	2	1	0	0	0	1	0	0	0	0	0	0	0
	Wh- in situ	0	0	0	0	0	0	1	0	0	0	0	0	0	0
	Simple target	0	0	2	3	2	7	2	1	0	2	2	5	1	2
	Simple without inversion	1	0	0	0	0	0	0	0	0	2	0	0	0	1
	Excluded	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Total	17	10	10	15	13	18	8	8	5	19	14	19	6	11

Table 17. Continued

Participant	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
Target LD	4	1	2	5	4	7	2	2	6	6	0	2	0	1	8	0
Missing do support	0	0	0	0	2	1	2	1*	0	0	1†	0	0	0	0	0
Emb. Inversion	1	1	0	0	2	0	3	2	2	1	0	2	1	1	2	4
Missing emb. auxiliary	7	0	0	1	0	2	4	4	0	1	9	2	0	0	1	0
That-trace violation	0	0	0	1	0	0	0	0	0	0	0	2	0	0	0	0
Medial wh-	1	0	0	2	0	2	0	0	0	3	0	0	1	0	0	2
Medial wh- & emb. inv.	1	0	13	1	1	0	0	1	0	1	0	0	14	2	0	2
Medial wh- ambiguous	1	0	0	0	0	0	0	6	0	0	0	0	0	2	0	0
Silent scope marking	1	8	0	0	0	0	0	0	0	0	0	0	0	0	0	1
Target emb. wh- question	2	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0
Yes/No question	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3
Wh-…if…	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Sequential no wh- in 2nd clause	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3
Non-target emb. wh- question	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0
Wh- in situ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Simple target	0	3	3	3	3	0	3	2	0	2	2	2	8	5	4	2
Simple without inversion	3	1	0	0	1	0	0	3	0	0	0	0	0	1	1	0
Excluded	0	0	0	1	0	0	0	0	0	0	1	3	0	1	1	2
Total	21	14	18	20	13	12	14	20	8	14	12	13	24	13	17	19

The rows "Target LD" through "That-trace violation" are grouped under **LD wh- movement**; the rows "Medial wh-" through "Simple without inversion" are grouped under **LD wh- movement avoidance**.

* Utterance also coded as "emb. inversion".
† Utterance also coded as "missing emb. aux".

As table 17 indicates, the data from the Bulgarian speaking participants are distributed in a similar way to the data from the French speaking participants (cf. table 15 above).[32] On the one hand, there are learners who used almost exclusively LD wh- movement. Particularly noteworthy in this respect are two participants (9 and 23) who show no evidence of avoidance strategies at all, even though both of them occasionally produce

---

[32] The only differences in terms of categories between table 17 (BG participants) and table 15 (FR participants) are as follows: the BG participants did not produce resumptives, clefts and *wh-… about/of* structures; the FR participants did not produce non-target embedded questions, and ambiguous *medial wh-* utterances were not included as a separate category (see footnote 15).

non-target inversion in the embedded clause. Ten other participants (2, 8, 11, 13, 14, 20, 21, 25, 26 and 29) also show a distinct preference for LD wh- movement versus avoidance strategies. As already mentioned, these results pattern very similarly with the French data discussed earlier. The one difference was that the French group had a larger number of participants (a total of 5) who produced exclusively LD wh- movement constructions.

On the other end of the spectrum are participants who showed a strong preference for LD wh- movement avoidance structures (1, 3, 4, 10, 12, 16, 17, 27, 28 and 30). Participant 3, for example, failed to produce LD wh- movement altogether: out of 10 utterances elicited, 7 involved *medial wh-* with embedded inversion (*indirect dependency*), 2 were simple wh- questions, and one was a complex yes/no question. This suggests that the participant either has not acquired LD wh- movement in the TL at all, or cannot use such structures in oral production at this stage of L2 development due to their derivational and processing complexity. Similar reasoning applies to participants 12 and 27 who produce almost exclusively avoidance structures as well. For participant 27, out of 24 utterances only 1 features LD wh- movement; the rest are simple wh- questions (8/24), *medial wh-* with embedded inversion (14/24) and *medial wh-* without embedded inversion (1/24). Participant 12 also has only 1 production with LD wh- movement; the rest are *medial wh-* with embedded inversion (7/19), ambiguous *medial wh-* (2/19), sequential questions without wh- in the second clause (3/19), simple questions (5/19), and the *wh-… if…* structure (1/19).

In the middle of the spectrum are participants who use both LD wh- movement and avoidance strategies at similar rates, or at least not with such a strong preference for one or the other (5, 6, 7, 15, 18, 19, 22 and 24). As already mentioned in the description of the individual results of *experiment 3,* it could be argued that such participants have some difficulty with long-distance wh- movement but acquisition of this mechanism is in process, and thus avoidance strategies are used only part of the time. Alternatively, it could be that the LD wh- movement mechanism is part of the grammatical competence, but cannot be made full use of during oral production because it requires a certain degree of automatization in order to overcome working memory or other processing limitations.

With regards to the question whether learners who resort to avoidance strategies focus on only one or two types, or whether they use a wide variety of avoidance structures, the data also pattern in parallel with the results from *experiment 3*. That is, there are learners who use almost exclusively one or two avoidance strategies, and learners whose productions show a wider variety. An example of the former are participants 3, 4, 16, 17 and 27 who focus mainly on the *medial wh-* sequential representation as an avoidance mechanism. These participants also produce some simple questions (especially participant 27), and occasionally an embedded question (target or non-target). Participant 16 focuses very strongly on the *silent scope marking* construction as an avoidance strategy. Of 14 elicited productions, this participant provided 8 tokens of *silent scope marking*, 4 simple questions and 2 LD wh- movement structures. As indicated earlier, the presence of the *silent scope marking* construction, repeated for convenience in (41), is quite interesting because it constitutes a learnability problem.

(41) *Silent scope marking*
    a. Do you think where John watching TV?   (participant BG 16)
    b. Do you think who Mary sent buy to buy eggs?   (participant BG 16)
    c. Do you think what is Silvia eating?   (participant BG 16)

Recall that such structures are illicit in both the L1 and the L2 of the participant, but are nonetheless attested crosslinguistically (see chapter 2 for examples of such structures in Bahasa Indonesia, Quechua and Kitharaka). Furthermore, they have been attested in the L1 acquisition of French (Oiry and Demirdache 2006; Jakubowicz and Strik 2008) and in the L2 acquisition of English (Wakabayashi and Okawara 2003). As such, this construction is attested in natural languages as well as in native and non-native acquisition. As already mentioned, *silent scope marking* is closely related to *wh- copying* and *wh- scope marking*, the difference being that no overt wh- word appears in the matrix clause. The learnability problem that this construction poses is explained under the account developed in this dissertation as follows: the production appears as it is a licensed UG option; it is also an LD wh- movement avoidance strategy, which provides further motivation for its occurrence; the limited overall rate of this type of production is explained by the fact that it is not part of the native grammar or the target input, and as such the probability that learners attach to it is relatively low (similar to *wh- scope marking* and *wh- copying*).

To return to the distribution of the avoidance strategies across the different individuals, apart from learners who use mostly one or two types of avoidance mechanisms, there are also ones who resort to a wider variety. These participants are, however, fewer (10, 12, 30). Particularly varied are the productions of participant 30: out of 19 elicited questions, 4 involve LD wh- movement, 2 are *medial wh-* without embedded inversion, 2 are *medial wh-* with embedded inversion, 3 are sequential questions without a wh- word in the second clause as illustrated in (37)*c*, 2 are simple wh- questions, 3 are complex yes/no questions, and 1 is a *silent scope marking* construction. The avoidance strategies used by participants 10 and 12 are also quite varied but they do show a preference for *medial wh-* (particularly the sequential representation). Note that this representation seems to be the most highly favoured avoidance strategy overall. This fact can be seen both in the group and in the individual results. I propose that this finding is due to a combination of several different factors. The *medial wh- indirect dependency* essentially represents a string of two consecutive simple wh- questions. Simple questions are widely attested in the input and usually form an integral part of explicit ESL/EFL instruction. Furthermore, from a processing perspective, such questions are easer to produce than LD wh- movement. Finally, a sequence of two wh- questions meets the pragmatic requirements of the elicitation task perfectly. That is, the first question "*What do you think?"* asks about the researcher's opinion, while the second question "*What is he drinking/Where is John/Who is John kissing, etc.*" asks about the activity, location, character, etc. involved in the particular elicitation context. Because of this combination of factors the *medial wh-* sequential representation is the most favoured avoidance strategy.

Before concluding the discussion of the individual results, figure 34 provides the distribution of LD wh- movement versus LD avoidance structures for each participant Note that the LD versus LD avoidance bars for each participant add up to 100% (i.e. they represent all questions produced by that participant).

Figure 34. Individual distribution of LD wh- movement versus avoidance structures. Oral elicitation task, Bulgarian speakers of L2 English, n=31
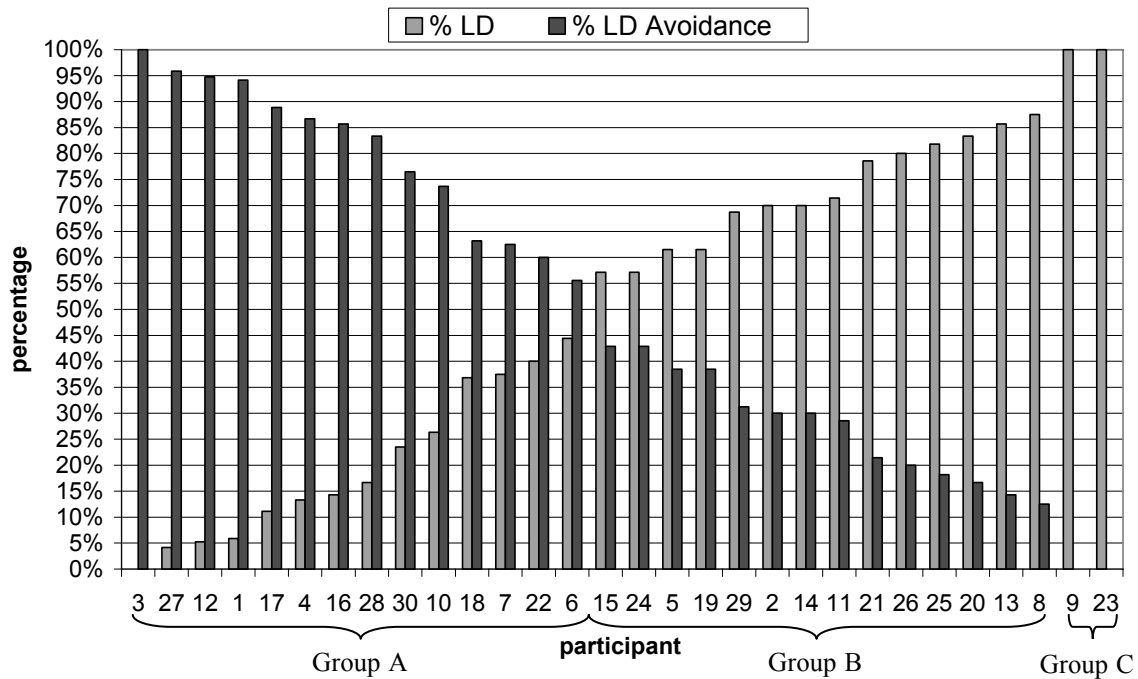
As with the French learners from *experiment 3*, three different groups can be discerned: *group A*, in which the learners use avoidance strategies more often than LD wh-movement; *group B*, in which the learners use LD movement more often than avoidance strategies; and *group C*, in which the learners show a facility for LD wh- movement and do not resort to avoidance structures at all. Overall, the data in this figure seems to reflect an acquisition process in which participants shift from exclusive use of avoidance strategies to full acquisition of the target LD wh- mechanism in the L2. As with the French participants, it should be noted that the EFL placement test for the Bulgarian learners of L2 English did not include an oral proficiency component. Thus, it is likely that even though all participants in *experiment 4* were placed at the (lower)-intermediate level, some of them had higher spoken proficiency than others. Adopting this assumption, the participants in *group A* have the lowest spoken proficiency level and as such LD wh-movement poses the highest degree of difficulty for them. To alleviate the processing pressure caused by the derivational complexity of target complex wh- questions, these

learners resort to avoidance strategies between 55% and 100% of the time. Group B is essentially a mirror image of group A as its participants use LD wh- movement between 57% and 87% of the time. As figure 34 indicates, the percentage of LD wh- productions steadily increases and six participants (21, 26, 25, 20, 13, 8) in this group use LD around 80% of the time. It can be assumed that these learners have acquired the LD wh-movement mechanism in the L2 sufficiently well and that processing pressures and derivational complexity have a limited effect on their productions. This is especially true with regards to participants 25, 20, 12 and 8, whose rates of LD avoidance are less than 20%. Finally, the two participants in *group C* use exclusively LD wh- movement in their productions and as such acquisition of the target structure has been achieved.

4.3.4.   Discussion and conclusion

Comparing the results of *experiment 3* and *experiment 4*, it is important to ask the question what these findings offer in terms of our understanding of the acquisition process. Although Bulgarian and French are two typologically distinct languages, the overall patterns of the respective two groups of learners of L2 English were essentially the same. The results of the Bulgarian group showed a lower rate of target LD wh-movement structures and a higher rate of avoidance strategies, but this can be attributed to the setting of acquisition and the possibility that this group was at an earlier stage of L2 development (recall that it was not possible to match the French and Bulgarian speakers for proficiency). Despite this, however, it seems that the two populations are following similar acquisition paths. Note that not only the type, but also the relative distribution of avoidance strategies was quite similar in the two populations, as indicated by the individual results.

The finding that speakers with distinct typological L1 backgrounds show similar acquisition patterns provides general support to the claim that UG plays a role in L2 acquisition. In addition to this, both populations resorted exclusively to structures licensed by the principles of UG. That is, no evidence of structures that fall outside of the UG hypothesis space were found. In some cases, the productions attested were options not supported by either the L1 or the L2 of the learners, but licensed options in other languages of which the participants reported no knowledge. The learnability problem

posed by this result finds a natural explanation if we assume that the learners, similarly to children acquiring their first language, rely on their abstract grammatical instincts.

The probabilistic learning component, which constitutes an integral part of the approach to L2 acquisition developed in this dissertation, comes into play in explaining the distribution of the different target and non-target productions attested in the two experiments. Essentially, with a wide array of UG supported options, a probabilistic device is important in the acquisition process as it analyses the input and determines which constructions should be upheld and which ones should be eliminated from the interlanguage grammar. This idea provides an explanation to the fact that the target LD wh- movement was among the most highly attested production in both populations. On the other hand, LD wh- movement avoidance strategies such as *medial wh-* (without embedded inversion) and *silent scope marking* were attested at lower rates because they are not supported by either the target input or the L1 base, and thus the probabilities attached to them would be much lower.

With regards to *medial wh-* productions, an interesting asymmetry was observed between a prevalence of the *direct dependency* in the written experiments (at least some groups) and a prevalence of the *indirect dependency* in the oral experiments. Recall that in chapter 3 I argued against a processing account of *medial wh-* constructions. The fundamental claim was that such productions seem to be grounded in the participant's abstract grammatical knowledge reflected in their judgements (choices) on the written grammaticality test. An explanation of the written data based strictly on processing was rejected because of, among other things, the relatively low occurrence of the *medial wh-* and embedded inversion option (i.e. the sequential question representation), which is the one that should be easier from a processing perspective. The fact that exactly this representation was preferred in the spoken productions, on the other hand, is a strong indication that a processing explanation should be pursued for *experiments 3* and *4*. In essence, since the *indirect medial wh- dependency* represents a less complex derivation than the *direct dependency*, it occurs more often in spoken productions where the processing resources of the learners are more limited. On the other hand, in written grammaticality judgement, many participants showed a stronger preference for the *direct dependency;* despite the fact that this structure is unattested in their L1s or the L2, it is

based on their abstract grammatical competence, represents a more efficient computation, and reflects a typical adult grammar preference for denser information packaging.

An important issue about grammatical competence and processing needs to be emphasized with regard to the oral experiments. The fact that I am arguing for a processing account of the spoken data does not mean that the learners' grammatical knowledge is lacking or impaired. Their competence is exhibited in their ability to produce target LD wh- movement structures, as well as other complex syntactic structures. At the same time, as proposed in this chapter, processing pressures limit the number of target LD wh- movement productions and cause the learners to resort to a variety of avoidance strategies that provide an overt intermediate "bridge", shorten, or eliminate the LD wh- dependency. In other words, the fact that both the target LD wh- constructions and the different avoidance strategies co-exist in the productions of the majority of the learners indicates that their grammatical competence in a sense engages in a "battle" with processing pressures. Consistent with the theoretical account adopted in this dissertation, as the language experience increases and processing resources free up, the grammatical competence will be able to take over and the productions will converge with the grammatically and contextually appropriate target.

Finally, the issue of whether producing derivationally less complex structures, as an alternative to LD wh- movement, really constitutes an avoidance mechanism that alleviates processing pressures was raised in section 4.2.5 above. Essentially, the idea was that these strategies could actually be alternative productions caused by some other factors, unrelated to LD wh- movement avoidance. This possibility, however, was rejected in *experiment 3* based on the fact that the native speakers did not produce such a variety of non-LD structures, the elicitation contexts provided a bias for complex wh- questions, and the elicitation follow-up prompts gave the participants strong indication as to what the required structure was. As additional support for this claim, *in experiment 3* I pointed out that some participants produced utterances in which hesitations, false starts, paraphrases etc. showed that LD wh- movement posed a high degree of difficulty and was thus replaced by alternative structures. Similar evidence is available in the data of

experiment 4 as well. Examples of such productions are provided in (42), where pauses, hesitations and paraphrases are shown.[33]

   (42) a. Who do you think uhm... who is going to (speaking BG: *kojto=comp.*) who is... (speaking BG: *kojto, ne znam kak e=I don't know how to say who/that*) who is going to buy …(eggs) eggs?   (participant BG 28)

        b. Where… where… where… ah… do you … where does she eat?   (BG 6)

        c. Where... where do you think that is ah… where… where is… ah… where is John at the moment?   (participant BG 19)

        d. What do you think uh… who… uh has uhm... kissing? (speaking BG*: dobre li go kazah=did I say it right?*) … What is John...Who is John kissing now?   (participant BG 4)

As the utterance in (42)*a* indicates, the participant begins an LD wh- question which looks like a *medial wh-* structure of the sequential type. Then the participant switches to Bulgarian and asks how to translate the Bulgarian complementizer *kojto* (who/that) into English; this shows evidence that the speaker is trying to compute a complex structure. Unable to produce the LD wh- question, however, the participant rephrases the utterance into a simple question. In (42)*b*, the participant pauses several times after *where* and it looks as if he intends to add *do you think*; however, the participant is unable to complete the question and thus resorts to a simple question. In (42)*c*, the fact that the participant attempts to generate an LD wh- question is even more easily discernable as the whole matrix clause including a complementizer is produced; however, after a pause midstream, the learner is unable to complete the utterance and resorts to a simple question. Finally, in (42)*d*, the participant attempts to produce an LD wh- question, but the second clause is interrupted by two pauses and is also missing a syntactic constituent (i.e. it is unclear whether the subject or object of *kiss* is missing); the participant shows doubt about this production (i.e. asking the researcher if it is correct) and then rephrases using a simple question (i.e. avoiding LD wh- movement). Overall, these examples provide additional support for the idea that the structures used as alternatives to LD wh- movement are indeed avoidance strategies and not productions that are attested for some other reason (see also Radford and Yokota 2006 for an independent argument that LD wh- movement poses a higher degree of difficulty than short wh- movement).

---

[33] Recall that a formal phonological or prosodic analysis of the utterances is beyond the scope of this dissertation and this point is used only as supplementary, impressionistic evidence.

This concludes the discussion of *experiments 3* and *4*. The following chapter offers an overview of the findings from both the written and spoken data, and some final remarks.

Chapter V
Conclusion

- ▪ SUMMARY: THIS CHAPTER CONTAINS AN OVERVIEW OF THE FINDINGS OF THE DISSERTATION AND PROVIDES A FINAL DISCUSSION OF THE L2 ACQUISITION ISSUES THAT THEY RAISE.

The main phenomenon that this dissertation set out to investigate is *medial wh-* constructions (both *wh- copying* and *wh- scope marking*) in L2 English. What makes the study of these structures worthwhile in the context of second language acquisition is their potential to provide insights with regards to issues such as learnability, variability, access to UG and the link between grammatical complexity and processing. Since *medial wh-* has only recently started to receive attention in the L2 literature, the data on this phenomenon is limited; thus, more studies, especially involving learners with different L1 backgrounds, are needed.[1] This dissertation contributes to this general research goal by offering data from two new populations of English L2 learners with typologically distinct L1 backgrounds: (Canadian) French and Bulgarian.

Medial wh- constructions were attested in the written experiments with the two learner populations (chapter 3). Although their rates were relatively low (e.g. average of 15% for *choices C* and *D* at the (lower)-intermediate level in group F2-3), they cannot be discounted as noise. *Medial wh-* was attested at the lower levels of proficiency while the advanced participants had essentially eliminated this representation from the interlanguage grammar. It is important to point out that the learners showed variability with regards to *medial wh-* acceptance rates. Not all learners at the lower levels of acquisition accepted such constructions and, among those who did, the rates differed. Variability raises important issues in L2 acquisition as it can be argued, among other things, that L2 learners are more prone to it than children acquiring their L1, and that L2 learners have more difficulty overcoming variability and converging with the target grammar (for discussion see Sorace 2000 and references therein). The data discussed in the written experiments of this dissertation, however, do not support this idea because the variability related to *medial wh-* phenomena was constrained in several important ways. The results for both the French and the Bulgarian speaker populations showed that the

---

[1] The study of *medial wh-* in second language acquisition should eventually also be extended to L2s other than English.

total *medial wh-* rate (choices *C* and *D*) never exceeded 26% at the group level and very rarely exceeded 50% in the responses of a particular individual. Apart from *medial wh-*, the variability of acceptance of the target complex wh- construction involving long-distance movement and deletion of all intermediate copies was constrained as well. At all stages of acquisition, these utterances were highly attested in the interlanguage grammar (between 60% and 93%) and always "won" convincingly in the competition with the non-target *medial wh-* constructions. Another interesting point that the written data show with regards to variability is that while learners may go through a *medial wh-* stage in early L2 acquisition, participants at the advanced stages show that the target representation is fully acquired. Thus, it seems that while different learners may take different routes, acquisition of the target construction is ultimately ensured.

Overall, these data show that the variability attested with regards to *medial wh-* constructions both within and across participants is constrained and does not prevent convergence with the target grammar. Furthermore, similar variability with regards to *medial wh-* constructions was observed by Thornton (1990) in child English L1 acquisition. Recall that Thornton did not attest *medial wh-* constructions with all participants, that the individual rates varied in the cases when *medial wh-* was attested, and that convergence with the target was ensured. In this respect, a parallel can be drawn between the data presented in this dissertation and Thornton's (1990) L1 research.[2] To return to variability in L2 acquisition, it is important to acknowledge that the patterns found in this dissertation may not necessarily be attested with other L2 phenomena. Most notably, it has been shown that morphological errors and omissions in L2 are subject to higher variability and fossilization, which has caused debates as to whether L1 and L2 acquisition are fundamentally similar or fundamentally different (see White 2007 for a recent overview; see Zobl and Liceras 1994 for discussion of morpheme acquisition orders).

Apart from discussing *medial wh-* with regards to variability, the question of the nature of the dependency investigated in the written experiments of this dissertation had consequences with regards to learnability issues. The *direct dependency medial wh-*

---

[2] Thornton's (1990) results differed from the current data in other ways, however. For example, co-occurrence of *medial wh-* and a complementizer was not attested in her data.

structure that was reliably attested in subsets of participants in the French and the Bulgarian populations is unavailable in both the L2 and the two L1s of the learners. At the same time, such representations exist in the adult grammars of other typologically distinct languages (e.g. German, Afrikaans, Romani, Bahasa Indonesia, to name a few). In light of the lack of possibilities of relevant L1 transfer and L2 input to account for *direct dependency medial wh-* structures in the L2 English of the participants, two learnability questions arise: 1) Why do such representations emerge in the interlanguage grammar of the two populations? and 2) What causes these constructions to disappear at the advanced stages of acquisition? To explain these issues the account developed in this dissertation combines insights from nativist and domain-general views on language acquisition (based on Yang's 2002, 2004, 2008 *variational model* of L1 acquisition). On the one hand, the L2 learners must be able to draw from a pool of UG-constrained grammatical resources that might cause the emergence of interlanguage representations unattested in either the L1 or the L2 (i.e. hypothesis testing). At the same time, the learners must also be able to engage in complex statistical analysis of the target input, in order to assign probabilistic values to competing representations in the interlanguage grammar and to be able to eventually eliminate structures not supported by the L2 (based on indirect negative evidence). While nativist and domain-general views on language may be seen as incompatible, recent proposals from both cognitive science (e.g. Perfors, Tenenbaum and Regier, 2008) and linguistics (e.g. Chomsky 2005a) increasingly argue against the validity of this dichotomy. The question that arises from integrated approaches, such as the one proposed in this dissertation, is exactly how much to attribute to the innate and the domain-general components, respectively. I believe that this question offers new potential avenues to be explored by future L2 research.

In addition to the written experiments, this dissertation also examined spoken data from an elicited production task carried out with (Canadian) French and Bulgarian speaking learners of L2 English (chapter 4). The main finding in the two spoken experiments was that many L2 learners have difficulty producing questions with long-distance wh- movement and resort to a variety of alternative target and non-target constructions instead. These involved simple questions with local wh- movement, embedded questions, complex yes/no questions, different types of sequential wh-

questions (including *indirect dependency medial wh-*), biclausal structures with an *if* or an *about/of* second clause (*wh-…if…*; *wh-…about/of…*), and *medial wh-* (*direct dependency*). Since all of these productions involved short or no wh- movement, or "bridging" the long-distance dependency by means of an overtly pronounced wh- word in medial position, they were classified as strategies avoiding standard long-distance wh-movement with deletion of all intermediate copies.[3] Their existence was explained by a version of the Derivational Complexity Hypothesis (Jakubowicz and Strik 2008 and references therein) adopted in this dissertation. Essentially, all avoidance strategies shortened or eliminated the LD wh- dependency, and as such were derivationally less complex and posed lower processing load than the target LD wh- movement structure. Thus, the proposal developed in chapter 4 was that the difficulty of LD wh- movement resides in both its complex grammatical nature and in the processing load associated with it.

Interestingly, the spoken data showed no evidence of *medial wh-* in infinitival clauses or other productions that go beyond the hypothesis space licensed by Universal Grammar. That is, all productions were explainable in terms of either the L1 and/or the L2, or in terms of constructions attested in other typologically distinct languages. Since the *medial wh-* (*direct dependency*, and *silent scope marking*) belong to the latter category of constructions, their existence in some of the learners' productions was interpreted as evidence for access to UG. At the same time, their lower rates of occurrence in comparison with some of the other avoidance strategies was accounted for by probabilistic learning. That is, since these structures are not part of the L1 and are not attested in the input, their rate of occurrence is expected to be lower than the rates of other avoidance strategies that involve structures available in the L1 and/or the L2. As such, the approach developed to account for the written experiments was able to explain and integrate the results of the oral production tasks.

This dissertation also raises the issue of the nature of the relationship between grammatical competence and processing. In chapter 3, it was argued that a processing explanation alone is not sufficient to account for the *medial wh-* constructions in the two

---

[3] Recall that in the context of this dissertation the term *avoidance strategy* is not meant to imply intentional avoidance of LD wh- movement but rather spontaneous or natural use of an alternative structure, which is less complex from a derivational and processing point of view.

written experiments. The argument was based on several different reasons but most notably on the fact that some learners (a subgroup in group F2-3) showed a strong preference for the *direct dependency medial wh-* in comparison with the *indirect dependency* (11% vs. 3%, respectively). Since the former represents long-distance wh-movement with a wh- word pronounced in an intermediate position and the latter is equivalent to sequential questions with independent local movements, the easier representation from a processing perspective would be the sequential one. Thus, it was proposed that the 11% *direct dependency* responses cannot be the result of processing alone and must be grounded in the learners' grammatical competence. In other words, such constructions must be part of what the learners at this level of acquisition hypothesise to be a representation allowed by the target grammar. In chapter 4, however, I argued that *medial wh-* representations were part of a larger pool of strategies that alleviate the difficulty posed by the long-distance wh- dependency. These strategies were proposed to result from derivational complexity and processing constraints. Thus, chapters 3 and 4 corroborate in showing that *medial wh-* constructions should receive an explanation in terms of both the learners' grammatical competence and processing profile. The picture that emerges from such data suggests that there may be a tight link between the L2 learner's interlanguage grammar and processing. This idea is consistent with Phillips' (1996) proposal that the dichotomy between grammar and processing should be dispensed with and that the parser is the same as the grammar and *vice versa*. Whether such a position can be adopted with regards to L2 acquisition is an interesting avenue left open to further research.

One final issue that needs to be revisited is the relationship between the data from the (Canadian) French and the Bulgarian speakers. Despite the fact that these two L1s share equivalent LD wh- movement mechanisms, they are different in a number of ways. For example, French has a wh- *in situ* option in addition to wh- movement, while Bulgarian does not. On the other hand, Bulgarian has multiple wh- fronting, while French does not. Despite these, and a number of other typological differences between French and Bulgarian, the L2 English results from the two populations were essentially equivalent. Similar rates and types of *medial wh-* were accepted in the written experiments (chapter 3) and both populations showed similar sensitivity to co-occurrence

of *medial wh-* with an overt complementizer and to negation in the matrix clause. Furthermore, the target LD wh- movement structure was properly acquired at the advanced stages of L2 development in both populations. With regards to the spoken experiments (chapter 4), both the French and the Bulgarian participants showed evidence of target LD wh- movement structures competing with LD avoidance strategies. The variation and types of strategies were strikingly similar. One difference was that in the Bulgarian population the rates of occurrence of *medial wh-* were higher and accuracy with the target LD wh- structures was lower than in the French population. This was explained by the EFL setting of acquisition, which gives the Bulgarian speakers more limited exposure to the TL; furthermore, since it was not possible to match the Bulgarian and the French speaking participants for proficiency, it could be that the Bulgarians were at a somewhat lower level of L2 development despite the label (lower)-intermediate that was assigned to them. Overall, it is reasonable to assume that the Bulgarian group had a lower level of proficiency, which explains their lower accuracy and higher *medial wh-* rates. Despite the differences in *medial wh-* and target LD rates, however, the patterns of acquisition remained essentially equivalent, and both populations could be grouped into participants who use predominantly avoidance strategies, participants who use predominantly LD wh- movement, and participants who use a mixture of both. This was interpreted as evidence that the L2 learners from both populations follow an acquisition path in which avoidance strategies are gradually replaced by LD wh- movement.

Before concluding this dissertation, I would like to point out what I consider to be its main contribution, and at the same time to acknowledge some of the limitations of the experiments. To start with the former, the dissertation was able to attest *medial wh- direct dependency* representations in the interlanguage of both the French and the Bulgarian speaking populations. Such data strengthen the evidence that UG plays an important role in L2 acquisition. The overall prevalence of and the ultimate convergence with the target LD wh- movement structure is another piece of evidence supporting the UG-based account that was adopted. The fact that the learners showed sensitivity to argument-adjunct asymmetries, an overt complementizer, and matrix negation indicates that L2 development is constrained at a deep level by abstract grammatical properties and that learner interlanguages are similar to natural languages. With regards to the spoken data,

the fact that LD wh- movement was found to pose difficulty for the learners and that it was replaced by avoidance structures, which some L2 learners were ultimately able to overcome, largely corroborates the overall approach developed in the dissertation.

As for some of the limitations, it was not possible to administer proficiency measures designed specifically for the purposes of this study and this did not allow for the level of the French and the Bulgarian speaking groups to be matched. Using such measures would have allowed for more reliable comparisons between the two populations. Furthermore, the fact that it was not possible to recruit an intermediate Bulgarian group in *experiment 2* leaves open the question whether learners at that level would show preference for the *direct dependency* in parallel to the equivalent French group. Larger sample populations for the spoken experiments would have increased the chance of attesting higher rates of *medial wh-* in production and of gathering further insights into its properties in L2 English. Psycholinguistic experimental techniques would have been able to shed more light on some of the issues with regards to processing and its interrelation with grammar and on the possible effects that individual working memory capacities may have on *medial wh-* productions. Finally, the account adopting access to UG in combination with a probabilistic learning mechanism provides a sound conceptual explanation of the data. However, a precise implementation of this account in the form of an L2 acquisition computational model remains beyond the scope and resources of the dissertation. These, as well as any other goals and issues that may remain outstanding, I leave for future work.

References

Alexiadou, A. and E. Anagnostopoulou (1998) "Parameterizing Agr: word order, V-movement and EPP-checking." *Natural Language and Linguistic Theory* 16, 491–539.

Aslin, R. N., J. R. Saffran and E. L. Newport (1998) "Computation of conditional probability statistics by 8-month-old infants." *Psychological Science* 9, 321–324.

Bates, E. and B. MacWhinney (1989) "Functionalism and the Competition Model." In E. Bates and B. MacWhinney (eds.), *The Crosslinguistic Study of Sentence Processing*, 3–73. Cambridge: Cambridge University Press.

Beck, S. (1996) "Quantified structures as barriers for LF movement." *Natural Language Semantics* 4, 1–56.

Beck, S. and S. Berman (2000) "*Wh-* Scope marking: Direct vs. indirect dependency." In U. Lutz, G. Müller and A. von Stechow (eds.), *WH- Scope Marking*, 17–44. Amsterdam: John Benjamins.

Beck, S. and S.-S. Kim (1997) "On Wh- and Operator Scope in Korean." *Journal of East Asian Linguistics* 6, 339–384.

Belikova, A. (2008) "Reciprocal Verbs in the L2 French of Native Speakers of Russian." In Roumyana Slabakova et al. (eds.), *Proceedings of the 9th Generative Approaches to Second Language Acquisition Conference (GASLA 2007)*, 1–14. Somerville, MA: Cascadilla Proceedings Project.

Bennet, S. (1994) "Interpretation of English reflexives by adolescent speakers of Serbo-Croatian." *Second Language Research* 10, 125–56.

Berwick, R. (1985). *The Acquisition of Syntactic Knowledge*. Cambridge, MA: MIT Press.

Berwick, R. and A. Weinberg (1984) *Grammars and models of language use: The grammatical basis of linguistic performance*. Cambridge, MA: MIT Press.

Bley-Vroman, R. (1990) "The logical problem of foreign language learning." *Linguistic Analysis* 20, 3–49.

Bley-Vroman, R., S. Felix, and G. Ioup (1988) "The accessibility of universal grammar in adult language learning." *Second Language Research* 4, 1–32.

Bošković, Ž. (1998) "LF movement and the Minimalist Program." In *Proceedings of the North East Linguistic Society* 28, 43–57. GLSA, University of Massachusetts, Amherst.

Brody, M. (1995) *Lexico-Logical form: a radically minimalist theory*. Cambridge, MA: MIT Press.

Bruening, B. (2006) "Differences between the *Wh-* Scope-Marking and *Wh*-Copy Constructions in Passamaquoddy." *Linguistic Inquiry* 37, 25–49.

Cheng, L. (1990) *On the typology of wh-questions*. Doctoral dissertation, MIT.

Cheng, L. (2000) "Moving just the feature." In U. Lutz, G. Müller and A. von Stechow (eds.). *Wh-scope marking*, 77-99. Amsterdam: John Benjamins.

Cheng, L. and J. Rooryck (2000) "Licensing *Wh*-in-situ." *Syntax* 3, 1–19.

Chomsky, N. (1957) *Syntactic structures*. The Hague: Mouton.

Chomsky, N. (1965) *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

Chomsky, N. (1981) *Lectures on government and binding*. Dordrecht: Foris.

Chomsky, N. (1986a) *Barriers*. Cambridge, MA: MIT Press.

Chomsky, N. (1986b) *Knowledge of Language: Its Nature Origin and Use*. New York: Praeger.

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

Chomsky, N. (2000) "Minimalist Inquiries: the Framework." In: R. Martin, D. Michaels and J. Uriagereka (eds.), *Step by Step: Essays in Minimalist Syntax in Honour of Howard Lasnik*, 89–155. Cambridge, MA: MIT Press.

Chomsky, N. (2001) "Derivation by phase." In M. Kenstowicz (ed.), *Ken Hale, a life in language*, 1–52. Cambridge, MA: MIT Press.

Chomsky, N. (2005a) "Three factors in language design." *Linguistic Inquiry* 36 (1), 1–22.

Chomsky, N. (2005b) On phases. Ms., MIT. Cambridge, MA.

Clahsen, H. (1984) "The acquisition of German word order: A test case for cognitive approaches to L2 development." In R. Andersen (ed.), *Second Languages: A Cross-Linguistic Perspective*, 219–242. Rowley, MA: Newbury House Publishers.

Clahsen, H. (1990/91) "Constraints on parameter setting: a grammatical analysis of some acquisition stages in German child language." *Language Acquisition* 1: 361–391.

Clahsen, H. and C. Felser (2006) "Grammatical processing in language learners." *Applied Psycholinguistics* 27: 3–42.

Clahsen, H. and M. Penke (1992) "The acquisition of agreement morphology and its syntactic consequences: new evidence on German child language from the Simone-corpus". In J. Meisel (ed.), *The acquisition of verb placement*, 181–224. Dodrecht: Kluwer.

Clahsen, H. and P. Muysken (1986) "The availability of UG to adult and child learners: A study of the acquisition of German word order." *Second Language Research* 2, 93–119.

Cole, P. and G. Hermon (1994) "Is there LF movement?" *Linguistic Inquiry* 25 (2), 239–262.

Cole, P. and G. Hermon (2000) "Partial wh-movement: evidence from Malay." In U. Lutz, G. Müller and A. von Stechow (eds.), *Wh-scope marking*, 101–130.

Crain, S. and R. Thornton (1998) *Investigations in Universal Grammar. A guide to experiments on the acquisition of syntax and semantics*. Cambridge, MA: MIT Press.

Croft, W. (1990) *Typology and Universals*. Cambridge: Cambridge University Press.

Dayal, V. (1994) "Scope marking as indirect wh- dependency." *Natural language semantics* 2, 137–170.

Dayal, V. (1996) *Locality in Wh- quantification: questions and relative clauses in Hindi*. Dordrecht: Kluwer.

Dayal, V. (2000) "Scope marking: cross-linguistic variation in indirect dependency." In U. Lutz, G. Müller and A. von Stechow (eds.), *Wh-scope marking*, 157–193. Amsterdam: John Benjamins.

De Villiers, J., T. Roeper and A. Vainikka. (1990) "The acquisition of long-distance rules." In L. Frazier and J. De Villiers (eds.), *Language Processing and Language Acquisition*, 257–297. Dordrecht: Kluwer.

Du Plessis, H. (1977) "Wh- movement in Afrikaans." *Linguistic Inquiry* 8, 723–726.

Duffield, N. (2006) "How do you like your doughnuts?" *Applied Psycholinguistics* 27, 56–59.

Eckman, F. R. (2004) "Optimality Theory, markedness and second language syntax: the case of resumptive pronouns in relative clauses." *Studies in Phonetics, Phonology and Morphology* 10 (1), 89–110.

Epstein, S.M., S. Flynn and G. Martohardjono (1996) "Second language acquisition: theoretical and experimental issues in contemporary research." *Behavioural and Brain Sciences* 19, 677–758.

Fanselow, G. and A. Mahajan (2000) "Towards a minimalist theory of wh-expletives, wh-copying, and successive cyclicity." In U. Lutz, G. Müller and A. von Stechow (eds.), *Wh-Scope Marking*, 195–230. Amsterdam: John Benjamins.

Felser, C. (2004) "Wh-copying, phases, and successive cyclicity." *Lingua* 114, 543–574.

Finer, D. and E. Broselow (1986) "Second language acquisition of reflexive-binding." In S. Berman, J.-W. Choe and J. McDonough (eds.), *Proceedings of NELS 16,* 154–68. University of Massachusetts at Amherst: Graduate Linguistics Students Association.

Flynn, S. (1987) *A parameter setting model of L2 acquisition: experimental studies in anaphora*. Dordrecht: Reidel.

Freedman, S. and K. Forster (1985) "The psychological status of overgenerated sentences." *Cognition* 19, 101-31.

Friedmann, N, A. Belletti and L. Rizzi (in press) "Relativized relatives: Types of intervention in the acquisition of A-bar dependencies." *Lingua.*

Gibson, E. and K. Wexler (1994) "Triggers." *Linguistic Inquiry* 25, 355–407.

Gregg, K. R. (2003) "The state of emergentism in second language acquisition." *Second Language Research* 19(2), 92-128.

Guasti, M. T. (1993/94) "Verb syntax in Italian child grammar: Finite and nonfinite verbs." *Language Acquisition* 3, 1-40.

Gutierrez, M. J. (2004a) "*Non-adult questions in the acquisition of L1 Spanish long-distance wh-questions: a longitudinal investigation.*" Paper presented at the Romance Turn Workshop on the acquisition of Romance Languages, Madrid, Spain.

Gutierrez, M. J. (2004b) "*The acquisition of L1 Basque Long Distance Wh-questions: a longitudinal investigation.*" Paper presented at BIDE04, Bilboa, Spain.

Gutierrez, M. J. (2005) *The acquisition of English LD wh-questions by Basque/Spanish bilingual subjects in a school context*. Unpublished PhD dissertation, University of the Basque Country.

Gutierrez, M. J. (2006) "Acquiring long-distance wh-questions in L1 Spanish: A longitudinal investigation." In V. Torrens and L. Escobar (eds.), *The Acquisition of Syntax in Romance Languages*, 251-287. Amsterdam: John Benjamins.

Hamann, C. (2006) "Speculations about early syntax: The production of *wh*-questions by normally developing French children and French children with SLI." *Catalan Journal of Linguistics* 5, 143-189.

Haznedar, B. (2001) "The acquisition of the IP system in child L2 English." *Studies in Second Language Acquisition* 23, 1–39.

Haznedar, B. and B. D. Schwartz (1997) "Are there optional infinitives in child L2 acquisition?" In E. Hughes, M. Hughes, & A. Greenhill (eds.), *Proceedings of the 21st Annual Boston University Conference on Language Development*, 257–268. Somerville, MA: Cascadilla Press.

Herschensohn, J. (2000) *The second time around: minimalism and L2 acquisition*. Amsterdam: John Benjamins.

Herschensohn, J. (2001) "Missing inflection in L2 French: Accidental infinitives and other verbal deficits." *Second Language Research* 17, 273–305.

Herschensohn, J. (2004) "Inflection, thematic roles and abstract Case." *Bilingualism: Language and Cognition* 7(1), 32–34.

Hiemstra, I. (1986) "Some aspects of wh-questions in Frisian." *North-Western European Language Evolution (NOWELE)* 8, 97–110.

Höhle, T. (2000) "The w. . .w construction: appositive or scope indicating?" In U. Lutz, G. Müller and A. von Stechow (eds.), *Wh-Scope Marking*, 249–270. Amsterdam: John Benjamins.

Horvath, J. (1998) "Multiple Wh-Phrases and the Wh-Scope-Marker Strategy in Hungarian Interrogatives." *Acta Linguistica Hungarica* 45, 31–60.

Horvath, J. (2000) "On the Syntax of 'Wh-Scope Marker' Constructions: Some Comparative Evidence." In U. Lutz, G. Müller and A. von Stechow (eds.), *Wh-scope marking*, 271–316. Amsterdam: John Benjamins.

Huang, C.-T J. (1982) *Logical Relations in Chinese and the Theory of Grammar*. Ph.D. dissertation. MIT.

Huang, C.-T. J. (1984) "On the distribution and reference of empty pronouns." *Linguistic Inquiry* 15, 531–74.

Hulk, A. and S. Zuckerman (2000) "The interaction between input and economy: Acquiring optionality in French *wh*-Questions." In S. C. Howell et al. (eds.), *BUCLD 24 Proceedings*, 438–449. Somerville, MA: Cascadilla Press.

Hyltenstam, K. (1984) "The use of typological markedness conditions as predictors in second language acquisition: The case of pronominal copies in relative clauses." In R. Andersen (ed.) *Second Languages: A cross-linguistic perspective*, 39–58. Rowley, MA: Newbury House Publishers.

Ionin, T. and K. Wexler (2002) "Why is 'is' easier than '-s'?: Acquisition of tense/agreement morphology by child second language learners of English." *Second Language Research* 18, 95–136.

Jaeggli, O. and K. Safir (1989) "The null subject parameter and parametric theory." In O. Jaeggli and K. Safir (eds.), *The null subject parameter*, 1–44. Dordrecht: Kluwer.

Jakubowicz, C. (2002) "Functional categories in (Ab)Normal language acquisition." In I. Lasser (ed.), *The process of language acquisition*, 165–202. Berlin: Peter Lang Verlag.

Jakubowicz, C. (2003) "Hypothèses psycholinguistiques sur la nature du déficit dysphasique." In Y. Brun and Ch.L. Gérard (eds.), *Les dysphasies*, 23–70. Paris: Masson.

Jakubowicz, C. (2004) "Is movement costly?" Presented at the JEL conference (Journée d'Etudes Linguistiques), Nantes.

Jakubowicz, C. (2005) "The language faculty: (Ab)normal development and interface constraints." Presented at GALA 2005, Siena.

Jakubowicz, C. and N. Strik (2008) "Scope-marking Strategies in the Acquisition of Long Distance *wh*-Questions in French and Dutch." *Language and Speech* 51, 101–132.

Jakubowicz, C., and Nash, L. (2001) "Functional categories and syntactic operations in (Ab) Normal language acquisition." In *Brain and Language* 77, 321–333.

Keenan, E. and B. Comrie (1977) "Noun phrase Accessibility Hierarchy and Universal Grammar." *Linguistic Inquiry* 8, 63–99.

Lardiere, D. (1998a) "Case and tense in the 'fossilized' steady state." *Second Language Research* 14, 1–26.

Lardiere, D. (1998b) "Dissociating syntax from morphology in a divergent end-state grammar." *Second Language Research* 14, 359–375.

Lardiere, D. (2000) "Mapping features to forms in second language acquisition." In J. Archibald (ed.), *Second language acquisition and linguistic theory*, 102–129. Oxford, UK: Blackwell.

Lehmann, C. (1986) "On the typology of relative clauses." *Linguistics* 24, 663–680.

Liceras, J. M. and L. Diáz (2001) "Triggers in L2 acquisition: The case of Spanish N-N compounds." *Studia Linguistica* 54 (2), 197–211.

Liceras, J., S. Perales and A. Bel (2007) "Modality, Non-Finite Forms and the Manifestation of the RI Stage in Null and Non-Null Subject Languages." In S. Baauw, J. van Kampen and M. Pinto (eds.), *The Acquisition of Romance Languages: Selected papers from The Romance Turn II*, 97–111). Utrecht: LOT.

Lutz, U., G. Müller and A. von Stechow (eds.) (2000) *Wh-scope marking*. Amsterdam: John Benjamins.

MacLaughlin, D. (1996) "Second language acquisition of English reflexives: is there hope beyond transfer?" In A. Stringfellow, D. Cahana-Amitay, E. Hughes and A. Zukowski (eds.), *Proceedings of the 20th Annual Boston University Conference on Language Development*, 453–64. Somerville, MA: Cascadilla Press.

MacLaughlin, D. (1998) "The acquisition of morphosyntax of English reflexives by non-native speakers." In M-L. Beck (ed.), *Morphology and its interfaces in second language knowledge*, 195–226. Amsterdam: John Benjamins.

MacWhinney, B. (1987) "The Competition Model." In B. MacWhinney (Ed.), *Mechanisms of language acquisition*, 249–308. Hillsdale, NJ: Lawrence Erlbaum.

MacWhinney, B. (2000) "Emergence from what: Comments on Sabbagh & Gelman." *Journal of Child Language* 27, 727–33.

MacWhinney, B. (2001) "The Competition Model: The input, the context, and the brain." In P. Robinson (ed.), *Cognition and Second Language Instruction*, 61-90. New York: Cambridge University Press.

MacWhinney, B. (2007) "A Unified Model." In N. Ellis and P. Robinson (eds.), *Handbook of cognitive linguistics and second language acquisition*. Hillsdale, NJ: Lawrence Erlbaum.

MacWhinney, B. and E. Bates (eds.) (1989) *The crosslinguistic study of sentence processing*. New York: Cambridge University Press.

Mahajan, A. (1990) *The A/A-bar distinction and movement theory*. Doctoral dissertation. MIT.

Marácz, L. (1990) *Asymmetries in Hungarian*. Doctoral dissertation. University of Groningen.

Marantz, A. (2005) "Generative linguistics within the cognitive neuroscience of language." *The Linguistic Review* 22, 429–445.

Mathieu, E. (1999) "French *wh*-in situ and the intervention effect." *UCL Working Papers in Linguistics* 11, 441–472.

McCloskey, J. (2006**)** "Questions and Questioning in a Local English." In R. Zanuttini, H. Campos, E. Herburger and P. H. Portner (eds.), *Crosslinguistic Research in Syntax and Semantics: Negation, Tense, and Clausal Architecture*, 87–126. Washington, DC: Georgetown University Press.

McDaniel, D. (1986) *Conditions on Wh-Chins*. Doctoral dissertation. City University of NewYork, New York.

McDaniel, D. (1989) "Partial and Multiple Wh- movement." *Natural Language and Linguistic Theory* 7, 565–604.

McDaniel, D., B. Chiu and T. Maxfield (1995) "Parameters for wh-movement types: evidence from child English." *Natural Language and Linguistic Theory* 13, 709–753.

Müller, G. (1997) "Partial wh-movement and Optimality Theory." *The linguistic review* 14, 249–306.

Muriungi, P. K. (2004) "*Wh*-movement in Kitharaka as focus movement." *Proceedings of the Workshop on the Syntax, Semantics and Pragmatics of Questions,* ESSLLI 16.

Mycock, L. (2006) *The Typology of Constituent Questions: a Lexical-Functional Grammar analysis of 'wh'-questions*. Doctoral Dissertation. University of Manchester.

Newmeyer, F. (2003) "'Interpretable' features and feature-driven A-bar movement." In C. Beyssade, O. Bonami, P. Cabredo-Hofherr and F. Corblin (eds.), *Empirical issues in syntax and semantics*, 255-271. Paris: Presses Universitaires de Paris Sorbonne.

Nowak, M.A. et al. (2001) "Evolution of universal grammar." *Science* 291, 114–118.

O'Grady, W. (2001a) "Language acquisition and language loss." University of Hawaii manuscript.

O'Grady, W. (2001b) "An emergentist approach to syntax." University of Hawaii manuscript.

O'Grady, W. (2008) "The emergentist program." *Lingua* 118 (4), 447–64.

Oiry, M. (2002) *Acquisition des questions à longue distance (On the acquisition of long-distance Questions)*. Mémoire de Maîtrise. Université de Nantes, France.

Oiry, M. and H. Demirdache (2006) "Evidence from L1 acquisition for the syntax of *wh*-scope marking in French." In V. Torrens and L. Escobar (eds.), *The Acquisition of Syntax in Romance Languages*, 289–315. Amsterdam: John Benjamins.

Okawara, I. (2000) *On the acquisition of long-distance wh-questions by Japanese learners of English*. BA thesis. Department of English Literature, Gunma Prefectural Women's University.

Park, H. (2004) "A minimalist approach to null subjects and objects in second language acquisition" *Second Language Research* 20 (1), 1–32.

Perfors, A., J. Tenenbaum and T. Regier (2008) "The learnability of abstract syntactic principles." Ms. MIT and University of Chicago.

Pesetsky, D. (1987) "Wh-in situ: movement and unselective binding." In E. Reuland and A. ter Meulen (eds.), *The representation of (In)definiteness*, 98–129. Cambridge, MA: MIT Press.

Pesetsky, D. (2000) *Phrasal Movement and Its Kin*. Cambridge, MA: MIT Press.

Phillips, C. (1996) *Order and Structure*. Ph.D. dissertation. MIT.

Prévost, P. & White, L. (2000a) "Accounting for morphological variation in L2 acquisition: Truncation or missing inflection?" In M.-A. Friedemann and L. Rizzi (eds.), *The acquisition of syntax*, 202–235. London: Longman.

Prévost, P., & White, L. (2000b) "Missing surface inflection or impairment in second language acquisition? Evidence from tense and agreement." *Second Language Research 16*, 103–133.

Progovac, L. (1992) "Relativized subject: long distance reflexives without movement". *Linguistic Inquiry* 23, 671–680.

Progovac, L. (1993) "Long-distance reflexives: movement-to-Infl vs. relativized subject." *Linguistic Inquiry* 24, 755–72.

Progovac, L. and P. Connell (1991) "Long-distance reflexives, Agr-subjects, and acquisition." Paper presented at the Formal Linguistics Society of Mid-America, University of Michigan, Ann Arbor.

Radford, A., and H. Yokota (2006) "UG-constrained wh- movement in Japanese learners' English questions." *Second Language: The Japan Second Language Association* 5, 61–94.

Reinhart, T. (1998) "*Wh*-In-Situ in the framework of the Minimalist Program." *Natural Language Semantics* 6, 29–56.

Reis, M. (2000) "On the parenthetical features of German was. . .w constructions and how to account for them." In U. Lutz, G. Müller and A. von Stechow (eds.), *Wh-Scope Marking*, 359–407. Amsterdam: John Benjamins.

Rizzi, L. (1986) "Null objects in Italian and the theory of pro." *Linguistic Inquiry* 17, 501–57.

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

Rizzi, L. (1997) "The fine structure of the left periphery." In L. Haegeman (ed.), *Elements of Grammar*, 281–337. Dordrecht: Kluwer.

Rizzi, L. (2000) "Remarks on early null subjects." In M. A. Friedemann and L. Rizzi (eds.), *The acquisition of syntax*, 269–292. London: Longman.

Rizzi, L. (2001) "Reconstruction, Weak Island Sensitivity, and Agreement." In C. Cechetto, G. Chierchia and M. T. Guasti (eds.), *Semantic Interfaces*, 145–176. Stanford: CSLI.

Rizzi, L. and U. Shlonsky (2005) "Strategies of Subject Extraction." Ms. Universitiy of Siena and University of Geneva.

Roeper, T. (2000) "Universal Bilingualism." *Bilingualism: Language and Cognition*, 2, 169–186.

Rooryck, J. (1988) "Relative clauses and binding in French: The solution of Bouchard (1984)." *Lingvisticoe Investigationes* 12 (1), 151–155.

Rudin, C. (1998) "On multiple questions and multiple Wh-fronting." *Natural Language and Linguistic Theory* 6, 445–502.

Sabel, J. (2000a) "Expletives as features." In R. Billerey and B. Lillehaugen (eds.) *Proceedings of the West Coast Conference on Formal Linguistics* 19, 441–424. Somerville, MA: Cascadilla.

Sabel, J. (2000b) "Partial wh-movement and the typology of wh-questions." In U. Lutz, G. Müller and A. von Stechow (eds.), *Wh-scope marking*, 409–446. Amsterdam: John Benjamins.

Saddy, D.(1991) "*Wh-* scope mechanisms in Bahasa Indonesia." *MIT Working Papers in Linguistics* 15, 183–218.

Saffran, J. R., R. N. Aslin and E. L. Newport (1996) "Statistical learning by 8-month old infants." *Science* 274, 1926–1928.

Schachter, J. (1988) "Second language acquisition and its relationship to Universal Grammar." *Applied Linguistics* 9, 219–235.

Schippers, A. (2006a) *Variation in the CP domain: Long-distance dependencies in the Dutch dialects*. Ms., Rijksuniversiteit Groningen.

Schippers, A. (2006b) *Wie wat or waar*. Ms., Rijksuniversiteit Groningen.

Schulz, B. (2006) *Wh-scope marking in English interlanguage grammars: Transfer and processing effects on the second language acquisition of complex wh- questions* Ph.D. dissertation. University of Hawaii.

Schwartz, B. D. and R. Sprouse (1994) "Word order and nominative case in nonnative language acquisition: a longitudinal study of (L1 Turkish) German interlanguage." In T. Hoekstra and B. D. Schwartz (eds.), *Language acquisition studies in generative grammar*, 317–68). Amsterdam: John Benjamins.

Schwartz, B. D. and R. Sprouse (1996) "L2 cognitive states and the full transfer/full access model." *Second Language Research* 12, 40–72.

Soares, C. (2003) "The C-domain and the acquisition of European Portuguese: The case of *wh*-questions." *Probus* 15 (1), 147–176.

Soares, C. (2006) *La syntaxe de la périphérie gauche en portugais européen et son acquisition*. Thèse de doctorat, Université Paris 8.

Sorace, A. (2000) "Syntactic optionality in non-native grammars." *Second Language Research* 16 (2), 93–102.

Stepanov, A. (2000) "Wh- scope marking in Slavic." *Studia Linguistica* 54 (1), 1–40.

Strik, N. and C. Jakubowicz (2005) "The acquisition of scope-marking strategies in French compared to Dutch." Paper presented at the Romance Turn II Conference, Utrecht, Netherlands.

Tenenbaum J. and T. Griffiths (2001) "Generalization, similarity, and Bayesian inference." *Behavioral and Brain Sciences* 24**,** 629–640

Thornton, R. (1990) *Adventures in long-distance moving: The acquisition of complex wh-questions*. Doctoral dissertation. University of Connecticut, Storrs.

Truscott, J. and M. Sharwood Smith (2004) "Acquisition by processing: A modular perspective on language development." *Bilingualism: Language and Cognition* 7 (1), 1–20.

Vainikka, A. and M. Young-Scholten (1998) "Tree growth and morphosyntactic triggers in adult SLA." In M. L. Beck (ed.) *The L2 Acquisition of Morphology*, 89–113. Amsterdam: John Benjamins.

van Craenenbroeck, J. and A. Lipták (2008) "On the Interaction between Verb Movement and Ellipsis: New Evidence from Hungarian." In C. B. Chang and H.J. Haynie (eds.), *Proceedings of the 26th West Coast Conference on Formal Linguistics*, 138–146. Somerville, MA: Cascadilla Proceedings Project.

van Kampen, J. (1997) *First Steps in Wh-movement*. Wageningen: Ponsen and Looijen.

van Kampen, J. and A. Evers (1995) "How greedy is the child? Long distance movement in small steps." In M. Verrips and F. Wijnen (eds.), *ASCLD 3 Papers from the*

Dutch-German colloquium on Language Acquisition 1994, 91–108. Amsterdam: Instituut voor Algemene Taalwetenschap.

van Riemsdijk, H. (1982) "Correspondence effects and the Empty Category principle." *Tilburg Papers in Language and Literature* 12, University of Tilburg.

Wakabayashi, S. and I. Okawara (2003) "Japanese learner's errors on long distance wh-questions." In S. Wakabayashi (ed.), *Generative Approaches to the Acquisition of English by Native Speakers of Japanese*, 215–245. New York: Mouton de Gruyter.

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

White, L. (1989a) *Universal Grammar and second language acquisition*. Amsterdam: John Benjamins.

White, L. (1989b) "The principle of adjacency in second language acquisition: do L2 learners observe the subset principle?" In S. M. Gass and J. Schachter (eds.), *Linguistic Perspectives on Second Language Acquisition*, 134-158. Cambridge: Cambridge University Press.

White, L. (2003) *Second language acquisition and universal grammar*. Cambridge: Cambridge University Press.

White, L. (2008) "Some Puzzling Features of L2 Features" In J. M. Liceras, H. Zobl and H. Goodluck (eds.), *The Role of Formal Features in Second Language Acquisition*, 301–326. Mahawah, NJ: Lawrence Erlbaum.

Yamane, Maki (2003) *On the interaction of first-language transfer and universal grammar in adult second language acquisition: WH-movement in L1-Japanese and L2-English interlanguage*. Ph.D. dissertation. University of Connecticut.

Yang, C. (2002) *Knowledge and Learning in Natural Language*. Oxford: Oxford University Press.

Yang, C. (2004) "Universal Grammar, Statistics, or Both." *Trends in Cognitive Sciences* 8 (10), 451–456.

Yang, C. (2008) "Three Factors in Language Variation." Ms., University of Pennsylvania.

Zobl, H. and J. M. Liceras (1994) "Functional categories and acquisition orders." *Language Learning* 44, 159–80.

Zuckerman, S. (2001) *The acquisition of "optional" movement*. Groningen: University Library Groningen.

Zukowski, A. (2001) *Uncovering grammatical competence in children with Williams syndrome*. Ph.D. dissertation. Boston University.

Zukowski, A. (in press) "Elicited production of relative clauses reflects intact grammatical knowledge in Williams syndrome." *Language and Cognitive Processes*.