VINETA RŪTENBERGA

SINTAKTISKĀS KRITERIĀLĀS PAZĪMES ANGĻU UN FRANČU VALODAS RAKSTVEIDA SNIEGUMA VĒRTĒŠANĀ

SYNTACTIC CRITERIAL FEATURES IN ASSESSING WRITTEN PERFORMANCE IN ENGLISH AND FRENCH

Promocijas darbs

filoloģijas doktora grāda iegūšanai valodniecības zinātņu nozares

lietišķās valodniecības apakšnozarē

Zinātniskā vadītāja: *Dr.phil.* V.Kalnbērziņa

Rīga 2014

ACKNOWLEDGMENTS

ANOTĀCIJA

Promocijas darba „Sintaktiskās kriteriālās pazīmes angļu un franču valodas rakstveida snieguma vērtēšanā" teorētiskajā daļā tiek analizētas dažādas pieejas valodas vienību apguves izpētē, kā arī aktualizēti galvenie principi rakstveida snieguma vērtēšanā un rezultātu pielīdzināšanā Eiropas kopīgajām pamatnostādnēm. Pamatojoties uz mūsdienu deskriptīvajās gramatikās lietoto angļu un franču valodas sintaktisko struktūru kontrastīvās analīzes rezultātiem, promocijas darba empīriskajā daļā tiek veikta angļu un franču valodas apguvēju lietoto sintakstisko struktūru kvantitatīvā un kvalitatīvā analīze. Iegūtie dati tiek izmantoti, lai izstrādātu katram valodas apguves līmenim (A1 – C2) atbilstošas sintaktiskās kriteriālās pazīmes angļu un franču valodā, kas garantēs testu rezultātu drošumu.

Atslēgvārdi: valodas apguve, rakstveida snieguma vērtēšana, sintaktiskās struktūras, uz korpusa datiem balstīta analīze, kriteriālās pazīmes.

ABSTRACT

The theoretical part of the thesis 'Syntactic Criterial Features in Assessing Written Performance in English and French' contains the analysis of different approaches towards linguistic unit acquisition as well as highlights the main principles in assessing written performance in compliance with the Common European Framework. The quantitative and qualitative research of the syntactic structures produced by the English and French language learners is based upon the results of the contrastive analysis of the syntactic structures retrieved from the contemporary descriptive grammars. The obtained data are used to develop syntactic criterial features for each language proficiency level (A1 – C1) both in English and French, which will ensure the reliability of test results.

Key words: language acquisition, written performance assessment, syntactic structures, data analysis based on corpus, criterial features.

CONTENTS

List of Tables

List of Figures

List of Abbreviations

ADV	adverb
A/Adj/ADJ/JJ	adjective
AdjP/ADJP/AP	adjectival phrase
AdvP	adverb phrase
ASCII	American Standard Code for Information Interchange
BNC	British National Corpus
C	complement/ conjunction (in French)
CC	coordinating conjunction
CD	cardinal number
CEFR	Common European Framework of Reference
CL	clitic
CLA	communicative language ability
CLC	Cambridge Learner Corpus
CLS	complex sentence
conj	conjunction
COORD	coordinated phrase
CS	compound sentence
D/DT	determiner
EFL	English as a foreign language
ELWT	English Learner Written Text sub-corpus
EFLWTC	English and French Learner Written Text Corpus
EPP	English Profile Programme
EU	European Union
FFL	French as a foreign language
FLWT	French Learner Written Text sub-corpus
IN	preposition or subordinating conjunction
JJR	adjective, comparative
L1	first langugage
L2	second language
LAD	Language Acquisition Device
LGSWE	Longman Grammar of Spoken and Written English
LOB	Lancaster-Oslo-Bergen corpus
LSWE	Longman Spoken and Written English corpus
MD	modal
N	noun
NCE	National Centre for Education
NL	native language
NLP	Natural Language Processing/Natural Language Parser
NP	noun phrase
NN	noun, singular or mass
NNP	proper noun, singular
NNS	proper noun, plural
O	object
O_d	direct object
P	preposition
POS	part-of-speech
PP	prepositional phrase
PRP	personal pronoun

PSRs	phrase structure rules
PRP$	possessive pronoun
PUNC	punctuation
RB	adverb
RLD	Reference Level Descriptors
S	subject
S/SENT	sentence
SBAR	clausal complement
Sint/Srel/Ssub	finite clauses
SS	simple sentence
TEI	Text Encoding Initiative
TO	to
UG	Universal Grammar
V	verb
VB	verb, base form
VBD	verb, past tense
VBG	verb, gerund or present participle
VBN	verb, past participle
VBZ	verb, 3rd person singular present
VN	verbal nucleus
VP	verb phrase
VBP	verb, non-3rd person singular present
VPinf	infinitive clause
WHADVP	wh-adverb phrase
WDT	wh-determiner
WHNP	wh-noun phrase
WRB	wh-adverb

INTRODUCTION

The study of how people acquire a foreign language prospered at the end of the previous century with the growth of the necessity to communicate, obtain education and compete in the job market in the international environment. Several different theories have been put forward to explain how we acquire language, what we already know before learning a foreign language, what the main principles or mechanisms we use to succeed in learning a language are. Behaviourists regard language as a mere habit formation, while mentalists attribute the process of language acquisition to the faculty for language learning known as the Language Acquisition Device. Formalists maintain that in every human language there is a set of universal principles and a finite number of parameters that define how to construct sentences, whereas typologists seek for universals across languages to predict the ease or difficulty in learning a particular grammatical feature. Accordingly, cognitive linguists focus on the building up of knowledge systems by distinguishing different stages in the process of acquiring the main grammatical structures, in which sentences occupy the highest level of grammatical attainment.

Since the language policy of the Council of Europe is aimed at 'achieving a greater unity among Member States' (Eiropas Padome, 2006: 11) by promoting mutual understanding, mobility, international communication, information access, etc., it is necessary to ensure quality and transparency in language learning. The European Union uses *the Common European Framework of Reference for Languages (CEFR): Learning, Teaching, Assessment* (Council of Europe, 2001), which defines what language learners have to learn, what knowledge and skills they have to develop, what levels of language proficiency they have to attain at each stage of learning. Moreover, the CEFR identifies the linguistic structures that foreign language learners should acquire at each level of language proficiency as all languages are acquired in a similar manner, moving from simple short sentences to more complex and longer ones, thus promoting a global approach towards language learning and assessing.

Though the CEFR provides a six-level global scale (A1 to C2) along with its descriptors to be utilized by language learners and assessors of any language, it turned out to be insufficient as the CEFR was meant to be functional and not language specific, i.e. it 'describes the communicative functions that learners should be able to perform at different languages per levels, but does not specify how those functions might be realised, e.g. in French' (Hendriks, 2008: 8). Accordingly, in 2005 the Language Policy Division of the

Council of Europe launched a project to produce Reference Level Descriptors (RLD) for national and regional languages. As a result, a collaborative programme known as English Profile (EP) was developed in 2011, which provides the RLD for English as a foreign language. Unfortunately, no RLD have been elaborated for French as a foreign language.

In Latvia as well as in some other European countries (e.g. Estonia, Lithuania, France) there is a centralised examination system, and the students graduating from a secondary school are to take an examination in one foreign language of their own choice, which is defined by the national curriculum. The specifications of the national examination determine the structure of the foreign language examination, i.e. it has to consist of five parts: reading, listening, writing, language use and speaking. Each part has to be assessed separately, and the total score is obtained by analyzing the amount of points in all parts, on the basis of which a corresponding level of language proficiency is to be awarded. In turn, educational administrators across Europe demand that language testers apply the CEFR levels as a reference tool and a calibrating instrument to produce comparable measurement systems that would function cross-linguistically and across age groups. Previous research (Kalnberzina, 2007 and Kunda, 2011) suggested that Latvian Year 12 examination level A could be related to the CEFR level C1, level B to level B2, level C to level B1, level D to A2 and levels E and F to level A1. These correspondences were implemented in 2013.

When assessing student language proficiency, markers have to rely on both the CEFR and EP to ensure cross-linguistic reliability of the judgment. Moreover, they have to be particularly objective when assessing students' written performance, since currently the role of *writing* has increased, i.e. it ensures not only effective communication and thinking skills valued highly in democratic society, but also plays a noteworthy role in judging our learning abilities and intellect. The test developers in Latvia have produced analytical marking scales in compliance with the CEFR to be used when assessing writing. Thus, student written performance is assessed along five descriptors (see Appendix 8), among which the linguistic competence of language learners is judged mainly relying on the syntactic complexity of the produced text. Nonetheless, the assessors do not have reliable syntactic indicators that discriminate between different language proficiency levels across languages. Therefore, a corpus-based quantitative, qualitative and contrastive research on different syntactic structures in the written texts produced by English and French learners is to be effected as it will ensure the retrieval of the syntactic patterns typical of each language proficiency level. Furthermore, the syntactic patterns will serve as *criterial features* (defined as 'linguistic properties that are distinctive and characteristic of each of

the levels' (English Profile, 2011: 2)) for attributing a certain level at the foreign language examination by guaranteeing that the test-taker grammatical competence is assessed objectively according to the elaborated *common* criterial features both in English and French.

The research subject

The subject of the research presented here is syntactic structures in bilingual (English and French) learner written text corpus.

The research object

English and French learner written texts of year 2009 centralised examination compiled in a corpus.

The research hypothesis

Syntactic structures are one of the criterial features used in discriminating between different levels of language proficiency when assessing foreign language learner written performance in English and French, thus ensuring the cross-linguistic reliability of the test results.

The research goals

1. To investigate the syntactic structures in written learner language.
2. To elaborate the syntactic criterial features for each language proficiency level applicable in assessing the linguistic competence of the test-taker written performance cross-linguistically.

The research objectives

1. To investigate and analyse different approaches towards describing the linguistic development of language learners and the assessment of written performance that guarantees the validity and reliability of the test results related to the CEFR.
2. To provide a theoretical basis for the syntactic analysis of different clause types produced by learners of English and French by effectuating a contrastive analysis of sample sentences from English and French contemporary grammars.
3. To compile English and French learner corpora containing structurally marked and syntactically annotated texts.

4. To carry out an empirical investigation of theoretically grounded syntactic structures in the compiled corpora by comparing quantitative and qualitative data across languages.

5. To effectuate an empirical validation of the elaborated criterial features.

Research methods in assessment of the theoretical publications

The theoretical background of the present research comprises the analytical survey and critical assessment of the 20th century and the beginning of the 21st century theories of general and applied linguistics. The theoretical part has been based upon:

1) language acquisition theories of Lado (1957), Brown (1973), Krashen (1974), Dulay and Burt (1974), Larsen-Freeman (1976), Pienemann (1998), Goldschneider and DeKeyser (2001);

2) linguistic theories of Chomsky (1965, 1981, 1995), Langacker (1972), Comrie (1977), Doughty (1988, 1991), Cook and Newson (1996), Dik (1997), Van Valin (1997), Croft (1990, 2001), Givón (1990), White (2003), Carnie (2007), Crystal (2007);

3) theories of assessing written performance and its relation to the CEFR by Hymes (1972), Bachman (1990), Weir (1990), Silva (1993), McNamara (1996), Weigle (2002), Hyland (2003), Alderson (2007), Hawkins and Filipović (2012);

4) contemporary grammars of English (Quirk et al. (1985), Biber et al. (1999), Blokh (2003), O'Grady (2005), Haegeman (2006), Carter and McCarthy (2006), Carnie (2007)) and French (Riegel et al. (1994), M. A. Jones (1996), Grevisse et Goosse (2011)).

Empirical research methods

1. The corpus compilation methodology has been applied to create a bilingual (English and French) learner corpus consisting of 191 electronic texts (44 387 tokens=words) of learners of English and 104 electronic texts (27 653 tokens) of learners of French. The compiled texts have been structurally marked according to the *Text Encoding Initiative (TEI)* standard and by creating a syntactic tagger for the text manual annotation.

2. The quantitative research method has been utilised to investigate the raw, normalised and ordinal frequency (Leech, 2011) of simple, compound and complex sentences as well as noun, adverbial and adjectival clauses in the compiled learner

corpus. The raw frequency count of annotated sentences was effected with the help of *WordSmith Tools v.5.0* and *Complete Lexical Tutor v.6.2*.

3. The qualitative descriptive analysis of various syntactic patterns and their functions in the sentences retrieved from the learner corpus has been effected by applying *the Stanford Natural Language (NLP) Parser v.3.3.0*, which provides phrase structure trees.

4. The contrastive analysis method has been applied to compare the quantitative and qualitative data cross-linguistically in order to find the similarities between the syntactic structures produced by the learners of English and French.

The research novelty

1. Specialised comparable bilingual (English and French) learner electronic corpus consisting of test-taker written essays (EFLWTC) has been compiled and syntactically annotated, which has not been accomplished previously.

2. The contrastive analysis of syntactic structures in English and French learner sub-corpora have provided evidence that foreign language learners produce, in general, the same syntactic patterns depending on their language proficiency level irrespective of the foreign language.

3. The development of syntactic criterial features ensures uniform reliable assessment of student written performance cross-linguistically, and enables the educational administrators to relate the national centralised examination test data to the Common European Framework levels.

Theoretical significance

The present research provides a theoretically substantiated and empirically validated analysis of different language acquisition theories, which enhances the comprehension of syntactic structure production order in the course of linguistic development. Furthermore, it specifies the major factors in assuring test reliability and validy in relation to the CEFR.

Practical significance

1. The structurally marked-up and syntactically annotated compiled electronic EFLWTC is a reusable source for other linguistic researchers.

2. A comprehensible quantitative and qualitative syntactic structure analysis of English and French learner written texts has been undertaken.

3. The elaborated syntactic criterial features discriminate between different levels of language proficiency and are applicable cross-linguistically to ensure reliable student written performance assessment.

The approbation of the research

1. The research results have been validated by English and French language teachers/test developers of Riga French Lyceum as well as by 100 English language assessors from all regions of Latvia.

2. The results of the present research have been trialled with the test developers and test administrators of the National Centre for Education of the Ministry of Education.

3. The theoretical significance and the empirical research results have been presented in six international scientific papers as well as reported at ten scientific conferences and seminars in Latvia and abroad (Belgium and Finland) from 2009 – 2014.

Outline of the research

The present research paper consists of an introduction, two parts, conclusions, bibliography and appendices.

Part 1 comprises two chapters. The first chapter, *Linguistic Units in Language Learning and Assessment,* gives insight into different approaches towards language acquisition by facilitating the comprehension of how we acquire a language, what preconditions determine the linguistic development of the language learner. Moreover, it emphasises the main factors in assuring the objective judgment of the learner written performance and relating the test results to the CEFR levels. The second chapter, *Syntactic Complexity*, reveals the formal and functional variety of syntactic patterns starting from a phrase level and arriving at a complex sentence via differentiating between coordination and subordination both in English and French. The analysed examples provide the theoretical basis for the syntactic structure research in the compiled learner corpus.

Part 2, *Learner Corpus in Research on Syntactic Features,* consists of three sub-chapters. The first sub-chapter, *Research Methodology and Procedure,* lists the main methods applied in carrying out the empirical part of the present thesis and describes in detail all the stages of the present research starting from the corpus compilation and finishing with the empirical validation of the elaborated syntactic criterial features. The second sub-chapter, *Research Findings and Discussions*, contains the quantitative,

qualitative and contrastive analysis of the syntactic structures retrieved from the English and French learner sub-corpora. The obtained data are compared with the English Profile, and used in elaborating the syntactic criterial features applicable when assessing student written performance in English and French. The third sub-chapter, *Empirical Validation of Syntactic Criterial Features*, analyses the validation procedure and the obtained results.

In *Conclusions* the author of the present research emphasizes the main findings and their expediency in the test development and assessment.

CHAPTER 1 LINGUISTIC UNITS IN LANGUAGE LEARNING AND ASSESSMENT

Language is generally considered to be a *system* of sounds that are linked to their *meaning* by using words and sentences, thus ensuring *thought* and *communication*. Owens (1992: 4) defines it as 'a socially shared code or conventional system for representing concepts through the use of arbitrary symbols and rule-governed combinations of those symbols.' By rules linguists understand *grammar,* which may be interpreted differently. Some consider grammar to be the correct way of speaking and writing, others relate it to the inflections common in many languages; yet others interpret grammar as the way humans organise their ideas into words (Kies, 2005). Nonetheless, it is irrefutable that grammar implies the study of how linguistic units governed by 'a finite set of underlying operational principles or rules' (Owens, 1992:4) are sequenced to express various meanings. Moreover, the linguistic knowledge is acquired in the lifelong process of language development starting from the smallest linguistic constituents – morphemes and finishing with the most complex ones – subordinate clauses (see, e.g. Pienemann, 1998).

Over the decades there have been different approaches towards understanding how grammar operates and how it accounts for the structure of language in one way or another by changing focus from mere prescription to describing its use in natural environment. An understanding of how language operates is an expedient pre-requisite in examining: i) how people learn languages; ii) what are the stages that we undergo in the development of the knowledge, referred to by Hawkins (2006: 1) as 'the developmental problem'; iii) why some structures are easier to learn than others, and why there are structures that remain difficult to master even for the advanced language learners.

Thus, *the first sub-chapter* of this study analyses different approaches towards the first and second language acquisition to comprehend the stages and processes people undergo in attaining a certain level of language competence. *The second sub-chapter* discusses the major approaches and principles in assessing learner written performance, and highlights the main pre-requisites in ensuring test result alignment across languages.

1.1 Approaches to Studying Language Learner Linguistic Development

The language study has been approached from various perspectives dictated by the domineering linguistic and psychological tendencies of each particular period of time. The

second language acquisition research boomed in the second half of the 20th century, when the focus was attributed to examining the factors influencing the children's acquisition of their native language, which grew into research on adults learning a foreign language.

The present chapter aims at disclosing the approaches towards the field of language acquisition studies whose scope of interest was the syntactic development of language learners, starting with the smallest grammatical units – morphemes, and finishing with the most complex ones – subordinate clauses. Moreover, each stage in the linguistic development was determined by various factors: biological, psychological, neurological and social factors, which served as the basis for different theoretical interpretations of the aforementioned phenomena. Thus, *Behaviourists* (e.g. Skinner, Lado 1957) studied how habit formation influenced the process of language learning. *Mentalists* (e.g. Brown, 1973; Krashen, 1974; Larsen-Freeman, 1976), having at their focal point the innate knowledge, researched in what order children acquired grammatical morphemes, which fostered the studies of the second language grammatical morphemes by Goldschneider and DeKeyser (2001). *Formalists* (e.g. Chomsky 1965) along with *Typologists* (e.g. Greenberg 1963; Comrie 1977, 1989; Givón 1990) searched for universals. Formalists studied the language form and the underlying mental processes, whereas typologists investigated languages to find typological universals. Doughty (1988, 1991) contributed to second language acquisition research by applying the Accessibility hierarchy for relative clauses in teaching English to 20 international university students. We can also mention the *processing* approach by Pienemann (1998) and the communicative approach adopted by the *Common European Framework of References* (2001) towards studying the developmental sequence of foreign language learners. Pienemann emphasised the influence of the processing capacity of humans at different stages of L2 linguistic development, whereas the CEFR focused on different communicative language competences (see *Figure 4*) relied upon when describing the language proficiency level of a foreign language learner.

1.1.1 Behaviourist Approach

Prior to the 1950s the main field of interest was attributed to teaching a foreign language under the guidance of structural linguistics, which conceives of language as consisting of a finite number of syntactic patterns. Since the 1950s the research object was relegated to understanding how people acquired the first language (L1) by examining child development. The aim of the study was to investigate how the mental grammar was built,

why the language learners 'know more than is present in the input' (Hawkins, 2006: 1). The research on L1 was a pre-requisite in carrying out the studies on L2 acquisition.

The research on first language acquisition was initiated by *behaviourists* – Mowrer (1957), Skinner (1957) and Osgood (1957) (representatives of the psychological theory in the 1950s and 1960s), who postulated that children acquired the language by using imitation and reinforcement. By imitating the language correctly, i.e. by receiving a positive impact, the children acquired new behaviour, which later could become a habit. Moreover, *transfer* was attributed the central role in the learning process (see Haskell, 2001) as it referred to how previously acquired knowledge influenced the future learning. Withal, transfer was not regarded as an instructional and learning technique but as 'a way of thinking, perceiving, and processing information' (Haskell, 2001: 23).

The behaviourists attributed the process of habit formation also to second language learning. The only difference was that a learner of a second language focused first on his natural language habits, which intervened with the ones necessary for second language learning. The success or failure in learning a language depended on the transfer of the previously acquired knowledge. If languages were similar, the learners would not encounter great difficulties, as they would be able to conform to the new situation by using the established habits. If the languages differed considerably (e.g. they had different syntactic structures), the learners encountered several interferences. For example, when learning French, the normal declarative sentence has subject-verb-object order (e.g. Le chat mange des poissons). The speaker of English encountered no problems in transferring the given sentence into French as English has the same order (e.g. The cat eats fish). However, if the object was replaced by a pronoun in French (e.g. Le chat les mange), the positive transfer was not possible because in French the pronoun is placed before the verb, whereas in English it occupies the same position, i.e. directly follows the verb (e.g. The cat eats them). Along with the positive and negative transfer we should mention also the 'avoidance' (Ellis, 1997: 51), which is typical for those learners of, for example, English whose native languages do not have relative clauses, e.g. Chinese. Lado (1957), whose research was directly related to the contrastive analysis of the native and foreign language, stated that

> individuals tend to transfer the forms and meanings, and the distribution of forms and meanings of their native language and culture to the foreign language and culture – both productively when attempting to speak the language and to act in the culture, and receptively when attempting to grasp and understand the language and the culture as practiced by natives (1957: 2).

Thus, within behaviourist framework second language learners had to overcome the differences between the first and second language. This led to the development of the *contrastive analysis* since linguists searched for similarities and differences that could either hinder or facilitate language learning. It should be specified that there were two traditions of contrastive analysis: 1) the North American tradition, which applied contrastive analysis in language teaching and learning to develop various teaching resources; 2) the European tradition, in whose scope of interest was the understanding of language itself. The representatives of the North American tradition investigated structure by structure to see which ones were easily transferrable to the foreign language and which ones had to be changed as they did not perform the required function. Thus, according to the Contrastive Analysis Hypothesis developed by Lado (1957), the similarities and diffences between the learned lagnuages determined how easy or difficult the learning process would be. Therefore, when acquiring a second language, one was expected to learn only the differences that existed between L1 and L2 since the learning process only involved acquiring new things. What is more, within the contrastive analysis framework the native language was considered to be a 'driving force' (ibid.) of second language learning, which resulted into criticism and renunciation of behaviourist theory. Ellis (1997) postulated that 'behaviourism cannot adequately account for L2 acquisition […] Learners frequently do not produce output that simply reproduces the input […] learning is not just a response to external stimuli' (1997: 32). Thus, further research was conducted to understand the syntactic and morphological development of L1 and L2 since it was not satisfactory to base the contrastive analysis on mere introspection.

1.1.2 Mentalist Approach

In the 1970s the notion of imitation and habit formation was replaced by the postulation of children's innate knowledge of L1 since the research carried out on child language acquisition during the 1950s and 1960s revealed some systematicity in learning a language, i.e. language was seen as a set of rules which were formulated by taking into consideration both innate principles and the exposure to the learned language. In order to explain how L1 was learned, the attention was attributed to the human mind, as *mentalists* considered that all human beings possessed a faculty for learning languages, referred to as a Language Acquisition Device (LAD). By receiving the input, the LAD was set in motion, i.e. it started operating. When acquiring L2, language learners constructed a completely different linguistic system known as *interlanguage* (the term coined by Selinker, 1972), which was

based upon the rules of both L1 and L2 linguistic systems. Thus, the language learner grammar was influenced by the inner grammar (L1) and the input from the outside, i.e. L2 grammar. The term *interlanguage* was used by the mentalists to explain how L2 was acquired.

Brown (1973) contributed greatly to second language acquisition research by initiating empirical studies to examine the development of grammatical morphemes in children's native language. He noted that there was a certain order in English children's morpheme acquisition (see *Table 1*), which could be predictable, e.g. the regular form of the plural was acquired in the second stage after the mastering of the present progressive.

Table 1 **Order of grammatical morpheme acquisition for L1 learners of English (Brown, 1973)**

1.	Present progressive - *ing*
2.	Plural -*s*
3.	Irregular past tense
4.	Possessive -'*s*
5.	Copula *be*
6.	Articles *a/an* and *the*
7.	Third person singular present tense -*s*
8.	Auxiliary *be*

Moreover, the average number of morphemes and their complexity increased along with children's linguistic development. R. Brown distinguished six stages in children's language development (*Table 2*), i.e. stage II (27-30 months) is characterised by morphological development; at stage III (31-34 months) – we could observe sentence form development, which was followed by stages IV - V+ with the development of embedded sentences and clause linking (for a more detailed linguistic development see *Appendix 1*). By the age of five the child had acquired the basic sentence competence.

Table 2 **R. Brown's (1973) stages of development (adapted from Owens, 1992: 308)**

Stage	Approximate Age (months)	Characteristics
I	12 - 26	Linear semantic rules
II	27 – 30	Morphological development
III	31 – 34	Sentence form development
IV	35 – 40	Embedding of sentence elements
V	41 – 46	Joining of clauses
V +	47 +	

Brown's research on *children's* morpheme acquisition was followed by the studies on second language acquisition of Dulay and Burt (1974) who regarded children's second language acquisition as similar to first language acquisition, which meant that there were

similar processes in the development of L1 and L2. Having studied a group of Spanish and Chinese speaking children learning English as L2, they found evidence for a certain order of morpheme acquisition different from the order of a child learning English as L1 (see *Tables 1* and *3*), e.g. in Dulay and Burts elaborated order children first learned the plural morpheme *-s* and only then the present progressive morpheme *-ing*; the copula *be* was mastered after the present progressive, while in Brown's order the copula *be* appeared a bit later, i.e. it occupied the fifth position.

Table 3 **Order of grammatical morpheme acquisition for L2 learners of English (Dulay and Burt, 1974)**

1.	Plural *-s*
2.	Present progressive *-ing*
3.	Copula *be*
4.	Auxiliary *be*
5.	Articles *a/an* and *the*
6.	Irregular past tense
7.	Regular past tense *-ed*
8.	Third person singular present tense *-s*
9.	Possessive *-'s*

In addition, the research of Dulay and Burt revealed little difference in the acquisition of English morphemes irrespective of L1. As a result, Dulay and Burt (1974) developed the theory of 'creative construction', which is

> the process in which children gradually reconstruct rules for speech they hear, guided by universal innate mechanisms which cause them to formulate certain types of hypotheses about the language system being acquired, until the mismatch between what they are exposed to and what they produce is resolved (Dulay and Burt, 1974: 37).

Hence, the focus of their research was on development rather than on the role of the native language. Moreover, they emphasised the mental processes, as for them the cornerstone for language learning was the innateness. They claimed that the processes involved in acquiring L1 and L2 were the same by trying to provide empirical data of second language learning by children.

Bailey, Madden and Krashen (1974) continued to investigate whether *adult* second language acquisition was influenced by the same processes as child L2 acquisition. They studied the oral form of the language, and the research was aimed at providing evidence that native language (NL) did not have a significant influence upon second language learning, which disclaimed the behaviourist viewpoint about L2 acquisition. Their studies resulted in positing the natural order of the acquisition of English morphemes. Krashen's Natural Order Hypothesis (1977) also supported the idea that we could predict the

morpheme acquisition order irrespective of L1 or L2. By modifying Dulay and Burt's (1974) morpheme acquisition order, Krashen (1977) organised them in four groups.

Group 1	Group 2	Group 3	Group 4
Plural -*s* Progressive -*ing* Copula *be*	Auxiliary *be* Articles	Irregular Past tense	Regular Past tense 3rd person Singular Present -*s* Possessive -'*s*
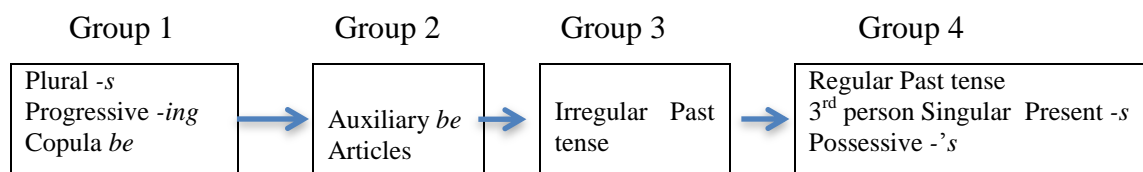
Figure 1 **Order of grammatical morpheme acquisition for L2 learners of English revised by Krashen (1977)**

Consequently, the morphemes in group 1 were acquired before those in group 2; the morphemes in group 2 before the ones in group 3, etc. However, there could be variations within a group.

Though Krashen's (1977) morpheme order studies did notice the role of NL, in general, the obtained results were similar with Bailey and Madden's (1974) findings, with the exception that Bailey and Madden were not able to neglect the role of NL in L2 acquisition.

Larsen-Freeman (1976) expanded the scope of the research by examining both oral and written language of 24 adults coming from different L1 backgrounds. The obtained results were similar to Dulay and Burt's (1974) findings for oral language, whereas for written language they differed radically. Thus, it was proved that oral and written forms should not be judged according to the same criteria, and that age was not the determining factor in language acquisition order.

The research on systematicity in acquiring a new language was continued 25 years later by Goldschneider and DeKeyser (2001), who investigated the order of acquisition of grammatical morphemes by conducting a meta-analysis of twelve English L2 studies. Their research resulted in establishing a set order of the acquisition of grammatical morphemes (see *Table 4*) similar to Dulay and Burt's order (see *Table 3*), though Goldschneider and DeKeyser based their explanation of the order on external factors (i.e. phonological, morphological, syntactic, semantic and numerical aspects of salience) rather than on internal factors (i.e. innate language learning mechanisms). Moreover, they also claimed that their morpheme acquisition order could be attributed not only cross-linguistically, but to any age group irrespective of the first language background.

Table 4 **Goldschneider and DeKeyser (2001) Morpheme acquisition order**

1.	plural -*s*
2.	progressive -*ing*
3.	copula *be*
4.	auxiliary *be*
5.	articles *the/a*
6.	irregular past tense
7.	third person *s*
8.	possessive -*'s*

However, if compared with French, there would be some discrepancies, which was already proved by Lightbown (1983) when studying French L1 speaking children learning English as L2 in Quebec. French speakers would not make any distinction between present simple and present progressive because in French present simple would be used in either case (e.g. I eat – Je mange; I am eating – Je mange); the possessive case would be denoted by the use of *de* instead of *'s* (e.g. Sarah's car – la voiture de Sarah), etc. Thus, the accuracy order would be different because of L1 influence, which was also supported by Ellis (2006), who argued that 'L2 acquisition is clearly affected by the transfer of learners' knowledge of their first language' (2006: 187).

Having examined different models of studying the morpheme order acquisition, we may conclude that the major contribution of morpheme acquisition research was to understand the nature of developmental sequences by moving from behaviourist interpretation of L2 acquisition to mental processes involved in learning L2. The research proved that there was a fixed order or natural sequence in learning grammatical morphemes similar to all learners irrespective of their age, L1 and their background knowledge of L2; thus, the existence of 'a psychologically natural order' (Littlewood, 2004: 40) due to L1 knowledge should not be denied. Nonetheless, the *accuracy* order of learning various linguistic units appeared to be *relatively* similar.

1.1.3 Formalist and Typologist Approach

The formalist and typologist approach to language acquisition incorporated the search for linguistic universals (known as Universal Grammar) and typological universals. The first approach was represented by Noam Chomsky (1965), who tried to comprehend how humans acquired the *knowledge* of language (also referred to as *competence* of language). The second approach was initiated by Greenberg (1963), who aimed at discovering similarities and differences in languages worldwide. Though the formalist and typologist approaches were 'diametrically opposed' (Croft, 1990: 3), they still shared some major

characteristics, i.e. both of them analysed language structure and both approaches utilized 'a considerable amount of abstraction' (ibid.: 3).

Noam Chomsky contributed enormously to explaining how humans were able to build mental grammars, which Chomsky (1986) referred to as I-language (internalized language). He considered that human beings were able to acquire a language due to the fact that human beings were born with a language acquisition device (LAD) (known also as Universal Grammar (UG)), which contained 'universal underlying principles or phrase structure rules', also known as 'linguistic universals'. In order to comprehend how language was acquired, Chomsky examined the linguistic development starting from a new-born baby and finishing with an adult L1 speaker. Chomsky (1980b: 37) postulated that every human being had to progress from S_o (initial zero state), to the final stage or steady state S_s. Moreover, one had to undergo a sequence of stages $S_1, S_2 \dots$ by processing the linguistic data with the help of LAD to arrive at S_s where the linguistic competence in the language, i.e. a 'generative grammar' was acquired. Hence, by comparing the language input, processed by the child's LAD, and the output of the linguistic knowledge (see *Figure 2*), Chomsky deduced what was happening in the LAD accessible to any child irrespective of their L1.

Input → Language Acquisition Device → Output
(primary linguistic data)   (a generative grammar)

Figure 2 **The LAD model of L1 acquisition (Cook and Newson, 1996: 80)**

In 1964 another model (*Figure 3*) was elaborated which incorporated Universal Grammar (previously known as LAD) and principles and parameters theory.

Input → Universal Grammar → Output
(primary linguistic data)   (a grammar consisting of principles, parameters and lexicon)
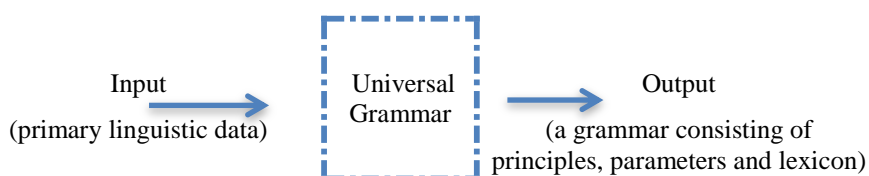
Figure 3 **The Universal Grammar model of L1 acquisition (Cook and Newson, 1996: 81)**

Chomsky (1965) maintained that Universal Grammar was an innate language faculty having 'a certain mental structure consisting of rules and principles that generate and relate the mental representations of various types' (Chomsky, 1980a: 48) that enabled the L1

acquirer to arrive at the output stage by using the input data. What is more, UG contained principles, e.g. the Projection Principle, Government Binding Principle, etc., invariant across languages (regarded as universal) and parameters which varied from one language to another. Though UG did not postulate that all languages were the same, which is self-evident, the question arose whether a principle that was not found in all languages could be regarded as a universal and in what sense that particular universal could be a universal. The most important aspect was not the need to have universals present in all languages but the fact that they might not be breached. 'If the principle can be ascribed to the language faculty itself rather than to experience of learning a particular language, it can be claimed to be universal on evidence from one language alone' (White, 2003: 29). Therefore, it was more important to discover the principles underlying the aforementioned rules.

A particular approach to the study of language was adopted known as '*Principles and Parameters theory*' (Chomsky, 1980), which aimed 'to reduce descriptive statements to two categories: language-invariant, and language-particular' (Chomsky, 1997: 25) (this theory was revised several times and resulted in the Minimalist Program (Chomsky, 1995)). According to *Principles and Parameters* theory, principles were considered to be language-invariant. Thus, the principles were the ones to define 'the structural architecture of human language' (Hawkins, 2006: 13), which was demonstrated by examining the structure of phrases (known as X-bar theory of phrase structure). It was assumed that phrases bore the same underlying structure across languages, i.e. they consisted of the core - *head* (noun, verb, adjective and preposition) that could be modified by *complements* and *specifiers*. For example, *Mike's report about his discoveries* constitutes the noun phrase, which comprises the specifier [Mike's], the head [report] and the complement [about his discoveries]. Though the normal pattern would be for the complements to follow the head, and the specifiers to precede the head, there are languages in which the specifier and the complement follows the head. For example, in French the head [Le rapport] precedes both the specifier [de Mike] and the complement [sur ses découvertes]. Hence, we may account for parametric variations, which Chomsky attributed to language-particular category.

The parametric differences were mainly associated with the functional categories, such as complementizers, inflections, determiners, negation, etc. According to Lydia White, 'parameters give the language acquirer advance knowledge of what the possibilities will be, that is they severely restrict the range of choices that have to be considered' (2003: 157). If we examine the word order parameter, then we see that, for example in French, a finite verb appears to the left of the negative *pas*, while in English it remains to the right:

Emilie ne regarde *pas* le film.

Emily does *not* watch the film.

The difference between the languages can be observed also in the position of the adverb in a sentence.

Emilie voit *souven*t Pierre.

Emily *often* sees Peter.

In French the finite verb appears before the adverb, while in English it is placed after the adverb. At an underlying level both languages have the same structure (SVO), the differences of verb placement appear because of the differences in feature strength (French has strong inflections, while English has weak). Thus, alterations in word order are determined by one parametric difference: the strength of verb related features in inflections.

Most languages have one of the following word orders: SVO, SOV, VSO. A few languages have VOS and OSV order, while no languages use OVS order (Greenberg, 1966). If we imagine that there is a parameter which determines the basic word order of the first four options excluding the last two ones, then the child acquiring English will choose SVO setting as it is innately available. 'Language variation thus reduces to learning the correct set of words and selecting from a predetermined set of options' (White, 2003: 23), which means that a child does not learn the rules of a particular language but only the parameter settings. As a result 'a core grammar' is being created.

For many years UG was considered to be the main driving force for child native language acquisition. However, research proved that it could also be applied to L2 acquisition as, according to Gass and Selinker (2008) 'if properties of human language are part of the mental representation of language, it is assumed that they do not cease being properties in just those instances in which a non-native language system is being employed' (2008: 161). However, learning L2 differs from learning L1 as people who start learning L2 already know L1, which means that many UG principles are present in their L1 grammar. Moreover, since UG is constant and distinct from the learner's L1 grammar, it also constrains the L2 learner's interlanguage grammar (White, 2003: 60). Consequently, when acquiring L2, learners are able to apply both L1 grammar and the parameters not present in L1, but provided by UG. However, not all learners of L2 attain the same proficiency level as in L1. This might be explained by several factors, among which the accessibility of UG occupies the primary position in denoting how successful the learning process will be. We may distinguish four possible positions (Ellis, 1997: 69): 1) L2

learners begin with L1 parameter settings and then switch to L2 parameter settings, which results in complete L2 competence; 2) learners of L2 do not have the access to UG; in such a case they rely mainly on general learning strategies and do not reach full competence in L2; 3) L2 learners have a partial access to UG, e.g. they have access only to the parameters of L1; though having received direct instruction, they are able to switch to L2 parameters; 4) L2 learners use both UG and general learning strategies; however, the latter one may disable the functioning of UG by causing the apparition of errors.

Though Chomsky's *Principle and Parameter* theory was proposed as 'a universal model of constituent structure' by representing to a certain extent the nuclear syntax of all languages, research on second language acquisition (e.g. White, 1985; Hilles, 1986; Park, 2004) induced ambivalence on the role of UG in acquiring a language. Though White (2003: 149) concluded that 'interlanguage grammars conform to parameters of UG', the researchers could not provide a uniformed answer whether L1 and L2 acquisition were fundamentally similar or not.

Overall, it makes us conclude that we cannot rely only on UG when explaining the success or failure of L2 acquisition since there are many more factors to be taken into consideration.

Another approach within the linguistic framework towards L2 studies was the typological one (sometimes referred to as functional-typological approach, since it was closely related to the functionalist viewpoint which tried to explain the linguistic structure in terms of linguistic function (Croft, 1990)). *Typologists* based their research upon the comparative studies of a variety of languages from all over the world. They studied the *language universals* (aspects common to all languages) or on the contrary - how languages differed by looking at *language typology* from outside, i.e. they made generalisations about certain linguistic phenomenon based upon empirical observations.

The typologists distinguished between *implicational* and *non-implicational* universals (e.g. all languages have vowels). By *implicational* we understand that if a language has X, then it also has Y, but not vice versa. For example, if a language has a category of dual, it also has a category of plural. Another example of an implicational universal could be the Accessibility Hierarchy elaborated by Keenan and Comrie (1977), which is related to relative clause formation. The basic claim is that we can predict the types of relative clauses in a particular language according to the order 'subject > direct object > non-direct object > possessor' (Comrie, 1989: 156), if we assume that subject relative clauses exist in all languages. Thus, if there is a relative clause of the type X (e.g.

direct object), then there will also be any relative clause type to the left of type X, i.e. there will also be the subject relative clause. Consequently, for each position there is a language that can form relative clauses on that position and to the left, but not vice versa. In addition, languages vary in some more aspects: 1) whether they have relative clauses (e.g. English and French), or not (e.g. Chinese and Japanese), which determines how easy or difficult it is for learners to acquire the relative clauses of L2 (highlighted also by the contrastive analysis hypothesis); 2) what is the position of the relative clause with respect to the main clause (e.g. in English and French the relative clause may follow the main clause or be embedded in the main clause). Thus, the existence of relative clauses in L1 facilitates the learning of L2; moreover, the linguistic structures of L1 denote the process of L2 acquisition since learners of L2 will first acquire the structures present in L1.

Though the research on relative clauses was initially conducted by Keenan and Comrie (1977), when studying the world's languages, Doughty (1991) continued to examine relative clauses within the scope of second language acquisition and classified them in *the Accessibility hierarchy for relative clauses in English*. The research demonstrated that there was a certain order in learning different types of relative clauses, i.e. first we acquired relative clauses with a subject pronoun and afterwards the ones with an object pronoun (see *Table 5*). Moreover, learners who were able to use the structures at the bottom of the hierarchy (e.g. relative clauses in which the relative pronoun functions as object of comparative) would be able to use the other structures as well, which might not be the case with those learners who had acquired the first two structures with the relative pronoun having the function of the subject or object since it was not evident that they would learn also any other relative clause types.

Table 5 **Accessibility hierarchy for relative clauses in English and French (adapted from Doughty, 1991)**

Relative pronoun function	*Example*
1. Subject	The boy *who is crying* lost his mobile phone. Le garçon *qui pleure* a perdu son téléphone mobile.
2. Direct object	The boy *whom I met* was Henry's friend. Le garçon *que j'ai rencontré* était l'ami de Henry.
3. Indirect object	The boy *whom I gave my toy-car* was Henry's friend. Le garçon *à qui j'ai donné mon jouet voiture* était l'ami de Henry.
4. Object of preposition	They found the shop *that the ladies in the village were talking about.* La chaise *sur laquelle je suis assis* est très ancienne.
5. Genitive	The woman *whose daughter is studying abroad* organises a party. La femme *dont la fille fait ses études à l'étranger* organise une fête.
6. Object of comparative	The car *that is bigger than yours* is mine. La voiture *qui est plus grande que le vôtre* est la mienne.

All in all, the accessibility hierarchy demonstrates that 'on the one hand, linguistic facts can be used to explain and even predict acquisition. On the other hand, the results of empirical studies of L2 acquisition can be used to refine our understanding of linguistic facts' (Ellis, 1997: 65). What is more, it confirmed the assumption that 'learners grammars are constrained in a similar way to natural language grammars' (Gass and Selinker, 2008: 199).

Although these two approaches towards the study of linguistic universals were initially opposed to each other, they became closer with time, since both studies were concerned with variations across languages, the only difference being that research on language universals was concerned primarily with limits on variation, whereas typological research was concerned more directly with possible variation (Comrie, 1989: 34). Thus, in practice these are two parallel processes.

1.1.4 Processing Approach

Another approach towards the studies of second language acquisition is known as the *processing approach* based upon a universal processability hierarchy of linguistic forms. In compliance with this approach, when acquiring second language the order of linguistic unit acquisition is determined by the operation of the processing mechanisms and capacities of the human brain in real time.

In order to substantiate the aforementioned assumption, Pienemann, Johnston, and Brindley (1988) studied how questions were acquired in English L2 by language learners from different L1 backgrounds. The research provided evidence that there was a certain order, or developmental stages, when acquiring questions in English L2. For example, if the learners attempted to produce Aux-second (see Table 6) already at stage 3, they would generate several interlanguage variants (e.g. Where he been?) as they were not 'ready' for that particular syntactic structure.

Table 6 **The sequence of question formation (Pienemann, 2008: 138)**

Stage	*Structure*	*Example*
Stage 1	SVO question	He live here?
Stage 2	*wh-* + SVO	Where he is?
Stage 3	Copula inversion	Where is he?
Stage 4	Aux-second	Where has he been?

Though the research revealed that in general the sequence of question production was similar in L1 and L2, there might be some discrepancies if question formation rules in L1

and L2 differed. This was due to the fact that at lower stages of language development (stages 1 and 2) question formation was influenced not only by the processing mechanisms, but also by the first language. For example, German learners of English would first use questions without subject-auxiliary inversion (*e.g. Speak you English?*), and only at stage 4 they would be able to understand that they had to produce the construction of S-V inversion to form questions. For French learners of English the question formation would not cause any difficulties since they had the same rule for S-V inversion.

Consecutive studies on language development determined by processability resulted in elaborating the 'processability theory' (Pienemann, 1998), in which Pienemann postulated that 'structural options that may be formally possible will be produced by the language learner only if the necessary processing resources are available' (1998: 2). It means that at a certain stage of development the learner can produce and understand only those linguistic forms that are accessible to the human mind. Consequently, when acquiring a language, the learner can acquire what he/she can process, i.e. language development is determined by processability. Pienemann, by applying *processability theory*, shows the order in which the main grammatical encoding procedures are activated in syntactic structures in the acquisition of English as a second language:

Table 7 **Hierarchy of processing procedures - summary (Pienemann, 1998: 87)**

1. Lemma access; words; no sequence of constituents;
2. Category procedure; lexical morphemes; no exchange of information – canonical word order;
3. Phrasal procedures; phrasal morphemes;
4. Simplified S-procedure; exchange of information from internal to salient constituent;
5. S-procedure; inter-phrasal morphemes; exchange of information between internal constituents;
6. Subordinate clause procedure.

The table illustrates that there is a time sequence involved in producing grammatical constructions, e.g. noun and verb phrases appear before sentences; category procedure appears before phrase procedure, etc. It follows that the basic thesis of Pienemann's theory is that

> in the acquisition of language processing procedures the assembly of the component parts will follow the above-mentioned implicational sequence. The hierarchical nature of this list arises from the fact that the procedure of each lower level is a prerequisite for the functioning of the higher level (ibid.: 7).

It means that every procedure in the hierarchy is obligatory for the next procedure to take place, e.g. the grammatical category of a lemma is needed before the category of procedure.

Overall, cognitive psychologists who investigate the information processing model of human learning and performance tend to see second language acquisition as the building up of knowledge systems. Moreover, as Pienemann postulates, 'the processability hierarchy is not language specific and, in principle, it applies to the transfer of grammatical information in any language' (Pienemann, 2008: 143). The basic reasons (ibid.: 141) why language learners follow this hierarchy are: a) the hierarchy is implicationally ordered, that is, every procedure is a necessary prerequisite for the next procedure; b) the hierarchy mirrors the time-course in language generation.

1.1.5 Communicative Approach (the Common European Framework)

Since the interest in language learning and teaching in Europe increased, another approach towards language studies was commenced in the 1970s, which moved away from the 'structure-dominated scholastic sterility' (VanEk and Trim, 1991: 1) to language use and language functions.

In 1972 Hymes developed the *communicative* approach, which distinguished between two kinds of knowledge essential to successful language use: linguistic knowledge and knowledge about what language was appropriate in particular social contexts. In 1990 Lyle F.Bachman elaborated the theoretical framework of communicative language ability (CLA), which comprised three components: language competence, strategic competence, and psychophysiological mechanisms (Bachman, 1990: 84). The language competence included morphology, syntax, vocabulary, etc. (*see* Appendix 2). In 2001 the Common European Framework of References (CEFR) was developed with the aim to provide language level proficiency definitions to ensure 'the transparency of certified language competence' (Alderson, 2007: 22) by bringing together 'a wide range of thinking and research in language education under a common umbrella' (ibid.).

The approach adopted by the CEFR was 'action-oriented […] as it views users and learners of a language primarily as 'social agents'' (2001: 9) who had to fulfil particular tasks (e.g. produce written texts) by using specific competences – both general (not specific to language, e.g. knowledge of the world, sociocultural knowledge, etc.) and communicative language competences (the ones which 'empower a person to act using specifically linguistic means' (ibid.)). Being influenced by Bachman's model of CLA, the

CEFR described the learner competence by breaking the communicative language competences into: linguistic, sociolinguistic and pragmatic competences (*Figure 4*).
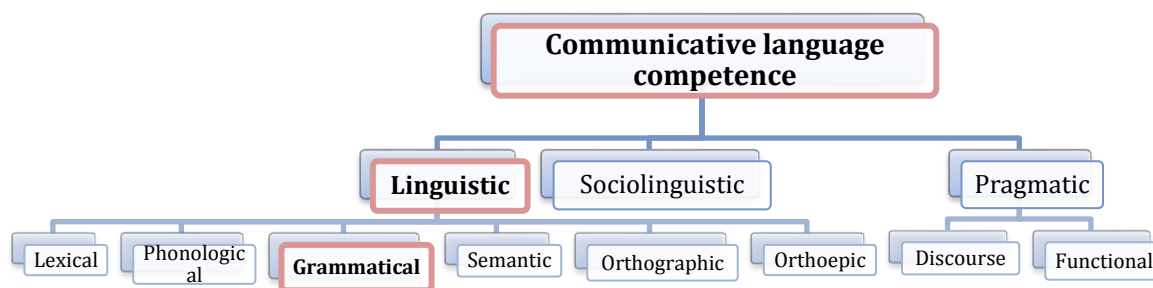
Figure 4 **Communicative language competences identified by the CEFR (2001)**

Since language is a very complex system (see Chapter 2), which evolves in the process of communication, the CEFR subdivided the linguistic competence into: lexical, grammatical, semantic, phonological, orthographical, and orthoepic competences, while the pragmatic competence distinguished between discourse and functional competence. Subsequently, the grammatical competence (the main objective of the present research) envisaged the 'ability to understand and express meaning by producing and recognising well-formed phrases and sentences' (ibid.:113). It is undeniable that the process of language learning is long-term and individual, during which the learners do not develop the same competences in a similar way. Therefore, in order to measure how language learners succeed in attaining a certain level of competences, the CEFR provides the language proficiency levels A1 (Breakthrough), A2 (Waystage), B1 (Threshold), B2 (Vantage), C1 (Effective Operational Proficiency) and C2 (Mastery) (2001: 30) with its descriptors (*see* Appendix 3) according to which the progress in learning a language could be measured over a certain period of time (further discussed in Chapter 1.2).

Conclusion

The analysis of different approaches towards studying L1 and L2 acquisition reveals the complexity of the processes and mechanisms involved in language acquisition. While behaviourists regard language learning as a mere habit formation whose success or failure depends on the transfer from L1 to L2 (either positive or negative), mentalists attribute the main role to the LAD by revealing a certain order in linguistic development, which could be predicted. Formal linguists investigate how people acquire and use the innate knowledge of language, whereas typologists analyse the patterns of language structure across languages. According to the processability theory, second language learners

undergo universal developmental stages along with individual variations constrained by psychological development, and the CEFR serves as a basis for measuring progress of language learners at a particular stage of their linguistic development in compliance with the common proficiency level descriptors. Consequently, we cannot reject any of the aforementioned approaches since each of them constitutes a piece in a big puzzle, known as language. The linguists are to decide which aspect of language to study - whether we want to find out what is common to all languages; how sentences are structured grammatically or what words, sentences and texts mean. In general, all the approaches provide a substantial theoretical grounding for contemporary researchers to base their findings and allegations upon.

While initially the research on linguistic development was related to second language acquisition in natural settings, with the advances in technology and globalisation the necessity to learn another language along with the native language increased, which changed the focus of much linguistic research from L2 acquisition to foreign language learning in the classroom setting or outside the local community of the language to be learned. This distinction was clearly marked also by the use of different concepts – *second language* vs *foreign language* and *acquisition* vs *learning*. Though there is not great difference whether we learn language in classroom or in a naturalistic environment as, in principal, the learning mechanisms in both cases are similar, within the framework of the present research the concept *foreign language learning* will be utilised. The main reason is that the majority of test-takers have learned English or French as a foreign language in the classroom settings. Moreover, the Republic of Latvia Language Law adopted in 1992 defines the Latvian language as the official state language and any other language (except the Liv language) used in the Republic of Latvia as a foreign language.

The following chapter outlines different approaches towards assessing writing ability by emphasizing the linguistic competence of language learners and discusses the three fundamental principles of assessment: reliability, validity and feasibility. Moreover, it accentuates the importance of relating test results to the CEFR to arrive at a common framework for measuring student written performance cross-linguistically.

1.2 Approaches and Principles of Written Performance Assessment

Nowadays writing is an integral part of the curriculum in schools from the earliest grades onward and is used as a medium for testing language learner knowledge. In Latvia students graduating from school are supposed to take the centralised examination in a foreign

language, one part of it being the written performance. Though writing is considered to be a model of a communicative language use (Grabe and Kaplan, 1996), accuracy and clear exposition are still regarded as one of the main criteria for good writing (Hyland, 2003: 9).

Traditionally, speech has been considered to be primary compared to writing. However, some educational research (e.g. Grabe and Kaplan, 1996) opposes this assumption and values the written form higher than the oral production, as the written form of the language:

> can be seen as a more standardized system which must be acquired through special instruction. Mastery of this standard system is an important prerequisite of cultural and educational participation and the maintenance of one's rights and duties... The fact that writing is more standardized than speaking allows for a higher degree of sanctions when people deviate from that standard (Grabowski, 1996: 75).

Writing as a distinctive area of studies emerged in the 1980s (e.g. Graves, 1984; Hayes and Flower, 1980; North, 1987). At that time the focus was on formal text units, which involved mainly linguistic knowledge, vocabulary, syntactic patterns and cohesive devices. This was due to the influence of structural linguistics and the behaviourist theories (1960s) that emphasized writing as a combination of lexical and syntactic forms by considering writing to be

> an extension of grammar – a means of reinforcing language patterns through habit formation and testing learners' ability to produce well-formed sentences. For others writing is an intricate structure that can only be learned by developing the ability to manipulate lexis and grammar (Hyland, 2003: 3).

However, syntactic complexity and grammatical accuracy are not the only aspects to be relied upon to weigh what good writing is. According to the functional approach (e.g. Firbas, 1992; Halliday and Matthiessen, 2004), it is considered that 'particular language forms perform certain communicative functions' (Hyland, 2003: 6), that structures should be related to meaning to achieve a particular purpose of writing. Thus, when developing a paragraph it should contain a topic sentence, supporting sentences and transitions. It is evident that this approach has also been influenced by the structural model since paragraphs are regarded as syntactic units like sentences in which 'writers can fit particular functional units into given slots' (Hyland, 2003: 6).

Writing in a foreign language is considered to be more problematic than writing in the first language because it tends to be 'more constrained, more difficult, and less effective' (Silva, 1993: 668). The candidates may need to search for appropriate lexical and syntactic units, they may not be aware of the social and cultural uses of writing as well as they may misuse various language functions. Moreover, writing in a foreign language

depends also on the proficiency in the first language. Saville-Troike (2006: 165) states that 'transfer of effective language-specific writing processes that have been acquired in L1 and L2 is not possible until a threshold level of L2 structural knowledge has been reached.' What is more, writing in a foreign language whose grammar, vocabulary and writing system is closer to the native language is easier than if the foreign language differs greatly from L1.

The present chapter focuses on elucidating the main approaches towards performance assessment by deciphering between competence and performance (researched by Chomsky, 1965; Hymes, 1972; Canale and Swain, 1980; Bachman, 1990 and McNamara, 1996), which are taken into consideration when assessing writing. Further on, the author discloses the fundamental concepts of assessment – reliability, validity and feasibility along with the four stages in the process of relating the examination to the CEFR, which provides common measurement scales and language level descriptors to ensure the cross-linguistic alignment of the test-taker written performance.

Bachman (2004: 6-7) defines assessment as 'the process of collecting information about something that we're interested in, according to procedures that are systemic and substantively grounded.' One of the first approaches towards language testing known as the 'skills and elements' model was represented by Lado (1961) and Carroll (1961, 1968), who made a clear distinction between skills and abilities. Lado considered that the abilities such as pronunciation, vocabulary, and grammar were 'integrated in the total skills of speaking, listening, reading and writing' (Lado, 1961: 25) and could be tested separately. For him, language knowledge incorporated only skills. When developing language tests, Lado and Carroll based them on the psychometric model, i.e. they applied a discrete-point approach (known also as item testing), which tested only one aspect of grammar at a time. This approach influenced a whole generation of language testing by incorporating notions of reliability and validity from psychometrics (Bachman, 2007: 47).

In 1980 Canale and Swain proposed another approach towards assessing language by separating the grammatical competence (phonology, lexicon, syntax and semantics) from all the other language competence components and developing a model of communicative competence that contained grammatical, sociolinguistic, discourse and strategic competences. Though they were not able to 'articulate the relationship between grammatical competence and the other competences in their framework' (Purpura, 2004: 53), their approach had a huge impact on foreign language education, and later on was developed by Bachman (1990) and Bachman and Palmer (1996) into a multi-componential

model of communicative language ability (see Appendix 2), which served as a basis for the CEFR communicative language competences. Thus, it is obvious that the initial approach towards assessing language knowledge has undergone noteworthy changes from mere skills assessment to performance assessment supported also by Weigle (2002), who states that 'any writing test that involves actual writing can be considered a performance test, [...] since the written product represents a performance of writing' (2002: 46). Subsequently, most writing assessments are concerned with performance assessment, which implies test-takers producing relevant pieces of writing and raters evaluating the results most validly and reliably.

Traditionally there have been two approaches towards *performance assessment* in a foreign language as stated by McNamara (1996: 6):

1) a simple application of the techniques of performance assessment developed in non-language contexts called *work sample* approach;

2) a tradition that sees performance as *a complex cognitive* achievement focusing on the *quality of the execution* and *the underlying state of language knowledge* in the individual being assessed.

According to Messick (1994), in the first approach the performance is the *target* of assessment while the second approach is cognitive and psycholinguistic in which the performance is the *vehicle* of assessment; the underlying knowledge and ability is the actual target of assessment. The latter one originated in the precommunicative testing and is considered to be 'the most common approach to general-purpose performance assessment' (McNamara, 1996:26).

Written performance assessment is particularly complex, as it encompasses both competence and performance. Various linguistic theories examine these concepts differently. Chomsky (1965) makes a clear distinction between competence and performance as for him the focus of interest is mainly in the knowledge of language, i.e. competence. Hymes (1972) relates the term *performance* to the actual use and actual events, while Canale and Swain (1980) view the *ability for use* as a part of the *communicative performance*. Bachman (1990) refers to both competence (knowledge) and its implementation in language use (performance) when using the term 'communicative language ability', which coincides with Widdowson's (1993) term 'capacity' in defining language proficiency. When referring to the performance assessment in language testing, McNamara (1996) goes further and distinguishes between its *strong* sense and *weak* sense. In the *strong* sense the focus is on the fulfillment of the task in which language ability is

not of primary importance. For example, in a persuasive essay the test-taker will be successful if he/she manages to persuade the reader. Whereas in the *weak* sense, the focus is on the language used, which means that the assessors are more interested in the linguistic aspects of the written performance than in whether they feel persuaded or not. However, McNamara considers language tests to be somewhere in between these two 'ends of a conceptual continuum' (1996: 45), which is reflected in *Figure 5*.
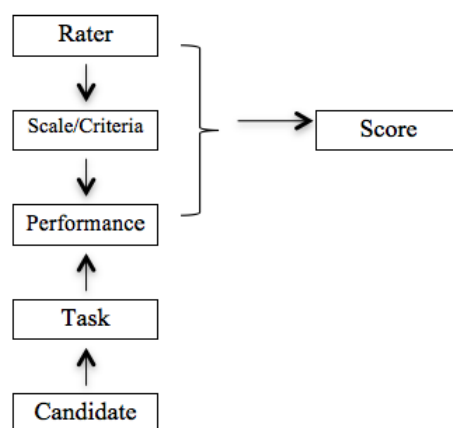
Figure 5 **'Proficiency' and its relation to performance (adapted from McNamara, 1995)**

This model can also be attributed to language testing in Latvia since competence can be evaluated only through performance, i.e. via assessing the written performance. Moreover, linguistic competence influences significantly the overall impression of the written performance as

> all good writing consists of good sentences properly joined. Since the sentence is the foundation or unit of discourse, it is all-important that the pupil should know the sentence. He should be able to put the principal and the subordinate parts in their proper relation: he should know the exact function of every element, its relation to other elements and its relation to the whole. He should know the sentence as the skillful engineer knows his engine, that, when there is a disorganization of parts, he may at once find the difficulty and the remedy for it (Bornstein, 1976: 44).

Consequently, when assessing written performance much attention is given to the linguistic knowledge including the basic syntactic patterns of a foreign language, which is the main area of the present research. Hence, to attain good results, we should describe:

1) The grammatical development for writing, i.e. what types of sentences students should be able to produce at a certain level of language proficiency, what should be the clause length, at which level we should expect the increase in embedded subordinate clauses and decrease in coordination or compound sentences. The CEFR provides syntactic descriptors for each level of language proficiency (A1 – Breakthrough; A2 – Waystage; B1 – Threshold; B2 – Vantage; C1 – Effective

Operational Proficiency; C2 – Mastery), which should be universal, i.e. applicable cross-linguistically, thus enabling a uniform interpretation of language learner success in acquiring either L1 or L2. For example, at B1 students should produce simple sentences, given that noun and verb phrases are not overloaded. They should be able to produce compound sentences. They can be expected to produce complex sentences limited to one subordinate clause of fairly simple structure with a main clause frame of a basic character (VanEk and Trim, 1991: 197). At level B2, learners should produce complex sentences containing more than one embedded clause, 'given that the internal structure of clauses and the relations between them are not unduly complicated' (VanEk and Trim, 2001: 175).

2) What kinds of writing are relevant for each group of language learners.

3) What are the criteria according to which, e.g. the language use, should be assessed.

These prerequisites are of utmost importance not only for the language learners but also for the assessors as writing is one of the most problematic areas of language use to assess. One of the reasons is that there are so many different situations in which people learn a foreign language (some do it as a part of an immersion programme, the others - to obtain a better job, students are obliged to learn a foreign language as it is stated by the national curriculum, etc.). Moreover, people come from different first language backgrounds, e.g. in Latvia students taking the examination can have Latvian, Russian, Lithuanian, etc. as their L1. Finally, assessing writing is partly subjective, which imposes yet greater responsibility upon the assessors to ensure reliable judgment.

Weir (1990) distinguishes two approaches for assessing writing ability. In the first one he divides writing into discrete levels, e.g., grammar, vocabulary, spelling and punctuation, which could be tested in objective tests, also referred to as 'indirect testing' (e.g. Alderson and Banerjee, 2002). In the second approach (direct testing) more direct extended writing tasks of various types have been envisaged, 'which would have greater construct, content, face and washback validity but would require a more subjective assessment' (Weir, 1990: 58). Accordingly, three fundamental concepts of assessment: validity, reliability and feasibility, should be discussed, as they are an integral part in the judgement of the quality of a performance.

1.2.1 Reliability of Written Assessment

Reliability is considered to be the most important measurement of quality when assessing the test-taker written performance, since it is the basis for making inferences about the

candidate abilities. Reliability is defined as 'consistency of measurement across different characteristics or facets of a testing situation, such as different prompts and different raters' (Weigle, 2002: 49). When assessing writing we cannot exclude the subjectivity factor; hence, the heaviest weight lies on the raters (also referred to as markers, scorers, assessors and judges), how consistent they are in their judgments. If candidates receive the same score from one rater to another and if they are ordered in the same way by different raters, on different occasions and situations, then the test is considered to be reliable. The most crucial concern is how to guarantee agreement between the markers, which is known as inter-rater reliability. It could be enhanced by establishing assessment criteria, which the markers should be trained to apply through standardization sessions. Moreover, when assessing written performance, there 'needs to be a degree of cross-checking to ensure that agreed standards are being maintained' (Weir, 1990: 32). However, we should admit that it is almost impossible to completely avoid inconsistency, as there will always be more lenient or more severe markers. Withal, reliability is an indispensable prerequisite for test validity.

1.2.2 Validity of Written Assessment

When developing a test one of the main concerns is to guarantee the *validity* of the test scores. Henning (1987) defines validity of a test as 'the appropriateness of a given test or any of its component parts as a measure of what it is purported to measure' (1987: 89).

Traditionally *validity* has been classified into three types: content, criterion, and construct validity (Hamp-Lyons (1990) also mentions face validity, which implies purely subjective judgment, i.e. 'what looks to an intelligent outsider as if it is valid' (Hamp-Lyons, 2003: 2). By *content validity* we understand how representative the test of the respective language skills is; the *criterion validity* implies the comparison of the test-taker scores with the scores obtained in other tests when measuring the student language ability (known as *correlation*) and other *criterial* measures; *construct validity* is directly referred to the meaning of the test scores, i.e. how the results of the test-taker could be interpreted. The American psychometricians (e.g. Bachman and Palmer) hold a more integrated view of validity and regard it as a unitary concept, which, according to Messick (1989), implies both empirical evidence and theoretical rationales in supporting '*the adequacy* and *appropriateness* of *interference*s and *actions* based on test scores' (1989: 13).

Hence, validity is regarded as a feature of the inferences based upon test scores and test use, and a more unified term *construct validity* has replaced the previous ones (by

construct understanding 'the specific definition of an ability that provides basis for a given test or test task and for interpreting scores derived from the task' whereas by *construct validity* 'the extent to which we can interpret a given test score as an indicator of the ability(ies), or construct(s), we want to measure' (Bachman and Palmer 1996: 21)). Thus, it is important to comprehend what ability, e.g., the ability to write, and to what extent the test is actually measuring it.

Validation also depends on the ability of interest of a particular testing context. According to Weigle (2002: 50), the ability of interest may be defined in a number of ways, e.g., 'the ability to generate well-formed sentences' or 'the ability to persuade an audience through the selection of appropriate evidence, tone, and rhetorical strategies.' To sum up the aforementioned aspects, it should be concluded that 'a test or assessment procedure can be said to have validity to the degree that it can be demonstrated that what is actually assessed (the construct) is what, in the context concerned, *should* be assessed, and that the information gained is an accurate representation of the proficiency of the candidate(s) concerned' (CEFR, 2001: 177).

What is more, Weir (1990) strongly believes that 'a test can be valid if it is also reliable' (1990: 33); therefore, he suggests 'to sacrifice a degree of reliability in order to enhance validity' (ibid.), since if the test cannot be regarded as valid, we are measuring not the things that we want to measure.

Therefore, in order to reduce the subjectivity and increase the validity and reliability, the following steps that should be taken into consideration have been listed in the CEFR (2001: 188):

- *a specification for the content* of the assessment should be developed;
- *pooled judgments* to select content and/or to rate performances should be used;
- *standard procedures* should be adopted;
- specific *defined criteria* should be provided;
- *multiple judgments* should be required;
- the quality of the assessment should be checked by *analyzing assessment data.*

The CEFR also accentuates that the most important aspect is *the accuracy of decisions made* in relation to the standard (e.g., the Foreign Language Standard for Secondary Education). This accuracy depends on the validity of the particular standard for the context. Moreover, the accuracy will depend on the *validity of criteria* used and the *validity of procedures* with which these criteria were developed.

However, over the years the CEFR level descriptors as well as the grammatical and lexical features elaborated by VanEk and Trim in the 1990s to describe different English

proficiency levels appeared to be 'underspecified' with respect to main properties that the assessors searched for when assigning the test-takers a particular level of proficiency (Hawkins and Filipović, 2012: 5). Therefore, the EPP was set out by using the resources of the Cambridge Learner Corpus (CLC) to provide 'reference level descriptions'. In such a way grammatical and lexical details would be added to CEFR's functional characterisation of the different levels. Thus, the researchers tried to find those properties that were typical for each foreign language proficiency level and could be used to distinguish between levels. As a result grammatical criterial features (*criteriality* means that 'the language features concerned serve as a basis for distinguishing one proficiency level from another' (English Profile, 2011: 6)) for each level of EFL were elaborated, which clearly stated, e.g. which syntactic structures were to be searched for in student written performance so that the assessors could attribute appropriate language proficiency levels to the test-takers. Since the research was accomplished only in English, the criterial features cannot be used when assessing student written performance cross-linguistically. Therefore, the objective of the present research is to develop the syntactic criterial features applicable across languages, i.e. English and French, which will ensure the reliability and validity of the written test results across languages.

1.2.3 Feasibility

The third concept to be examined is the *feasibility or practicality* of the test, since without it the test has little reliability and validity. The test could be regarded as practical only 'if the resources available for test-related activities meet or exceed the resources required to develop and/or administer the test' (Weigle, 2002: 56). For example, in assessing large-scale writing one should be aware of the fact that the number of tasks to be accomplished is limited because of time constraints. Besides, the scoring of writing is time-consuming and demands a lot of assessors. Therefore, the availability of sufficient resources is of great importance.

What test developers have to consider is all the aforementioned aspects, i.e. reliability, validity and feasibility, to be able to determine an appropriate balance among them in each particular situation. Moreover, test developers should regard the process of validation 'as ongoing, as the continuous monitoring and updating of relevant information […] as a process that is never complete' (Alderson and Banerjee, 2002: 79).

1.2.4 Relating Writing to the CEFR Levels

Nowadays, when taking the centralised examination in a foreign language in Latvia, writing is assessed according to the *performance level descriptors* provided by the CEFR, which enable the relation of the language examination results to the six language competence levels: A1 - A2 - B1 - B2 - C1 - C2 (A1 refers to the basic level of language proficiency while C2 describes the language competence at academic level). In many European countries, also in Latvia, these levels serve as the aims in the curricula, which the students have to attain. The Regulations of the Cabinet of Ministers Nr.715 (September 2, 2008) envisage that the students learning a foreign language from form 10 of upper secondary school are to attain level B1 when graduating from secondary school, whereas those learning a foreign language from the primary or lower secondary school are to have the proficiency levels B2 and C1. For many years (also in 2009 when the present research was initiated) the test-taker language proficiency was assessed according to levels A – F (A being the highest and F – the lowest level of language proficiency). In order to ensure the relation of the examination results to the CEFR it was decided to implement competence levels B1 – C1 as the basic descriptors for assessing language learner proficiency level cross-linguistically starting from September 1, 2012 (VISC, 2011b).

Relating foreign language examinations to the CEFR entails the comparison 'between different systems of qualification' (Council of Europe, 2009: 3), which is a complex process as both theoretical and empirical data are to be provided. It is basically 'a judgmental method of linking in which the results of one assessment can be compared to another by linking both separately on the same proficiency level, i.e. the CEFR scales' (Council of Europe, 2003: 5). This procedure is also known as the 'validation of the claim' (Council of Europe, 2009: 7). The aim of the linking process is 'to enable a categorization of test takers in terms of the proficiency levels of the CEFR, in such a way that this categorization reflects in a truthful way what is meant by the CEFR' (ibid.: 36).

The process of relating language examinations to the CEFR is based on four stages: familiarization, specification, standardization (training, benchmarking, standard setting) and finally validation (see Appendix 4), which is the most important aspect in this process as it is directly linked to the quality of the examination (linking to the CEFR is valid if the examination itself proves to be valid).

Familiarization is the starting point in the process of relating a language examination to the CEFR during which the persons involved are to obtain a profound knowledge of the CEFR criterion descriptors as well as of all the rating activities and

instruments used when accomplishing the set task. The outcome of this stage is essential because 'if those involved in the CEFR have only superficial knowledge of the CEFR and exhibit low agreement with its levels, then the quality of the subsequent stage is suspect' (Figueras et al., 2005: 268).

Specification is a qualitative method to validate the content of an examination in relation to the CEFR 'from the point of view of their coverage' (Council of Europe, 2003: 29). It aims at increasing awareness among developers of quality language examinations and defining minimum standards in terms of both quality of content specification in language examinations and the process of linking examinations to the CEF (ibid.: 29).

Papageorgiou (2009: 12) regards the *standardization* as 'transition from test content description to description of examinee performance in relation to the CEFR.' This stage is of particular importance as it ensures the reliability of the examination, i.e. new forms of examination do not differ in content and difficulty and the results do not vary among test users (Alderson and Banerjee, 2002). The standardization comprises three phases: training, benchmarking and standard-setting (Papageorgiou, 2009: 12). In the training phase the assessors use already calibrated samples available from the Council of Europe, thus ensuring the ability to relate the samples to the CEFR. In the benchmarking phase, which is the procedure applicable for standard setting only, the raters utilize samples from the examination to be linked to the CEFR. Kaftandjieva (2004: 2) describes benchmarking as 'a special case of standard setting procedure, which requires no cut-off point establishments,' which is not the case with the standard setting phase where the main aim is to establish the cut-off points that distinguish between different CEFR levels.

When linking the examinations to the CEFR, the standards are to be set, which are also referred to as 'Performance Level Descriptors' or 'cut scores' (Council of Europe, 2009: 7), which is understood as 'a border between two adjacent categories on some scale' (ibid.: 58). We can distinguish between criterion-referenced testing, in which the reference should be made to 'specified rating criteria' (Allison, 1999: 175), and norm referenced-testing, i.e. reference should be made to 'work produced by other learners in the same or comparable setting' (ibid.). There might also be a third option when the levels of performance are described by comparing both these forms. In Latvia test-administrators apply the criterion-referenced testing, which implies such criteria as: content, organization, grammar, vocabulary and orthography. Thus, a detailed description of each level of performance is an important prerequisite when allocating the student performance to a particular CEFR level as, e.g. a cut score for B2 is 'the minimum score that will lead to the

decision/classification that the ability of the candidate is at level B2 or higher' (Council of Europe, 2009: 58). If the candidate obtains lower scores, then he/she will be attributed a lower level than B2.

For *standard setting* the decision has to be made within a group (also called *a panel*) whose members have undergone appropriate training beforehand, which comprises illustration, controlled practice and individual assessment. There are several standard setting methods - the extended Tucker-Angoff method, the contrasting group method, the borderline group method, the body of work method, the item-descriptor matching method and the basket method. The most recent ones are the item-descriptor matching method proposed by Ferrara, Perie and Johnson in 2002 which requires the panelists to decide to which level category each item belongs (B1, B2, etc.), and the basket method which was especially developed for linking language tests to the CEFR, and which was also used by Alderson in the Dialang project (2005). The latter focuses on the capacities at a certain level of an abstract candidate, i.e. at what level an examinee can answer the particular item correctly. The group members have to decide to what level each item corresponds; if an item corresponds to level B1, then a test-taker should be able to answer this item correctly (or use a particular syntactic structure correctly) at level B1. Moreover, the person at a higher level of language proficiency should also be able to provide a correct response, which is not the case with a lower level at which the *correct* response is not expected.

This method is considered to be the most practical as 'practicability refers to the ease with which a standard-setting method can be implemented, computed, and interpreted' (Berk, 1986: 143-144). However, it does not take into consideration the empirical information. Therefore, additional information should be provided which incorporates the empirical data as well. In this case the Bookmark method or the Item Descriptor Matching Method (Cizek and Bunch, 2007: 155-207) should be applied as the items are arranged in a certain order in terms of difficulty.

1.2.5 Cross-linguistic Standard Setting

Since there are several methods applicable when setting standards for allocating a certain level of performance, the most challenging aspect is to use the appropriate methods to show that the examinations in different languages, e.g. English and French can be linked to the CEFR in a comparable way applying the same standards since the first question to be asked is whether my B1 is equal to your B1, and only afterwards to find out the meaning of B1. This linking process is known as *cross-linguistic standard setting* (Council of Europe,

2009: 85), and it requires plurilingual panelists who are able to judge the candidates' performance in both languages. It means that the language and training background of the group members has to be considered, e.g. one half of the panelists are teachers of English who have a sufficient language proficiency level in French or vice versa. Another pre-requisite is to guarantee that all members will utilize the same standard, i.e. the same criteria to each language. Finally, the results of the bilingual standard setting procedure should be compared with the ones obtained from a monolingual procedure of independent experts in English and in French.

One of the pre-requisites in assuring the objective assessment of test-taker written performance cross-linguistically is to construct a *common measurement scale* and then set standards that will be applied equally to all languages. McNamara (1996) notes, 'the scale that is used in assessing performance tasks such as writing tests represents, implicitly or explicitly, the theoretical basis upon which the test is founded; that is, it embodies the test (or scale) developer's notion of what skills or abilities are being measured by the test' (Weigle, 2002: 109). Thus, the scale and the descriptors for each scale level are of great value for the validity of the assessment. Weigle (2002) distinguishes four main types of rating scales: primary trait scales, holistic scales, analytical scales and multiple-trait scales, though holistic and analytical scales are most widely used.

Traditionally, when developing a rating scale, language ability has been considered as 'a single unitary ability, and yields a single score, called a 'global', or 'holistic' rating' (Bachman and Palmer, 2010: 339). In *holistic* assessment the judgment is based on the impression of the whole text, therefore it might be problematic to assign levels and differentiate weighting of components (different raters may consciously or unconsciously weigh the components differently as the judgments will be more subjective). Weigle distinguishes two major disadvantages of the holistic scoring:

> One drawback is that a single score does not provide useful diagnostic information about a person's writing ability [...] Another disadvantage of holistic scoring is that holistic scores are not always easy to interpret, as raters do not necessarily use the same criteria to arrive at the same scores [...] Holistic scores have also been shown to correlate with relatively superficial characteristics such as length and handwriting [...] Holistic scoring has also come under criticism in recent years for its focus on achieving high inter rater reliability at the expense of validity (2002: 114).

Nevertheless, holistic scales are preferred where the main concern is with evaluating the communicative effectiveness of candidates' writing in different languages.

In contrast, *analytic scales* are based on two principles: 1) the specific construct, which has been defined beforehand, is to be measured; 2) the levels are criterion referenced in different areas of language ability. Bachman and Palmer (2010) envisage two

practical advantages of analytical scales. First, they allow the test developer to provide a 'profile' of the areas of language ability that are rated. For example, in writing the candidate might organize the text well but does not have much control of registers of formal writing. A second advantage is that they reflect what raters actually do when rating samples of language use such as grammar, vocabulary and content. Hence, analytical scales provide more detailed information about a candidate's performance and are more applicable in assessing written performance. In Latvia test-developers have elaborated a common marking scale for English and French in compliance with the CEFR level descriptors, and it has been utilized since 2012 for assessing student written performance cross-linguistically (see Appendix 8).

The final stage in the test relation to the CEFR implies *empirical validation,* which comprises internal and external validation. Internal validation provides evidence on the quality of the examination as test developers need to ensure that test-takers receive 'valid, reliable and fair test scores' (Papageorgiou, 2009: 13), which is a pre-requisite for relating the examination to the CEFR. In external validation the cut-off scores set in the standardization stage need to be confirmed empirically. It aims at 'providing evidence from *independent sources* which corroborate the results and conclusions of one's own procedures' (Council of Europe, 2009: 108). The evidence can be obtained in two ways: 1) from the results of the same candidate in another examination or from the results of other candidates in the same or another examination; 2) from another standard setting procedure in which the same panel and the same staff is involved or in which another panel and another staff contribute. It should be stated that the process of external validation is complex and time consuming, and that 'the result is never perfect, due to measurement error in the test and residual variance in the judgment of the panel members' (ibid.: 108). Moreover, external validation is beyond the focus of the present research.

Writing is regarded to be the easiest skill for cross-linguistic alignment because the raters who are proficient in two foreign languages (English and French) can directly compare the written performance of test-takers in the two corresponding languages (SurveyLang, 2011). What has to be pointed out is that the comparison can be done only of similar writing tasks, e.g. essays. In Latvia test administrators together with test developers have accomplished the first stage in relating the centralised examination in foreign languages to the CEFR, i.e. a common marking scale has been developed (Appendix 8) and is being used when assessing student written performance in English, French, Russian and German. Moreover, the written part contains the same task types irrespective of the

language (the theme of the essay may vary but the type of the essay, e.g. an opinion essay, is similar, which is determined by the State examination programme (VISC, 2012). What is still needed is a plurilingual group of experts and assessors, which would guarantee the reliability of judgment across languages pointed out also during the Surveylang project initiated by the European Commission to support the development of language learning policies across Europe. Thus, we may fully agree that 'the logic of a multilingual framework is such as to severely constrain the freedom of judgments relating to individual languages' (Jones, 2009: 38).

Conclusion

To conclude, written performance assessment is considered to be subjective assessment as it comprises both competence and performance (known also as communicative language ability). Since language use has proved to be one of the most problematic areas to assess, the markers should familiarize themselves with the stages of language learner grammatical development as defined by the CEFR. Moreover, they should apply the common marking scale with its level descriptors as well as the criterial features that allow distinguishing among different levels of linguistic proficiency. All these aspects will guarantee the objectivity of assessment at high stakes examination. Furthermore, the standardization procedure is of utmost importance as it ensures inter-rater reliability, which is a pre-requisite for validity. Finally, relating the writing test to the CEFR levels, which encompasses both cross-linguistic standard setting and empirical validation, will exclude variations in understanding the CEFR levels and the test-taker written performance will be judged along the same criteria irrespective of language.

As the CEFR provides a common basis for assessing written performance across languages, and the syntactic features not lexis have proved to be the discriminatory indicators of language proficiency levels, the next chapter will examine syntax and its major constituents – phrase, clause and sentence. Moreover, the comparison of different clause types as well as their functions and position within a complex sentence will be effectuated, since subordinate clauses are the ones that allow distinguishing the lower levels of language proficiency from the higher ones. The comparison will be based upon the rules and examples from English and French descriptive grammars (e.g. *Cambridge Grammar of English* (2006); *Longman Grammar of Spoken and Written English* (1999); *A Comprehensive Grammar of the English Language* (1985); *Grammaire méthodique du français* (1994); *Modern French Grammar* (2004); *Le Bon Usage* (2011), etc.), which have

been developed on the database retrieved from the contemporary spoken and written English and French native speaker and learner corpora.

CHAPTER 2 SYNTACTIC COMPLEXITY OF WRITTEN TEXTS

For a more profound comprehension of how language operates, it can be broken into its components. According to Bloom and Lahey (1978), language consists of: form, content and use. Form comprises syntax, morphology and phonology, whereas content is related to semantics, and use to pragmatics, thus constituting the basic rule system of a language.
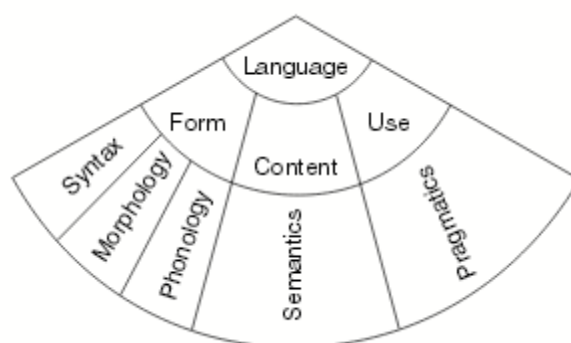
Figure 6 **Components of Language (Owens, 1992: 15)**

For the native language learners it is natural to produce sentences without paying any attention to the rules underlying each phrase they produce. Saville-Troike (2006: 32) postulates that all natural languages are systemic - there is an infinite number of sentences in all languages and the majority of these sentences are produced unconsciously according to certain rules or principles. Traditionally, language study was concerned with form only, in which syntax had a significant role as it provided a set of rules that could be used in producing an infinite number of sentences. However, in order to ensure communication and interaction among humans, the language learners are supposed to learn how words and phrases form sentences in order to produce logical and complex utterances. Moreover, the student written performance (the main objective of the present research) principally encompasses linguistic knowledge, vocabulary, syntactic patterns and cohesive devices. Moreover, student written performance in Latvia is assessed according to five parameters (one of them being grammar defined as 'a finite set of underlying operational principles or rules that describe the relationships between symbols that form the structure of a language' (Owens, 1992:11)). Therefore, the present chapter *aims* at disclosing the system of rules, i.e. linguistic competence, by defining 'syntax' and examining the syntactic complexity of English and French starting from the smallest constituents – phrases (focusing on their form and function within a clause), and ending with the most complex ones – sentences.

2.1 Syntax

There have been various definitions or interpretations of what the term 'syntax' means. Traditionally, syntax refers to the branch of *grammar* dealing with the ways in which words, with or without appropriate inflections, are arranged to show connections of meaning within the sentence (Matthews, 1981:1). However, the origin of the term 'syntax' can be traced back to Ancient Greek, where *sýntaxis* literally means 'arrangement' or 'setting out together' (ibid.).

Chomsky defines syntax as 'the study of the principles and processes by which sentences are constructed in particular languages' (2002: 11). In his Principles and Parameters approach to syntax (1981; 1995), he attempts to combine the UG theory with the formal grammar that focuses on the structural studies of a particular language. Givón (1979) looking from the discourse perspective considers that 'syntax is not an independent, autonomous level of sentence organization, but rather a dependent, functionally motivated entity, whose formal properties reflect […] the properties of the explanatory parameters which motivate its rise' (1979: 208).

Modern linguists, e.g. Haegeman (2006), define syntax as a 'branch of linguistics that concentrates on the formation of sentences' (2006: 4); Hawkins (2006) describes syntax as 'a set of properties which determine the construction of sentences in that language' (2006: 2). According to Carnie, 'syntax […] studies the level of Language that lies between words and the meaning of utterances: sentences. It is the level that mediates between sounds that someone produces (organized into words) and what they intended to say' (2007: 4). Liddicoat (2007) states that syntax deals with how to put words together to form sentences which mean what we want. A similar definition is given by Tallerman (2005) where syntax means 'sentence construction': how words group together to make phrases and sentences. Matthews (1997) adds the grammatical relations to his definition of syntax, namely, 'syntax is the study of grammatical relations between words and other units within the sentence' (1997: 397). Thus, all the above mentioned definitions allow the author to conclude that words are 'the basic tissue of syntax' (Teubert and Cermáková, 2007: 5), and that syntax is a cover term for studying how to construct sentences.

As for language learners, whose goal is to improve the linguistic knowledge of L2, the most important aspect is to comprehend which sentence is grammatical and which one is ungrammatical, e.g. *He is happy* has to be distinguished from the ungrammatical *He happy is* since it is assumed that language learners do not learn 'all possible word combinations. Instead, they learn rules that *govern* these combinations. Knowing the

linguistic rules allows each language user to understand and create an infinite variety of sentences' (Owens, 1992: 14) to ensure the successful communication in different contexts. However, O'Grady (2005) suggests that 'the theory of sentence structure can and should be unified with the theory of sentence processing' (2005: 47). Thus, when analysing the syntactic structure of a sentence, the linguist has to identify and determine to what category it belongs and what function it serves. Hence, the present framework for the research is based upon *integrative functionalist* or *formal-functional* interpretation of syntax. Moreover, language comparison (in our case – the comparison of syntax) aims at revealing those features that are shared by English and French as well as those that are language specific, thus providing a theoretical basis for the empirical studies of the syntactic structures produced by the test-takers. The results of this study are to be applied when developing the criterial features for each language proficiency level.

2.2 Sentence

The sentence is considered to be the highest unit in the grammatical hierarchy (see, e.g. Pienemann, 1998; Crystal, 2007). A dictionary usually refers to the *sentence* when defining syntax, for example, syntax is 'the study of grammatical relations between words and other units within the *sentence*' (Matthews, 1997). Similar definitions were already proposed by the antique scholars, for example, Priscian defines an 'oratio' ('sentence' more literally 'utterance') as 'a concordant ordering of words' ('ordination dictionum congrua') which 'expounds a complete idea' ('sententiam perfectam demonstrans') (Keil, 1855 in Matthews, 1981: 27). Thus, the main focus is to the thought expressed by the given sentence, for example:

> Don't disturb me! (the hearer shouldn't disturb the other person)
> *or*
> I'm reading. (the speaker is reading).

Matthews (1981) suggests defining a sentence as 'the maximal unit of syntax, or the largest unit over which constructional relations hold' (1981: 29). Le Goffic (1993) considers a sentence to be an independent sequence in which an utterer (speaker) highlights two terms, a subject and a predicate. Kroeger (2004: 5) representing a lexical-functional approach does not consider a sentence to be simply a string of words as 'speakers think of sentences as having a fairly complex structure, with certain words grouped together to function as units, larger groups formed from smaller groups, and important relationships defined

between one group and another.' Blokh (2003) in his definition also combines syntactic structures with their communicative applicability:

The sentence is the immediate integral unit of speech built up of words according to a definite syntactic pattern and distinguished by a contextually relevant communicative purpose. Any coherent connection of words having an informative destination is effected within the framework of the sentence (2003: 255).

Thus, from the perspective of grammatical theory the sentence is the central aspect of syntax since the latter is viewed as 'a device [...] for producing the sentences of the language under analysis' (Chomsky, 2002: vii). From the functionalist viewpoint it is built up out of different language units in the course of communication, i.e. speech process. Crystal (2007) claims that the aim is not to define a sentence but 'to analyse the linguistic constructions that occur, recognising the most independent of them as sentences' (2007: 249). Crystal proposes a five-level hierarchy as a model of analysing syntactic structures:

sentences
which are analysed into
clauses
which are analysed into
phrases
which are analysed into
words
which are analysed into
morphemes

morphemes
which are used to build
words
which are used to build
phrases
which are used to build
clauses
which are used to build
sentences
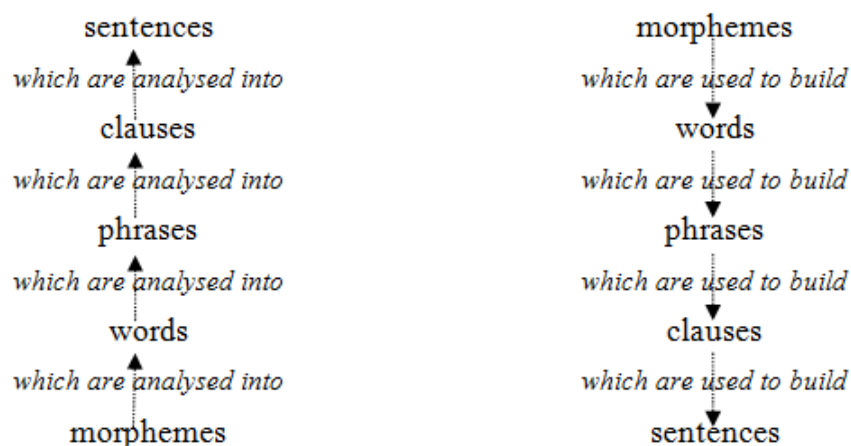
Figure 7 **Five-level hierarchy (Crystal, 2007: 251)**

Morphemes are considered to be the lowest element of grammatical enquiry, as they have no grammatical structure. Even so, they are the smallest linguistic element to which a meaning and function can be attributed. In Crystal's hierarchy sentences is the highest level of grammatical study for 'they usually do not form a part of any larger grammatical unit' (Crystal, 2007: 251).

Phrase level

Generative grammar does not clearly define what a sentence is, as for generative linguists syntax foremost is the study of sentence *structure*. By applying the structure-dependency principle, the linguistic competence relies not on the sequence of words but on the structural relationship within a sentence. Chomsky (1995: 5) claims that 'a sentence is not

just a string of words in a linear sequence but is structured into phrases, all of which connect together to make up the whole.' Thus, it is vital to understand the internal structure of a sentence and what associations this structure conveys in our minds. Blokh (2003: 17) divides syntax into 'minor syntax', which implies the phrase description, and 'major syntax' – the study of sentences. However, they are not mutually exclusive as phrases help us to understand how sentences have been built up by them. Thus, when analysing a sentence, linguists start with looking at different kinds of words, in particular the word's part of speech or syntactic category (noun, verb, adjective and preposition), which are grouped into units (called constituents – the basic tissues in syntactic theory), e.g. [the girl / la fille]. These units tell us how a word functions in the sentence and it is possible 'to make generalisations (scientific ones) about the behaviour of different word types' (Carnie, 2007: 38). Subsequently, the constituents are further grouped into the larger ones by forming a sentence in the end. It means that constituents are embedded inside one another by making a *hierarchical structure*. For example,
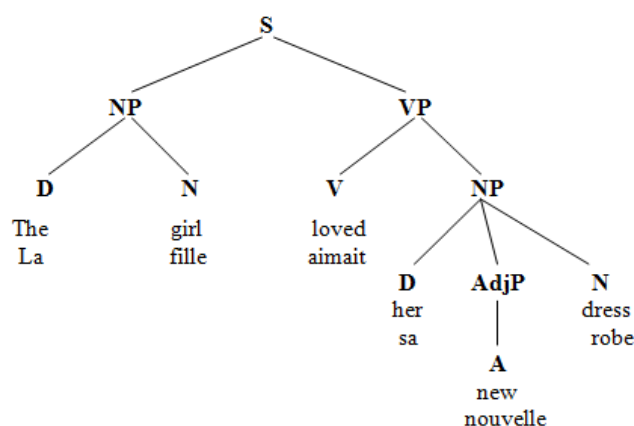
Figure 8 **The hierarchical structure of sentence constituents**

[s[NP The girl [VP loved her new dress]].
[s[NP La fille [VP aimait sa nouvelle robe]].

The sentence constituent (S) consists of two constituents: a subject noun phrase (NP) [*the girl/la fille*] and a predicate or verb phrase (VP) [*loved her new dress/aimait sa nouvelle robe*]. The NP contains a noun (N) [*girl/fille*] and a determiner (D) [*the/la*], while the VP contains a verb (V) [*loved/aimait*] and an object NP [*her new dress/sa nouvelle robe*]. The object NP is further split into a determiner [*her/sa*], adjective [*new/nouvelle*] and noun [*dress/robe*].

This syntactic aspect is known as a constituent structure or a *phrase structure*. In generative grammar, generalizations about structure are represented by rules. These rules are said to 'generate' the tree. The rules [...] are called *phrase structure rules* (PSRs) because they generate the phrase structure tree of a sentence. (Carnie, 2007: 66). The phrase is given its category by the head of a phrase, e.g. the head of the verb phrase (VP) is the verb (V). We can distinguish: 1) *noun phrases* (NPs) consisting of only a noun or a pronoun (N), e.g. *Mark*; a determiner and a noun (NP → D N), e.g. *the box / la boite*; a determiner, an adjective and a noun (NP → (D) (AdjP) N), e.g. *the big girl / la grande fille*; a determiner, an adjective, preposition and a noun (NP → (D) (AdjP) N (PP), e.g. *the best chocolate of Laima / le meilleur chocolat de Laima*; 2) *adjective phrases* (AdjPs) and *adverb phrases* (AdvPs), e.g. *very bright / très claire* represented as AdjP → (AdvP) Adj or *very quickly / très vite* represented as AdvP → (AdvP) Adv; 3) prepositional phrases (PPs) comprising a preposition followed by a noun phrase (PP → P NP), e.g. *behind the house / derrière la maison*; 4) *verb phrases* (VPs), which might consist of a single verb, e.g. *came /est venu* (VP → V), or a verb modified by adverbs, e.g. *came very fast / est venu très vite* (VP → V (AdvP)).

Thus, sentences are constructed according to particular phrase structure rules, which in their turn determine what the possible kinds of phrases are, what constituents might follow them, and finally all elements are arranged in linear order.

Moreover, lexical units may be freely placed in the generated structures 'forming the base component of the grammar [...] known as DEEP STRUCTURE, to be modified in various ways, to yield a SURFACE STRUCTURE' (Farrel, 2005: 174). Chomsky refers to this process as movement, which is 'not just a matter of recognizing phrases and then of moving them around but of moving the *right* element in the right phrase: movement depends on the structure of the sentence' (1996: 8). This principle is especially obvious when comparing active (A) and passive (P) sentences:

A   The girl may take the cake from the table. (deep structure)

P   The cake may be taken from the table by the girl. (surface structure)

A   Cette fille peut prendre le gâteau de la table. (deep structure)

P   Le gâteau peut être pris de la table par cette fille. (surface structure)
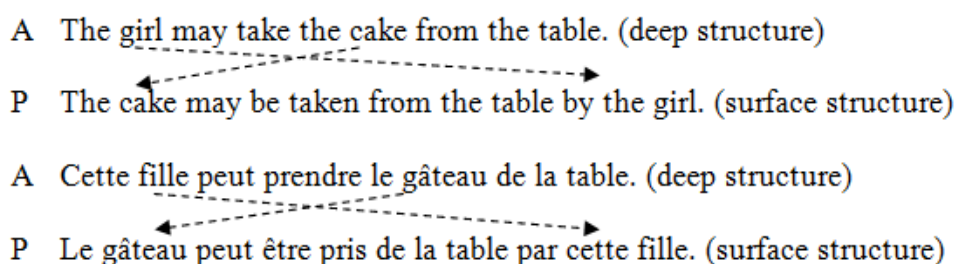
The above examples demonstrate that both English and French are structure-dependent languages; therefore, when analysing a sentence, we should look at the order of structural elements within a sentence.

The order of the elements that constitute phrases in a sentence is also of great value for syntactic pattern analysis as they may have different functions. Formally there are five categories of phrases that may have different functions in a clause, which are illustrated in the figure below.
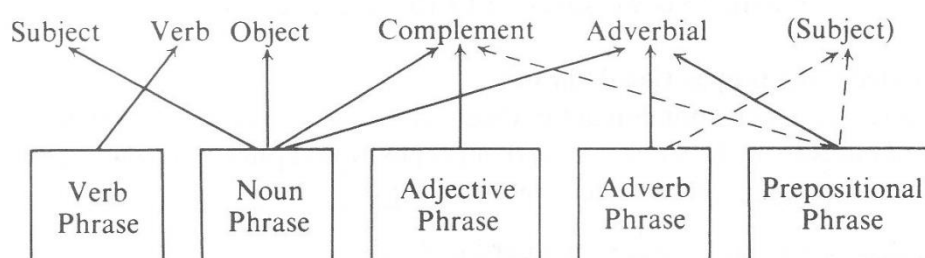
Figure 9 **Phrases as clause elements (Quirk et al., 1985: 60)**

Each phrase bears the name of the class of word which has the most important function within it. Consequently, the structure and functions of phrases within a clause are:

1) The verb phrase (VP) consists of a main verb, which might be preceded by auxiliary verbs (e.g. The cat *has* [eaten]. Le chat *a* [mangé]), and has only the function of a verb in a clause.

2) The noun phrase (NP) comprises a head (a noun) and the elements that determine and modify the head, e.g. I enjoyed *those sunny* [days]. J'aimais *ces* [jours] *ensoleillés*. Thus, *those sunny* determines and pre-modifies the head *days* in English, while in French *ces* determines *jours* and *ensoleillés* post-modifies the noun *jours*. In addition, the noun phrase may function as a subject, object, complement and adverbial.

3) The adjective phrase (AdjP) contains an adjective which may be preceded or followed by some modifiers, e.g. The day was *awfully* [cold]. La journée a été *terriblement* [froid]. Hence, *awfully/terriblement* modifies the head *cold/froid* and performs the function of a complement.

4) The adverb phrase (AdvP) consists of an adverb preceded or followed by its modifiers, e.g. She met him *quite* [often] *indeed.* Elle l'a rencontré *vraiment assez*

[souvent]. The head *often/souvent* is modified by *quite/assez* and *indeed/vraiment*). In general the AdvP functions as an adverbial.

5) The prepositional phrase (PP) is constructed by a preposition followed by a prepositional complement – NP, e.g. He came [in] *the morning*. Il est venu [dans] *la matinée*. The PP *on/dans* is modified by the NP *the morning/ la matinée*. It should be noted that prepositional phrases might function as subjects and complements, which is rather rare.

All in all, the analysis of the phrase constituents and their functions within a clause is a necessary pre-requisite before initiating the further study of a higher level of linguistic development, since it provides the overall understanding of the basic tissues of a sentence.

Clause level

According to Crystal (2007), the next level in the syntactic structure hierarchy is allocated to the clause, which is considered to be 'a more clearly defined unit than a sentence' (Quirk et al.,1985: 47), as it 'expresses a finished predication' (Holvūts, 2012: 132) and is 'structured around a verb phrase [...] accompanied by one or more elements which denote the participants involved in the action, state, etc.[...], the attendant circumstances [...], the attitude of the [...] writer to the message, the relationship of the clause to the surrounding structures, etc. Together with the verb phrase, these are the **clause elements** [...] realised by phrases or by embedded clauses' (Biber et al., 1999: 120). Thus, when analysing the clause structure the distinction between form and function has to be made to identify the main elements of clause structure: subject (S), verb (V), object (O), complement (C), and adverbial (A).

Generally, the clause consists of the verb preceded by the subject. The verb might be followed by one or two objects. Object is further subdivided into direct and indirect object; complement – into subject complement and object complement; adverbial – into subject-related and object-related. Quirk et al. (1985) establish seven major clause types (see *Table 8*) which are based on the allowable combinations of the clause elements and their functions. Thus, we may have a clause containing only two elements – subject and verb (e.g. The sun is shining; Le soleil brille) or the clause may be enlarged by adding objects, complements and adverbials. For example, the clause [*I must send my parents an anniversary card*] / [*Je dois envoyer mes parents une carte d'anniversaire*] comprises the

subject *I / Je*; the verb *must send / dois envoyer*; direct object *my parents / mes parents* and indirect object *an anniversary card / une carte d'anniversaire.*

Table 8 **Major clause types (adapted from Quirk et al., 1985)**

	Subject	Verb	Object	Complement	Adverbial
SV	The sun Le soleil	is shining brille			
SVO[1]	That lecture Cette lecture	bored ennuyait	*direct object* the students les étudiants		
SVC	Your dinner Votre dîner	seems semble		*subject complement* ready prêt	
SVA	My office Mon bureau	is est			*subject-related adverbial* in the next building dans l'immeuble voisin
SVOO	I Je	must send dois envoyer	*direct object* my parents mes parents *indirect object* an anniversary card une carte d'anniversaire		
SVOC[2]	Most students La plupart d'étudiants	have found ont trouvé	this rule cette règle	*object complement* reasonably helpful raisonnablement utile	
SVOA	You Vous	can put pouvez mettre	the dish le plat		*object-related adverbial* on the table sur la table

[1,2] In French there is a different word order if the direct or indirect object is expressed by a clitic pronoun.

It should be pointed out that the verb is considered to be the 'central' element which is preceded by the subject and followed by one or two objects, or a complement as '(i) its position is normally medial rather than initial or final; (ii) it is normally obligatory; (iii) it cannot normally be moved to a different position in the clause; and (iv) it helps to determine what other elements must occur' (Quirk et al., 1985: 50). Thus, in declarative sentences the regular order is SVO. The adverbial, on the contrary, can appear in front of the subject or after the verb, and after the object or complement. Therefore, '(i) their position is most frequently final; (ii) they are usually optional: (iii) they are mostly mobile; and (iv) they do not determine what other elements occur' (ibid.: 50).

In general, the regular order of clause elements in English and French is the same – SVO (see *Table 8*). However, if we examine the direct object, its placement is different in relation to the main verb. In English the direct object follows the verb (SVO), while in French it may occupy either the initial or final position in relations to the verb (SVO or

SOV) depending on whether the object is a clitic pronoun (i.e. the pronoun attached to the first verb of the clause) or another part of speech. If the object (direct/indirect) is expressed by a clitic pronoun, it will precede the verb in the sentence. Moreover, there is a special order of clitic pronouns in which they must appear before the verb in declarative sentences:

(1) La plupart d'étudiants *l'*[ont trouvé] raisonnablement utile.

The direct object expressed by the clitic pronoun *l'* is placed before the verb *ont trouvé*, thus having the combination SOV. If the direct or indirect object is another part of speech, then it will occupy verb final position:

(2) La plupart d'étudiants [ont trouvé] *cette règle* raisonnablement utile.

The direct object *cette règle*, which is expressed by the demonstrative pronoun and the noun, appears after the verb *ont trouvé*, thus forming SVO combination.

As for the position of adverbs in French and English, the only difference will be observed when used with the simple tense form verb. In French the adverb will be placed immediately after the verb or at the end of the VP, after any complements, whereas in English it will occupy the verb initial position or as in French – will be placed at the end of the VP, after any complements:

a) Christine [apprendra] *très vite* cette chanson.   SVAO
   Christine [apprendra cette chanson] *très vite.*   SVOA

b) Christine *very quickly* [learnt] this song.   SAVO
   Christine [learnt this song] *very quickly*.   SVOA

Thus, the only difference is observed between the first variant of the above examples, i.e. in French the order of clause elements is SVAO while in English – SAVO.

Although clause elements might occupy different positions in a sentence, which sometimes do not correspond in English and French (e.g. adjectives can either occupy the position before a noun or after a noun in French), both languages belong to the group of fixed-order languages. In 1783 Antoine de Rivarol when praising the French language in his essay 'French is (in)corruptible' wrote:

> What distinguishes our language from other ancient and modern languages is the order and the structure of the phrase. This order must always be direct and necessarily clear […] there is the logic that is natural to all men; there is what constitutes common sense. […] The French language, by unique privilege, is the only one to remain faithful to the straightforward order […] French syntax is incorruptible […] What is not clear is not French; […] to learn the French language, we must also retain the word order (De Rivarol, 1784: 49).

Having examined the form and function of the clause constituents, it may be concluded that each element is allotted a definite place within a clause in relation to the

main element – the verb, which in the majority of cases is preceded by the subject and followed by the object or the complement of the clause. Moreover, the order of the clause elements is rather stable, except for the movements in French due to the different placement of clitic pronouns with respect to the main verb. In order to comprehend how clauses constitute various types of sentences as well as what functions they perform within a sentence, we will now turn to the analysis of subordinate clause types and their function, along with their placement in relation to the main clause within a sentence.

Simple sentence

We might distinguish different models of sentence classification. A majority of French grammarians (e.g. Riegel, 1994; Kannas, 2003; Delatour et al., 2004; Grevisse et Goosse, 1993) divides sentences into simple and complex depending on how many clauses they contain. Carter and McCarthy (2006: 488-489) classify sentences into simple, compound and complex depending on whether the sentence contains only a main clause or also a subordinate clause. Quirk et al. (1985: 719) distinguish between simple (sentences consisting of a single independent clause) and multiple sentences (compound and complex sentences which consist of more than one clause). Though all the models arrive at the complex sentence division into compound and complex sentences, within the framework of the present research the model proposed by Quirk et al. (1985) will be relied upon.

The most obvious kind of clause is the *simple* sentence. According to Jackson (2007), a simple sentence is a sentence composed of a single main clause (also called 'matrix clause'), which means that simple sentences stand alone, and are not attached to any other clause; they are independent. An independent clause contains a finite verb, which is marked to show tense, person, number or aspect, for example, in the sentence *John waited* (*John attendait*) the verb *waited* expresses tense information in English, whereas in French along with tense information the verb *attendait* expresses number and person, which is third person singular. In English and French only one verb in each clause can be finite, whether it is an auxiliary or a main verb. As for the structure, all the elements of a sentence are arranged in a hierarchy in which each of them has some modifying role (see *Figures 10* and *11*).
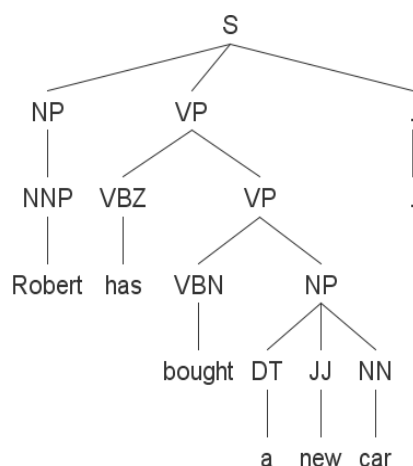
Figure 10 **Structure of a simple sentence in English**
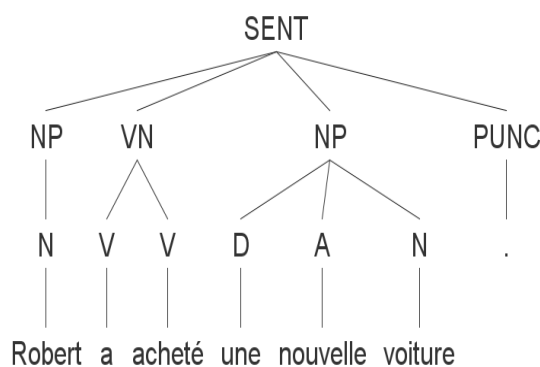
Figure 11 **Structure of a simple sentence in French**

In *Figure 10* the sentence consists of the NP [Robert]; the VP [has bought] and another NP [a new car] in the head [car] is pre-modified by the determiner *a* and adjective *new*. In *Figure 11* the sentence structure is similar, though the presentation of the tree diagramme differs because of the Stanford NLP parser's peculiarities, i.e. the syntactic tags depend on the Treebank used for each language. Thus, for the sentence parsing in English the Penn Treebank has been applied, whereas in French – the French Treebank. However, it does not hinder us performing the contrastive analysis of the syntactic structures in English and French as the position of the main elements (S, V, O) within a clause are similar, which was accentuated previously.

2.3 Multiple Sentence

A multiple sentence consists of more than one clause. When analysing multiple sentences, we have to examine how clauses are combined to form either *compound* or *complex* sentences.

The most common grammar reference books (e.g. *Cambridge Grammar of the English, 2006*) offer only two analytic tools – *coordination* and *subordination* – to describe the syntactic hierarchical relations, which are defined by using morphosyntactic criteria: embedding and non-finite verb forms, such as infinitives, gerunds, etc. Quirk et al. (1985: 918) state that 'in coordination the units are constituents at the same level of constituent structure, whereas in subordination they form a hierarchy, the subordinate unit being a constituent of the superordinate unit […]' Thus, coordination and subordination are special cases of two types of syntactic arrangement traditionally known as PARATAXIS ('equal arrangement') and HYPOTAXIS ('underneath arrangement')' (ibid.: 918).

Coordination

In *coordination* the clauses act as independent or main clauses, i.e. syntactically they do not depend on any other clause, they have an equal status 'bound together at the same level of structure by means of a linking device' (Dik, 1997: 189). In coordination the coordinators are constrained to the initial position of a clause, e.g. *He plays basketball, and his brother plays football (Il joue au basketball, et son frère joue au football)*. Moreover, in the majority of cases the clauses having the coordinators *and (et)*, *or (ou)*, and *but (mais)* are consecutively related to the preceding clause. Therefore, unless changing the relationship between them, they cannot be interchanged, e.g. *They are enjoying the opera, or they are chatting* (*Ils apprécient l'opéra, ou ils bavardent*) cannot be transposed as *Or they are chatting, they are enjoying the opera* (*Ou ils bavardent, ils apprécient l'opéra)*. What is more, the coordinators cannot be preceded by another conjunction, e.g. *yet/pourtant, so/donc, nor/ni*, etc.). Coordinators can link:

1) main clause + main clause

$_1$[*He wants to have his own house*] and $_2$[*he is willing to build it himself*].

$_1$[*Il veut avoir sa propre maison*] et $_2$[*il est prêt à construire lui-même*].

2) subordinate clause + subordinate clause

$_1$[*If you like this place*], and $_2$[*if you want to buy it*], we can negotiate the price.

$_1$[*Si vous aimez ce lieu*], et $_2$[*si vous voulez l'acheter*], nous pouvons négocier le prix.

3) more than two clauses

$_1$[Tom played basketball], $_2$[John worked in the garden], *but* $_3$[Tim helped his mother].

$_1$[Tom jouait au basketball], $_2$[John travaillait dans le jardin], *mais* $_3$[Tim aidait sa mère].

However, it is possible to have coordination without a particular conjunction, especially in literary style, in which case only commas draw a borderline between coordinate clauses.

Subordination

In *subordination* a dependent clause is not able to form a sentence as it depends on the main clause. According to Quirk et al. (1985), subordination is 'an assymetrical relation: the sentence and its subordinate clauses are in a hypotactic relationship' (1985: 987). By hypotactic relationship we understand the arrangement of the sentence constituents in a hierarchy in which the subordinate clause occupies a particular place within a sentence.

Thus, we distinguish between *subordinate* clause and *superordinate* clause (ibid.: 988) in a complex sentence. For example:

a[David declared b[that he would be back c[when he had finished his studies]c]b]a

a[David a déclaré b[qu'il reviendrait c[quand il aurait fini ses études]c]b]a
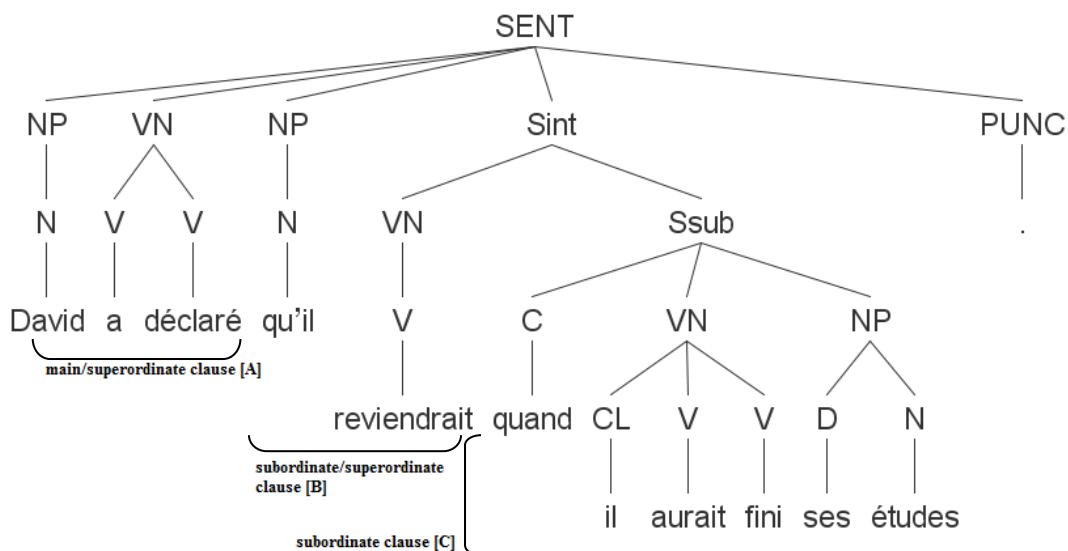
Figure 12 **The structural presentation of subordination**

The main clause [A] contains the subordinate clauses [B] and [C], where clause [C] functioning as adverbial is subordinate or immediately superordinate to [B] while clause [B] functioning as object is subordinate to [A]. Clause [A], in its turn, is immediately superordinate to [B], which is further superordinate to [C].

Hence, we can say that as clauses do not always stand on their own, they are embedded inside another. Traditionally, when identifying a subordinate clause, we usually take into consideration the two main concepts: *dependency* and *embedding*. Dependency is 'the impossibility of a clause occurring in isolation [...] usually indicated by phenomena such as the presence of particular conjunctions such as English *that* or *because*' (Lyons, 1968: 178), or morphosyntactic reduction (e.g. elimination of tense, aspect, mood or agreement inflection from the dependent verb) (Langacker, 1991: 439-45). There are cases when embedded constructions are joined to the main clause by no subordinator:

> I think *(that) Noam is the guilty one.*
> Je pense *que Noam est le coupable*.

Such constructions are typical for the English language but do not occur across languages, thus, in French the conjunction *que* is essential as a clause linking device.

If we consider *dependency* as a deciphering factor for subordination, then in the sentence 'I can't buy this house [because I don't have enough money]' (Je ne peux pas acheter cette maison [parce que je n'ai pas d'argent]) the clause *because I don't have enough money* (*parce que je n'ai pas d'argent*) is subordinate as it cannot stand on its own, which is not the case with *I can't buy this house* (*Je ne peux pas acheter cette maison*). Sometimes semantic relations between the clauses are the ones which denote the dependency. For example, in a sentence *I think that she will succeed* (*Je pense qu'elle va réussir*) both clauses *I think* (*Je pense*) and *that she will succeed* (*qu'elle va réussir*) cannot stand on their own because semantically the complement clause *that she will succeed* (*qu'elle va réussir*), which describes the predicate in the main clause, refers to the main clause. However, no one doubts that *I think* (*Je pense*) is the main clause. These examples prove that dependency is not the only aspect to be used to determine the subordination.

Another concept is *embedding*, which was introduced within the scope of generative grammar, but nowadays is used also by non-generative linguists, e.g. Foley and Van Valin, 1984; Haspelmath, 1995; Dik, 1997. According to Dik (1997: 140), 'a subordinate clause is an embedded construction which, apart from the subordinate device which signals its dependent status, could also appear as a main clause.' Cristofaro (2003: 17) provides a simpler definition by stating that it is 'a clause functioning as a constituent of another clause':

s[I know s[it is my sister's watch s[I gave her on her birthday]s]s]s

s[Je sais s[que c'est la montre de ma sœur s[que je lui ai donné pour son anniversaire]s]s]s

Another name for embedded clause is *subordinate clause*, which is defined as 'a clause that does not normally occur on its own, but either in combination with a main clause to form a complex sentence or a part of another clause as an 'embedded' element' (Jackson, 2007: 54). As the number of embedded clauses is not limited, it is impossible to say how long a sentence might be. One thing is obvious - in a complex sentence there is just one main clause but an infinite number of subordinate clauses.

Dik (1997) has elaborated the taxonomy of embedded constructions, which will be relied on when analysing the syntactic structures in the learner written corpora as this taxonomy has 'quite general cross-linguistic validity' (Dik, 1997: 143).

In his taxonomy Dik distinguishes between finite and non-finite embedded constructions as all main clauses contain a finite verb, but this is not the case with all embedded clauses.
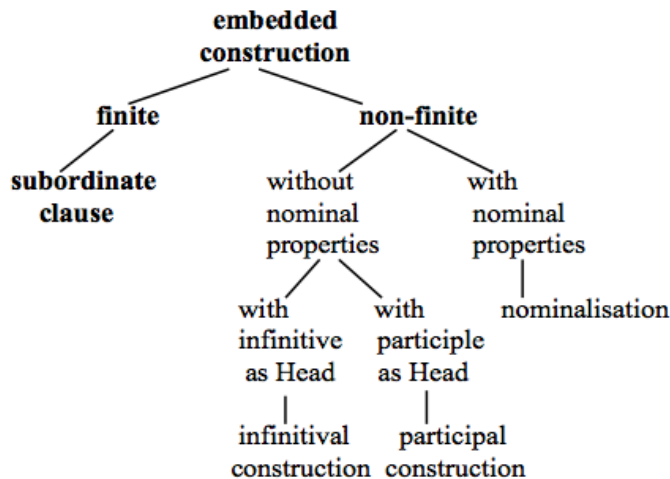
Figure 13 **Taxonomy of embedded constructions (Dik, 1997: 143)**

According to Dik (1997: 144), *finite embedded* constructions are 'those […] in which the predicate can be specified for the distinctions which are also characteristic of main clause predicates' while '*non-finite embedded* constructions contain a non-finite verb' (infinitive, participle, and nominalized verb), which cannot be used as a main verb of a clause. A more simplified definition is provided by Biber et al. (1999), which states that 'a finite dependent clause contains a verb which is marked for tense or modality' (Biber et al., 1999: 193) whereas 'non-finite clauses are regularly dependent [...] more compact and less explicit [...] they are not marked for tense and modality, and they frequently lack an explicit subject and subordinator' (ibid.: 198). M.A.Jones (1996), who studied French syntax, states that subjunctive and indicative forms are finite verb forms, which constitute finite clauses, whereas non-finite or infinitive forms ending in *–er, -ir, -re* or *–oir* build up a non-finite clause. What is more, only finite embedded constructions form subordinate clauses. This is obvious when the subordinate clause (without any changes) may function as an independent main clause (2).

1) People believe that *the economic situation will improve.*

   Les gens croient que *la situation économique va s'améliorer.*

2) *The economic situation will improve.*

   La situation économique va s'améliorer.

On the one hand Dik (1997) attributes the function of an independent clause to the finite subordinate clause, but on the other hand he rejects it because it is assumed that in

56

majority of cases subordinate clauses are marked by some subordinating device (Dik, 1997: 144).

Quirk et al. (1985) add one more clause type, thus distinguishing *three* structural types of clauses: *finite*, *non-finite* and *verbless* clause. They provide rather simplified definitions of the different embedded clauses. They define the finite clause as 'a clause whose verb element is finite', non-finite clause as 'a clause whose verb element is non-finite' and verbless clause as 'a clause that does not have a verb element, but is nevertheless capable of being analysed into clause element' (1985: 992). For example:

1) finite clause:

   I can't go out with you *because I am studying this evening*.
   Je ne peux pas sortir avec toi *parce que j'étudie ce soir*.

2) non-finite clause:

   Knowing my temper, I didn't reply.
   Connaissant mon caractère, je n'ai pas répondu.

3) verbless clause:

   Although always helpful, he was not much liked.
   Bien que toujours disponible, il n'était pas beaucoup aimé.

It is of utmost importance to recognise these structural types when analysing the internal structure of the clauses as only finite embedded constructions, defined by morphosyntactic criteria, are realised as subordinate clauses. However, one does not have to discard the structural types that do not contain a finite verb.

The studies on clause linkage have also recognised that each type may be 'more or less coordinate-like or subordinate-like depending on the parameter taken into account' (Cristofaro, 2003: 23). Foley and Van Valin (1984: 242) distinguish three basic types of relation based on embedding and dependency. Coordination involves neither dependency nor embedding; subordination involves both embedding and dependency; co-subordination involves dependency but not embedding, e.g. the adverbial clause in the sentence *I could not see him [because it was dark] (Je ne pouvais pas le voir [parce que il faisait noir])* is not embedded within the main clause, though in a way it is dependent on it. Cosme (2006) considers coordination and subordination to be universal. The only difference is that they are 'employed with varying frequencies across languages (2006: 99).

In many languages (including English and French) clause linkers (also referred to as connectives) are the most important means when denoting clause relations in syntax. Such word classes as conjunctions, adpositions, adverbs, relative pronouns, complementizers as

well as non-finite verb-forms serve as clause linkers to establish either coordinate or subordinate relations in a sentence.

2.3.1 Compound Sentence

One of the major types of multiple sentences is the compound sentence. '*Compound sentence is a sentence composed of at least two main clauses in a relation of coordination*' (Jackson, 2007: 29), for example, *Robert went to the cinema and his sister stayed at home (Robert est allé au cinéma et sa sœur est resteé à la maison)*. Moreover, 'the clauses of a compound sentence provide classic instances of a paratactic relationship, that is they have equal function' (Quirk et al., 1985: 987).
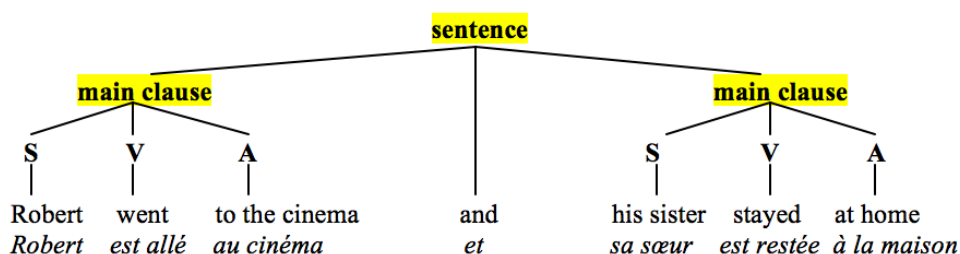
Figure 14 **The representation of paratactic relationship**

The above sentence consists of two main clauses [*Robert went to the cinema / Robert est allée au cinéma*] and [*his sister stayed at home / sa soeur est restée à la maison*] conjoined in paratactic relationship by the coordinating conjunction *and / et*.

When denoting the clause type, we have to be aware of clause linkers. 'The clauses of a coordinate structure are said to be **conjoined**, and each conjoined constituent is said to be a **conjunct**' (Langacker, 1972: 138). In English and French the most common coordinators are: *and/et, or/ou, but/mais* and they appear in the central position of the clause.

[Yesterday I went to the baker] *and* [I bought ten pies].

[Hier, je suis allé à la boulangerie] *et* [j'ai acheté dix croissants].

Along with the above mentioned coordinators, we also distinguish *for, nor, so, yet* in English and *ni, or, car, donc, pourtant, c'est pourquoi, d'ailleurs, puis* in French. Of course, there are many more, but they are more like conjunctive adverbs, e.g., *accordingly/en conséquence, likewise/également, besides/d'ailleurs, etc*. However, there are cases when no coordinator is used to link the clauses; the coordination is marked by commas. Such clauses are known as 'contact clauses' (Carter and McCarthy, 2006: 556).

2.3.2 Complex Sentence

A complex sentence is another major type of multiple sentence. There are various interpretations to what a complex sentence is. Lyons (1968: 178) considers a complex sentence to consist of 'those in which one of the clauses ('main clause') is 'modified' by one or more subordinate clauses grammatically dependent upon it and generally introduced by a subordinating conjunction.' Langacker (1972) proposes defining a complex sentence in terms of surface structure, i.e. consisting of more than one clause, or its underlying representations, i.e. 'as a constituent consisting of a predicate together with its assorted adjuncts, such as its subject, its object, adverbs, and so on' (1972: 137). Quirk et al. (1985) compare a complex sentence with a simple sentence as it contains only one main clause, but 'unlike a simple sentence it has one or more SUBORDINATE clauses functioning as an element of the sentence' (1985: 987). Blokh (2003: 333) defines a complex sentence as 'a polypredicative construction built up on the principle of subordination. It is derived from two or more *base sentences* one of which performs the role of a matrix in relation to the others, the *insert sentences*.' Thus, Blokh refers to the main clause as 'a base sentence' and to the subordinate clause as 'an insert sentence', which is not characteristic of other linguists when describing syntactic relations. What all the definitions do imply is that the main clause and the subordinate clause are integral constituents of a complex sentence. Thus, we can say that a *complex sentence* is 'a sentence composed of a main clause and at least one subordinate clause' (Jackson, 2007: 28).

When analysing a sentence consisting of one subordinate clause, first we analyse the sentence as a whole and then the subordinate clause, which is part of the sentence.
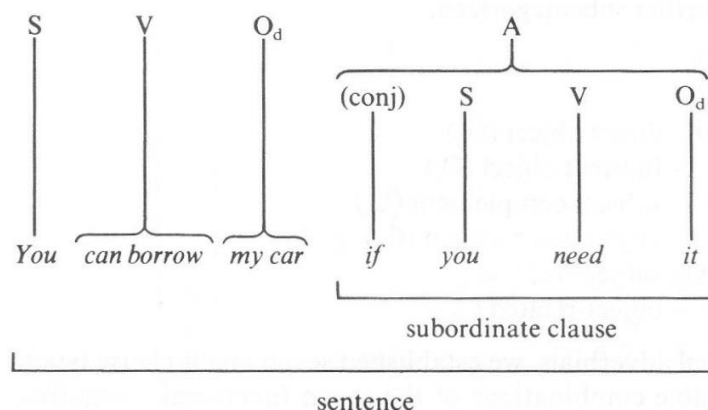
Figure 15 **Sentence and clause elements (Quirk et al., 1985: 719)**

Thus, the above sentence consists of the main clause [*You can borrow my car*] containing the subject [*You*] + verb [*can borrow*] + direct object [*my car*] and the

subordinate clause [*if you need it*] introduced by the conjunction [*if*] followed by the subjet [*you*] + verb [*need*] + direct object [*it*].

As for the order of subordinate clause constituents, it is said that 'in general subordinate clauses never have more freedom of constituent ordering than main clause' (Dik, 1997: 138), which means that the freedom of the word order in the subordinate clause has to be at least equal to that of the main clause or in majority of cases it is restricted to SVO (e.g. in English).

One of the most important devices to denote complex sentences is subordinators, which 'forming the core of the class consist of a single word' (Quirk, 1985: 998). Along with simple subordinators there are also complex subordinators and correlative subordinators (a combination of two markers of subordination). All of them are attributed to a particular clause type (see *Table 9*).

Table 9 **An overview of subordinators in English**

Simple subordinators	*after, although, as, because, before, if, once, since, that, though, till, unless, until, when(ever), where(ever), whereas, whereupon, while, whilst*		
Complex subordinators	ending with *that*		*but that, in that, in order that, insofar that, the event that, save that, such that*
	ending with *optional that*	participle form	*assuming, considering, excepting, given, granted, granting, provided, providing, seeing, supposing + (that)*
		others	*except, for all, now, so + (that)*
	ending with *as*		*according as, as far as, as long as, as soon as, forasmuch as, inasmuch as, insofar as, insomuch as*
	others		*as if, as though, in case*
Correlative subordinators	*as… so*		
	as…as, so… as, such…as, so…(that), such…(that), less…than, more (/er)…than, no sooner…than, barely…when/than, hardly…when/than, scarcely….when/than		
	the…the		
	whether…or, if…or		
	subordinator plus optional conjunct		*although, even if, (even) though, while…yet, nevertheless, etc.; if, once, since, unless…the, in that case; unless, because, seeing (that)…therefore*

In French finite subordinate clauses are immediately identifiable since they are typically marked by the conjunction *que*. When the subordinate clause expresses cause, condition, purpose, time, concession, it consists of a subordinating conjunction with final *que*, e.g. *Je dors parce que je suis fatigué*. Thus, *que* can be generalised as 'an all-purpose finite complementiser' (Rowlett, 2007: 105).

Table 10 **An overview of subordinators in French**

Conjonctions simples de subordination	que, quand, comme, si	
Conjonctions composées de subordination	*cause*	*parce que, puisque, du moment que, attendu que, vu que, sous prétexte que*
	concession	*bien que, quoique, quelque...que, si...que, pour...que, tout...que, encore...que, tandis...que, alors...que, quand même*
	but	*pour que, afin que, de peur que*
	consequence	*de façon que, de manière que, de sorte que, au point que, si bien que, tellement que*
	temps	*lorsque, aussitôt que, des que, avant que, pendant que, tant que, tandis que, après que, en même temps que, depuis que, jusqu'à ce que, en attendant que*
	condition	*à condition que, pourvu que, à moins que, soit que..., au cas où, en cas que*
	comparaison	*de même que, ainsi que, à mesure que, selon que, suivant que, comme si*

When using a subordinator in French, we have to consider which mood to use in the subordinate clause that follows: either *indicative* or *subjunctive*, as in French there is a significant distinction between these two moods: indicative is *assertive* whereas subjunctive mood is *non-assertive* (ibid.: 106). Moreover, the subjunctive mood is characteristic of subordinate clauses. In simple and compound clauses indicative mood is found. The mood in French is determined by the verbs of desire (e.g. Marie *veut* qu'il fasse beau demain.), verbs, adjectives and nouns which express emotion and are non-assertive:

a) Je *regrette* [qu'elle soit malade].

b) Je suis *heureuse* [qu'il parte aujourd'hui].

c) Je ne regrette pas mon *désir* [qu'il parte aujourd'hui].

Since subordinators indicate the semantic relationship between the subordinate clause and the clause it is dependent on, they are an integral part of syntactic structure analysis.

Position of subordinate clauses

Carnie has said that 'one of the most interesting parts of syntax is comparing the sentence structure of different languages' (Carnie, 2007: 174) as embedded constructions may appear in different syntactic forms depending on the language, either before or after the main predicate. According to Quirk et al. (1985), subordinate clauses may be positioned initially, medially, or at the end of the main clause (1985: 1037). For example:

INITIAL: [*When you're ready*], we'll go to my parents' place.

[*Quand tu seras prêt*], nous irons chez tes parents.

MEDIAL: We'll go, [*when you're ready*], to my parents' place.

Nous irons, [*quand tu seras prêt*], chez tes parents.

END: We'll go to my parents' place [*when you're ready*].

Nous irons chez tes parents [*quand tu seras prêt*].

The embedded constructions that appear in the initial position are possible mainly in conditional sentences:

[*If you learn the irregular verbs*], then we will go to the cinema.

[*Si tu apprends les verbes irréguliers*], nous irons au cinéma.

In a majority of cases we observe the subordinate clauses at the final position and they are the easiest to understand. According to Dik (1997: 126), the positioning of embedded constructions is also determined by the categorical complexity of embedded constructions. Thus, constituents prefer to be placed in order of increasing complexity, where complexity is defined as follows: Clitic < Pronoun < Noun Phrase < Adpositional Phrase < Subordinate Clause (ibid.: 127). This means that subordinate clauses are to be placed at the end of a complex clause, which is a dominant tendency both in English and French.

It should be noted that the positioning of the embedded constructions depends not only on the order of the constituents, but also on the semantic function. This was already postulated by Greenberg (1963: 84) when he formulated the following universals:

Universal 14. In conditional statements, the conditional clause precedes the conclusion as the normal order in all languages.

Universal 15. In expressions of volition and purpose, a subordinate verbal form always follows the main verb as the normal order except those languages in which the nominal object always precedes the verb.

Thus, most English and French subordinate clauses are positioned at the end of the main clause and are initially marked by subordinators, while conditional clauses appear more in the initial position of the main clause:

We should learn languages, *because* it will be very important for our future.

On doit étudier des langues, *parce que* cela sera très important pour notre avenir.

Complex sentences that contain two or more subordinate clauses (or multi sub-clauses) can have two basic arrangements: *parallel* and *consecutive* (Blokh, 2003: 357). The subordinate clauses that refer to one and the same main clause are subordinated in parallel, e.g. I'll meet you [*when you have finished your work*] and [*when you have done shopping*].

62
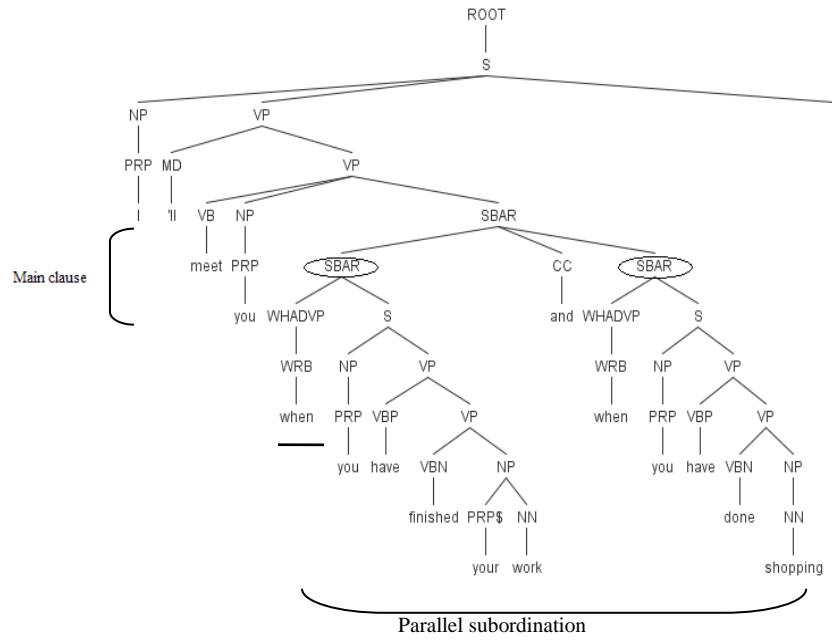
Figure 16 **Homogeneous parallel subordination**

Moreover, parallel subordination may be both 'homogeneous', i.e. they depend on the main clause as a whole, are connected by a coordinating conjunction and have the same function (see the sentence above) and 'heterogeneous', i.e. parallel subordinate clauses refer to different elements in the main clause (ibid.: 358), e.g. The producer [*who introduced us with the script*] was very excited about the film [*that he was going to shoot*].
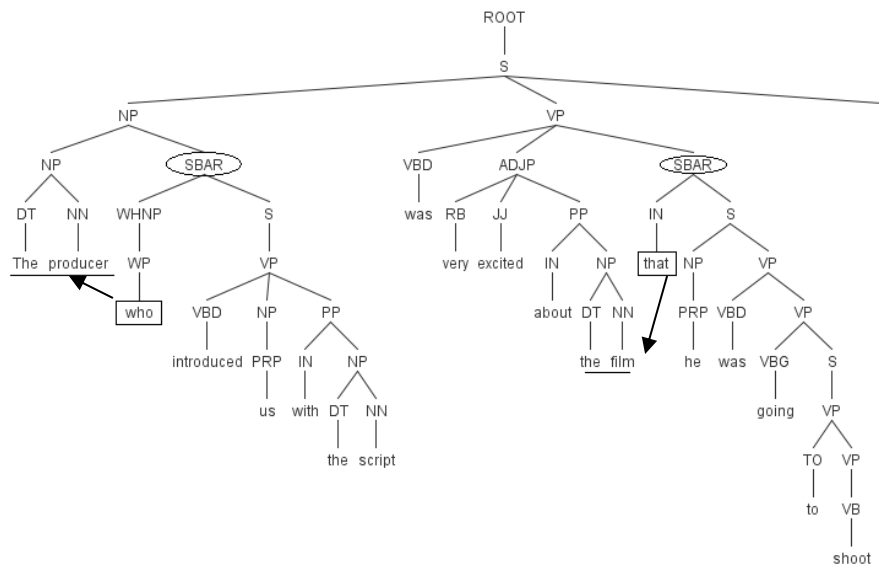
Figure 17 **Heterogeneous parallel subordination**

In consecutive subordination one subordinate clause is subordinated to another making 'direct' consecutive subordination (ibid.), e.g. She had no idea [*why he had not arrived at the time*] [*(that) they had agreed upon*].
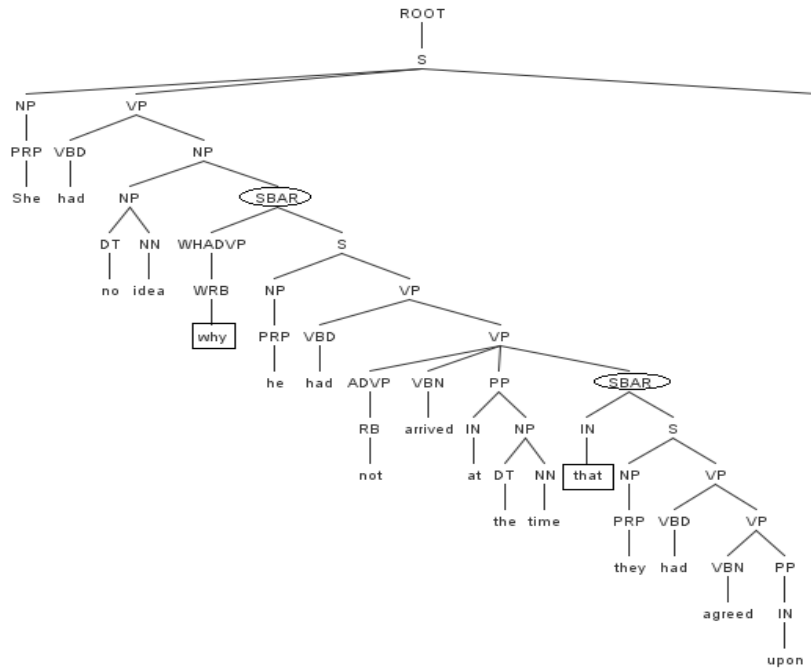
Figure 18 **Consecutive subordination**

Blokh distinguishes one more type of consecutive subordination, i.e. 'oblique' consecutive subordination (ibid.: 359). This is the case when the main clause comprises a subject or predicative clause, e.g. [*What he heard*] made him turn red [*as if he had not expected it*].
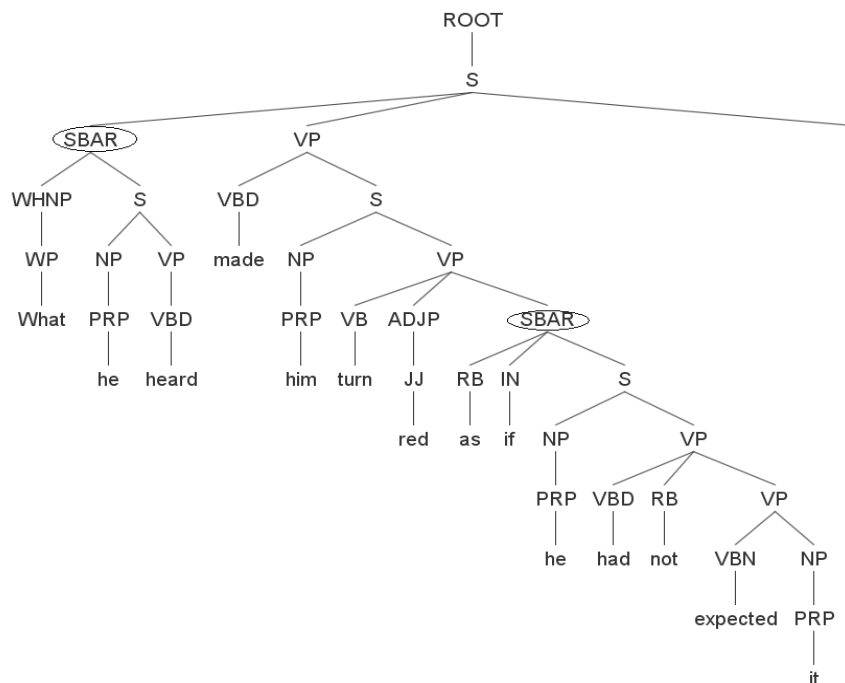
Figure 19 **Oblique consecutive subordination**

Graphically the structure of the sentence is consecutive, though not directly as [What he heard] is the subject of the main clause.

Classification of finite subordinate clauses

Subordinate clauses may be classified according to two principles: the *categorical* and the *functional* principle (Blokh, 2003). According to the categorical classification, 'the nominative properties of notional words' (Blokh, 2003: 334) are to be relied upon, which are reflected by the part-of-speech classification. Thus, we can distinguish among *noun, adjectival* and *adverbial* clauses (see also, e.g. Huddleston, 1984; Biskri et Desclés, 2005). This principle was mainly applied by traditional grammars when classifying complex sentences. The functional principle bases the classification upon 'the analogy of the positional parts of the simple sentence, since it is the structure of the simple sentence that underlies the essential structure of the complex sentence' (Blokh, 2003: 333). For example, Chevalier et Blanche-Benveniste, 1964; Grevisse, 1969; Wilmet, 1997 distinguish: subject, attribute, object, circumstantial adverb, noun complement, etc.; Blokh himself (2003) proposes a more complicated sub-division of subordinate clauses: 1) *clauses of primary nominal positions*, which function as subject, predicate and object of the main clause; 2) *clauses of secondary nominal position*, which include attributive clauses; 3) *adverbial clauses*; 4) *parenthetical or insertive constructions*, e.g. Jack has called here twice this morning, *if I am not mistaken*. Lyon (1968) subdivides the subordinate clauses by function into *nominal, adjectival* and *adverbial*; Quirk et al. (1985) also on the basis of their function distinguish *nominal, adverbial, relative*, and *comparative* clauses while Fischer et al. (2000) enumerate *adjective (= relative)*, *complement* and *adverbial* clauses. Cristofaro (2003), O'Grady and Archibald (2004) divide the subordinate clauses into *complement, relative* and *adverbial* clauses; Huddlestone (1984) classifies them into *relative, comparative, content* clauses and *clauses introduced by because, before, unless, etc.*; Le Goffic (1993) differentiates among four clause types: 1) complement clause (e.g. Je crois [que c'est inutile]); 2) relative clause with antecedent (e.g. Le livre [qui est sur la table]); 3) integrative pronominal clause containing: relative clause without antecedent (e.g. [Qui dort] dîne) and integrative adverbial clause (circumstantial with *qu-* or *si*), e.g. [Quand tu veux] tu peux; 4) percontative (indirect interrogative/exclamative clause), e.g. Jean ne sait pas [comment il pourrait réussir]). Dik (1997) takes into consideration the semantic, perspectival and pragmatic functions of embedded constructions; Carnie (2007) groups the embedded clauses into *specifier* (ones that serve as the subject of a sentence), *complement* (they are directly linked to the verb), and *adjunct* (e.g. relative clauses) clauses. For example:

1) Everybody knows [that Tina loves chocolate]. (complement clause)

Chacun sait que [Tina aime le chocolat].

2) The lady [I saw talking with the old man] is my sister. (adjunct clause)

La dame [que j'ai vue parler] avec un home âgé est ma soeur.

3) [[Lady selling cabbages] caused the accident in the street]. (specifier clause)

[[La dame qui vend du chou] a provoqué un accident de route].

Relative clauses do not function by themselves as subjects or objects as in example 2, they rather perform a function of a modifier of a noun phrase. Thus, *I saw talking with the old man* is a relative clause that modifies *the lady*. In the majority of aforementioned classifications the linguists have subdivided the subordinate clauses into nominal, adverbial and adjectival (or relative); however, in some cases the concept 'complement' is used instead of 'nominal'. This could be explained by the fact that nominal clauses are 'normally selected or **controlled** by a preceding verb, adjective, noun or preposition' (Biber et al., 1999: 194). Hence, when discussing their complement role, this clause type is frequently called complement clause.

Having examined the main principles of subordinate clause classification proposed by contemporary linguists of English and French it was decided to apply the traditional or categorical division of complex sentences into noun, adverbial and adjectival clauses for the present research as it could be used cross-linguistically.

2.3.2.1 Noun Clause

Noun clause, known also as *nominal* clause is 'a type of subordinate clause that functions in sentence structure where noun phrases usually occur' (Jackson, 2007:44) or, according to Matthews (1997) definition, it is 'a clause whose syntactic role is seen like that of a noun or noun phrase: thus, in particular, 'a subject clause' or 'object clause' (Matthews, 1997: 269). Within a sentence noun clauses may perform the function of a subject, object and complement. Moreover, they may function as adjective complementation, which is not typical for noun phrases. Semantically noun clauses refer to abstract notions - facts, dates, events. The only exception is the *nominal relative clause*, 'which may refer to objects (including persons) and which has some of the properties of a noun phrase consisting of head and postmodifying relative clause, the head and relative pronoun coalescing to form a single *wh*-element" (Quirk et al., 1985: 1047). Consequently, finite noun clauses can be subdivided into:

1) subordinate declarative clauses;

2) subordinate interrogative clauses;

3) subordinate exclamative clauses;

4) nominal relative clauses.

Subordinate declarative clause

Finite subordinate declarative clauses in English and French are instantly identifiable as they are marked by a subordinating conjunction *that* (in English) and *que* (in French), which is a typical declarative complementiser. *That* clauses may have different functions in a sentence, i.e. they may function as subject, direct object, subject complement, adjectival complementation and adverb complementation (only in French):

1) subject:

(a) [*That the invading troops have been withdrawn*] has not affected our government's trade sanctions. (Quirk et al.,1985: 1049)

(b) [*Qu'il vienne*] m'étonnerait beaucoup. (Riegel et al.,1994: 827)

When having the function of a subject, the *that*-clause appears at the beginning of a noun clause (sentence (a)). In French, when a *that*-clause functions as a subject at the beginning of the sentence, the introductory pronoun *que* requires the verb in subjunctive mood (e.g. *vienne* in sentence (b)). There are cases when the *that*-clause is extraposed, i.e. the pronoun *it (il* or *ce* in French*)* occupies the place of *that (que).* In such a case the pronoun can be called 'an anticipatory *it* (or preparatory *it* because it comes before the clause to which it refers' (Online 1). For example:

[*It is true*] that you received a top mark.

[*C'est vrai*] que tu as reçu la meilleure note. / [*Il est probable*] que tu as reçu la meilleure note.

Semantically the extraposed clause does not add any new information to the original sentence, it only facilitates the understanding of the idea. Structurally it has its own subject – *it (ce)* and predicate – *is (est).*

2) direct object:

I noticed [*that you were not able to do the task*].

J'ai remarqué [*que vous n'étiez pas capables de réaliser la tâche*].

When the subordinate declarative clause functions as direct object, it immediately follows the predicate of the main clause, and the structure of the subordinate clause is *subject + verb + complement,* as in the above examples. However, it is possible to have the inversion *verb + subject* in French if the verb is intransitive and the subject is a NP that does not

contain a personal pronoun. For example, in the sentence *J'aime [que surviennent de nombreux rebondissements]* (Riegel et al., 1994: 825) the verb *surviennent* in the subordinate clause is placed before the NP *de nombreux rebondissements*, thus having the structural pattern [verb + subject].

Sometimes the subordination is formally expressed without any subordinator, which may be the case in English. The conjunction *that* might be omitted when: (i) the *that-clause* functions as direct object or complement, e.g. I know *you will be late*; (ii) a subject *that*-clause is extraposed, e.g. *It's* obvious *you don't know Chinese.* Sometimes there might be confusion in deciphering whether *that* is an adverbial and belongs to the main clause, e.g. They told us once again (*that) the conflict was worsening* or it introduces a noun clause, e.g. They told us *(that) once again the conflict was worsening.* However, this phenomenon is rather rare across languages (e.g. in French the subordinate clause is always introduced by a subordinator).

In French one has to be aware of the fact that there are verbs (*accepter, aimer, conseiller, craindre, defender, demander, désirer, détester, douter, ordonner, permettre préférer, regretter*, etc.) and expressions *(il/c'est étrange, important, normal, ridicule, etc.; il faut, il semble, il est dommage, etc.; cela/ca m'agace, m'étonne, m'inquiète,* etc.) that require the subjunctive mood in the subordinate clause. Moreover, some verbs that normally require the indicative mood, when used in the negative demand the subjunctive mood:

Les sauveteurs *sont certains* qu'il y *a* encore des survivants.

Les sauveteurs *ne sont pas certains* qu'il y *ait* encore des survivants. (Delatour et al., 2004: 216).

3) subject complement

When the noun clause has the function of subject complement, it appears directly after the linking word *is (est)* both in English and French:

The problem is [*that he is not willing to work*].

Le problème est [*qu'il n'est pas prêt à travailler*].

4) adjectival complementation:

They are happy [*that you will be able to come to their birthday party*].

Ils sont heureux [*que vous puissiez venir à leur anniversaire*].

In the above examples the subordinate clause functions as complement of the adjective *happy (heureux)*. One should be aware of the fact that in French after the adjectives that

express feelings, e.g. *être heureux, malheureux, triste, surprise, content*, the verb that follows the subject of the subordinate clause has to be used in the subjunctive mood.

5) adverb complement (in French):

Il fait de plus en plus chaud. Peut-être [*qu'on pourra aller à la mer*].

When having the function of adverb complement, the subordinate clause follows directly the adverb, which is the only element in the main clause.

Subordinate interrogative clauses

In English the subordinate interrogative clauses are referred to as *wh*-interrogative clauses (as they are introduced by *who, what, when, where, why, which, how*) while in French they are simply called interrogative clauses. However, their function in a complex sentence is similar. The interrogative clauses may function as:

1) subject:

[*How you will play at the concert*] depends on yourself.

[*Comment tu joueras au concert*] dépend de toi-même.

The subordinate interrogative clause functioning as subject appears at the initial position of the sentence and contains the interrogative word *how (comment)* followed by the *subject + verb*.

2) direct object:

I don't know [*what they are looking for*].

Je ne sais pas [*ce qu'ils recherchent*].

In French the interrogative form *est-ce que* is suppressed, and the subordinate clause is introduced by *ce que*. When the interrogative subordinate clause functions as direct object it appears at the final position of the sentence directly after the predicate of the main clause.

3) subject complement:

The question is [*who will give water to the elephants*].

La question est [*qui donnera de l'eau aux éléphants*].

In the above sentences the subordinate clause follows directly the linking word *is (est)*.

4) adjectival complementation:

He is not sure [*which one she will choose*].

Il n'est pas sûr [*lequel elle choisira*].

In both languages the subordinate interrogative clause functioning as adjectival complementation is preceded by the adjective *sure (sûr).*

5) prepositional complement:

They did not consult us on [*what we had to do*].

Ils ne nous ont pas renseignés sur [*ce qu'il fallait faire*].

Having examined various functions of subordinate interrogative clauses it has to be concluded that: 1) indirect questions resemble the *wh*-interrogative clauses both in English and French; 2) the word order in the subordinate clause is not the same as in interrogative sentences, though introduced by the *wh*-element which demands the inverted word order; 3) the subordinate interrogative clause has the same functions in relation to the main clause as the declarative subordinate clause (subject, direct object, subject complement and adjectival complementation) with the exception of the function as prepositional complement; 4) when having the function of subject the noun clause is placed at the beginning of the sentence.

Yes-no and alternative interrogative clauses

Along with the *wh*-interrogative clauses we distinguish also the subordinate *yes-no* and alternative interrogative clauses introduced by the subordinators *whether* and *if* in English and *si* in French. These clauses may perform the same functions in a sentence as *wh*-interrogative clauses, e.g. I wonder *if you can help me*. Je me demande *si vous pouvez m'aider*. Here the interrogative clause functions as direct object.

The correlatives *whether ... or* or *if ... or* in English and *si ....ou* in French form the alternative clauses:

(a) I don't know [*whether/if* he has arrived or *whether/if* he is still at work].

Je ne sais pas [*s'*il est arrivé ou *s'*il est toujours au travail].

If the second part of the subordinate clause is a full sentence as in (a), then it has to be repeated, which is not the case in (b – d) as it is the abbreviated form of the clause:

(b) I couldn't decide [*whether it was cheap or expensive*].

Je ne pouvais pas décider [*si c'était cher ou pas cher*].

(c) The waiter asked them [*if they wanted potatoes or rice*].

Le serveur leur a demandé [*s'ils voulaient des pommes de terre ou du riz*].

(d) It doesn't matter [*if they come or not*].

Peu importe [*s'ils viennent ou non*].

As for the sentence structure, the *yes-no* and alternative interrogative clauses occupy the same position (initial or final) within a sentence as the interrogative subordinate clause because they perform the same functions.

Exclamative clauses

In both languages the subordinate exclamative clauses have the same introductory words as in independent exclamative clauses: *what* predetermines a NP (e.g. What a nice dress she has (It's unbelievable *what a nice dress she has*) / Quelle jolie robe elle a (Il est incroyable *quelle jolie robe elle a*)); *how* intensifies an adjective, adverb or the whole clause (e.g. How beautiful the dog is (It's surprising *how beautiful the dog is*) / *Comme* le chien est beau (Il est surprenant *comme le chien est beau*). There is an exception in French for *que,* which changes its form when introducing the subordinate exclamative clause (e.g. *Que* tu es belle/ J'adore *à quel* point tu es belle). Although there are some differences in the form of exclamative clauses in English and French, their function is similar, except for the function of prepositional complement since it is typical for English only. Thus, the exclamative clauses may function as:

1) extraposed subject:

   It's unbelievable [*how much she can eat*].

   Il est incroyable [*combien elle peut manger*].

2) direct object :

   I remember [*what good friends we were*].

   Je me souviens [*quels bons amis nous étions*].

3) prepositional complement:

   I read an account of [*what an impression you had made*]. (Quirk et al., 1985: 1053)

As for the position in the sentence, the subordinate exclamative clause is always placed after the main clause.

Nominal relative clauses

Nominal relative clauses are rather difficult to distinguish from *wh*-interrogative clauses as they are also introduced by a *wh*-element (e.g. That's *what she likes doing* / C'est *ce qu'elle aime faire*). Moreover, nominal relative clauses are 'more like noun phrases, since they can be concrete as well as abstract and can refer even to persons [...] we can paraphrase them by noun phrases containing a noun head with general reference that is modified by a relative clause' (Quirk et al., 1985: 1056):

   I took [*what they offered me*]. → I took *the things that they offered me*.

   J'ai pris [*ce qu'ils m'avaient offert*]. → J'ai pris *les choses qu'ils m'avaient offertes.*

Thus, we can say that the nominal relative clause is 'basically a noun phrase modified by an adnominal relative clause [...] and can function as an element in a superordinate clause'

(ibid.: 1056). Moreover, nominal relative clauses perform the same functions as noun phrases; in addition, they can function as (a) indirect object and (b) object complement:

(a) He lent [*whoever asked*] his new car.

Il a prêté [*quiconque a demandé*] sa nouvelle voiture.

(b) You can call me [*what(ever) you like*]. (ibid.: 1058).

Vous pouvez m'appeler [*comme vous voulez*].

In English it is sometimes quite difficult to distinguish between the nominal relative clause and the relative clause, as they seem to be rather similar. In that case we have to replace *that* by *which*:

I love the idea *that* cats can talk. ➔ I love the idea *which* cats can talk.

That is not replaceable by a *wh*-pronoun in the nominal relative clause; it could have been done in relative clauses without changing the meaning of the sentence. Therefore, it is obvious that the first clause is a complement clause rather than a relative clause and that the second clause is incorrect. What is more, the nominal relative clause can have two positions in a sentence: medial (sentence *a*) and final position (sentence *b*).

All in all, finite noun clauses examined beforehand perform the same functions both in English and French, the only exceptions being: (i) the subordinate declarative clause functioning as adverb complement in French and (ii) the subordinate exclamative clause having the function of prepositional complement in English. These two cases cannot be attributed across languages when analysing the linguistic knowledge of language learners. As for the position of noun clause in relation to the main clause it always appears after the main clause, except for the subordinate declarative and interrogative clauses functioning as subject, which are placed at the beginning of the main clause.

2.3.2.2 Adverbial Clause

The definition provided by Koptojevskaja-Tamm (1993: 23) describes an adverbial clause as 'one functioning as an adverb with respect to the main predicate.' O'Grady and Archibald (2004: 117) regard adverbial clauses as 'subordinate clauses that, like adverbs, modify a verb.' Thus, the main function of this particular clause is that of adverbial.

Adverbial clauses are sometimes divided into two types: *adjuncts* and *disjuncts* (Carter and McCarthy, 2006: 561). Semantically 'adjuncts denote circumstances of the situation in the matrix clause whereas disjuncts comment on the style or form of what is said in the matrix clause [...] or on its content [...] syntactically disjuncts are peripheral to the clause to which they are attached' (Quirk et al.,1985: 1070). In writing disjuncts are

separated by a comma from the main clause. Quirk et al. (1985) further subdivide adjuncts into *predication* (obligatory and optional) *adjuncts* and *sentence adjuncts* whereas disjuncts - into *style disjuncts* and *content disjuncts*. Being either *obligatory* or *optional* predication adjuncts are positioned after the main clause. Moreover, *obligatory adjunct* clauses:

a) are related to the verb *be* of the main clause, e.g. Your English text book <u>is</u> [*where you left it yesterday*]. Votre livre d'anglais <u>est</u> [*où vous l'avez laissé hier*];

b) may be related to other verbs of the main clause, e.g. Jane may stay [*as long as she wants*]. Jane peut rester [*aussi longtemps qu'elle le souhaite*];

c) may function as the adverbial, e.g. He found his trousers [*where he had left them*]. Il a trouvé son pantalon [*où il l'avait laissé*].

Optional adjuncts relate more to the verb than to the whole main clause, e.g. I found the wallet [*where I left it*] (The wallet was where I left it). J'ai trouvé le portefeuille [*où je l'ai laissé*] (Le portefeuille était où je l'ai laissé). As for *sentence adjunct* clauses, they may occur initially or finally since they are not directly related to the verb of the main clause, e.g. [*Whenever I meet Jane*], she looks tired. [*Chaque fois que je rencontre Jane*], elle a l'air fatigué.

Style disjuncts contain a verb and the subject *I* and they 'refer to the circumstances of the speech act' (Quirk et al., 1985: 1072), e.g. I have no money in my wallet, [*because I gave it to my son*]. Je n'ai pas d'argent dans mon portefeuille, [*parce que je l'ai donné à mon fils*], whereas *content disjuncts* refer to the main clause, e.g. [*Since you know French*], you can translate this instruction. [*Puisque vous connaissez le français*], vous pouvez traduire ce mode d'emploi.

Sometimes adjuncts and disjuncts may cause confusion, as there are cases when they have the same subordinator and they appear either before or after the main clause, e.g. in English the subordinator *since* may have the meaning of time (a) *or* reason (b):

a) He has been watching TV [*since his wife left*]. (adjunct)

Il avait regardé la télévision [*depuis que sa femme est partie*].

b) He took an umbrella, [*since it was raining*]. (disjunct)

Il a pris un parapluie, [*parce qu'il pleuvait*].

Syntactically there is no difference between adjunct and disjunct clauses as they may have the same subordinator, and they may have either initial or final position in relations to the main clause.

Semantically we might distinguish rather many types of adverbial clauses taking into consideration the role of adverbials in a sentence in general. Thus, there are adverbial clauses denoting time, purpose, place, manner, etc; Delatour et al. (2004: 13-14) distinguish the adverbial clauses of cause, consequence, purpose, time, contrast, condition and comparison; Biber et al. (1999) recognise the adverbial clauses of place, time, manner, reason, cause, purpose and condition; Blokh (2003: 348-351) classifies them into four groups: 1) clauses of time and place; 2) clauses of manner and comparison; 3) 'classical' clauses of different circumstantial semantics: *attendant event, condition, cause, reason, result (consequence), concession, purpose;* 4) parenthetical or insertive constructions. It is obvious that there is no uniform classification provided by linguists. What is more, unlike noun clauses, which either use *that (que)* or *nothing* in English to denote the subordination, adverbial clauses make use of a range of subordinators which introduce clauses with different meaning, thus, indicating the semantic relationship between the main and subordinate clause. However, as aforementioned, one and the same subordinator may be used to denote different semantic relations, e.g. *since* may introduce a temporal or causal clause in English; *tandis que* in French can be used to denote time, concession or sometimes - even contrast, which may cause confusion when denoting to which clause type the particular clause belongs to. Therefore, Riegel et al. (1994: 846) suggest that it would be better to base the classification of adverbial clauses upon formal criteria, such as mobility, the use of indicative or subjunctive mood, or the possibility of being substituted by gerunds or participle clauses. However, the author of the present research will classify the adverbial clauses according to the roles of adverbials in general into: time, place, contrast, condition, concession, reason/cause, purpose, result and comment clauses.

Clause of time

Adverbial finite clauses of time in English are introduced by the subordinators: *after, as, before, once, since, till, until, when, whenever, while, whilst, now (that), as long as, as soon as, immediately, directly:*

[*Once you finish reading the book*], you may go out.

In French the following subordinators are used: *au moment où, en même temps que, pendant que, tandis que, alors que, lorsque, aussi longtemps que, tant que, quand, comme, avant que, jusqu'à ce que, en attendant que, après que, dès que, aussitôt que, une fois que, chaque fois que, toutes les fois que, à mesure que, depuis que, maintenant que, à présent que, à peine...que:*

[*Dès que vous finissez de lire le livre*], vous pouvez sortir.

It should be noted that many conjunctions in French, such as *avant que,* consist of two elements: preposition and the subordinator *que.*

As for the position in a sentence, adverbial clauses of time usually appear either before or after the main clause; though, there are cases when they are placed in the mid position, e.g. The lady, [*after she had finished her coffee*], left the cafeteria. La dame, [*après qu'elle avait fini son café*]*,* est sortie de la cafétéria.

According to Quirk et al. (1985: 1080), 'an adverbial clause of time relates the time of the situation denoted in its clause to the time of the situation denoted in the matrix clause.' Hence, there may be different time relations between the main and the subordinate clause, i.e. the situation in the main clause may happen previously, simultaneously or subsequently in respect to the subordinate clause. Givón (1990: 827-37) has used the concepts: *temporal anteriority, temporal overlap* and *temporal posteriority* when subdividing the adverbial clauses of time into three groups, which will be utilised also within the framework of present research. What is more, these time relations are not expressed only by certain subordinators, but also by tense, aspect, adverbs, adjectives, nouns expressing time, etc.

Clause of temporal anteriority

In the relation of temporal anteriority the situation in the main clause occurs before the situation in the subordinate clause or is leading up to the situation in the subordinate clause. This is indicated by the subordinators *until, till* and *before* in English. It should be noted that *till* is less frequent than *until* although they both are used in the same way. In French the notion of 'anteriority' is marked by the subordinators *avant que, jusqu'à ce que, en attendant que*, which all demand the use of the verb in subjunctive mood. Another aspect that should not be neglected is that the main clause (a) followed by *until*-clause or *jusqu'à ce que*-clause must be continuous, which is not the case with *before*-clause or *avant que*-clause (b):

a) The children refused to go to bed [*until their mum returned from the concert*].

Les enfants refusaient d'aller au lit [*jusqu'à ce que leur maman soit revenue du concert*].

b) They went to bed [*before their parents arrived*].

Ils sont allés se coucher [*avant que leurs parents soient arrivés*].

Moreover, example (a) expresses the result of the activity in the main clause as well as time whereas example (b) only has a time relation between the main and subordinate

clause. Thus, the adverbial clause of temporal anteriority may perform some more functions along with the main one, i.e. expressing time. It may:

i) denote the exact time from which the situation in the main clause starts:

He started watching TV [*before his wife entered the living-room*].

Il a commencé à regarder la télévision [*avant que sa femme entre dans la salle de séjour*].

ii) express preference:

He'll suffer [*before he'll forgive her*]. (He would rather suffer than forgive her.)

Il va souffrir [*avant qu'il lui pardonne*].

iii) express implausibility:

The tail of owl will blossom [*before he'll become a doctor*]. (He'll never become a doctor.)

La queue de la chouette s'épanouira [*avant qu'il devienne docteur*].

iv) imply purpose and result as well as time:

I had to put my complaint in writing [*before they would take any action*]. (...so that they would...; They wouldn't take any action until I put my complaint in writing.) (Quirk et al., 1985: 1081)

J'ai dû déposer ma plainte par écrit [*avant qu'ils ne prennent aucune   mesure*].

Clause of temporal overlap

In English the temporal overlap or simultaneity of the situation in the main clause and the subordinate clause is indicated by subordinators - *as, as long as, so long as, while, whilst, when, whenever, now (that)* when used as time conjunctions. In French the notion of simultaneity is expressed by subordinators - *au moment où, en même temps que, pendant que, tandis que, alors que, lorsque, aussi longtemps que, tant que, quand, une fois que, comme.* The subordinators - *as long as, so long as, while, whilst* (in English) and *pendant que* (in French) function as duration adverbials:

My dad arrived [*while I was sunbathing*].

Mon père est arrivé [*pendant que je me bronzais*].

In French the conjunctions *tandis que* and *alors que* have the same meaning as *pendant que*, but they imply also the nuance of contrast :

[*Tandis que le ministre parlait aux députés*], on entendit des protestations sur les bancs de l'opposition (Delatour et al., 2004:257).

The subordinators *as* in English and *au moment où, en même temps que, comme* in French characterise the two situations happening at the same time:

[*As the sun set*], it became colder.

[*Comme le soleil se couchait*], il a fait plus froid.

The subordinators *as long as* and *so long as* in English may denote condition as well as time:

The government is happy [*as long as its people are happy*].

The subordinators *when, whenever*, *now* in English and *quand, à chaque fois, maintenant* in French may also denote simultaneity:

[*When I have a headache*]*,* I see demons attacking me.

[*Quand j'ai mal à la tête*], je vois les démons m'attaquer.

Now / maintenant may occur in the combination with *that / que* expressing reason and simultaneity:

My mother is happy [*now that my father is back from Ireland*].

Ma mère est heureuse [*maintenant que mon père est de retour d'Irlande*].

All the above examples demonstrate the overlap of subordinate clause functions in relations to the main clause, though the main function is to reveal the simultaneity with the main clause.

Clause of temporal posteriority

There are a number of subordinators which denote the sequence of events in the main clause after the events in the subordinate clause: *after, as soon as, directly, immediately, once, since, when, whenever, now (that)* in English and *après que, des que, aussitôt que, une fois que, depuis que* in French. Four of the subordinators in English (*as soon as, directly, immediately, once*) and two subordinators in French (*dès que, aussitôt que)* emphasize the proximity of the events in both main and subordinate clauses. The proximity effect could be attained by adding such modifiers as *just/juste, right/tout de suite, immediately/immédiatement*, etc. to the subordinators *after/après* and *as soon as/dès que*:

I went to sleep [*soon after my mum returned from the theatre*].

Je suis allé dormir [*peu après que ma mère est rentrée du théâtre*].

The subordinator *since / depuis que* indicates the beginning of the period of time when the situation in the main clause happens:

She feels better [*since she left the hospital*].

Elle se sent mieux [*depuis qu'elle a quitté l'hôpital*].

When, *whenever* and *once; quand, une fois que* and *chaque fois que* may express various notions in a sentence. They may:

a) indicate a sequence of events (She was surprised [*when he entered the room*]. Elle était surprise [*quand il est entré dans la pièce*].)

b) imply a cause (The girls started crying [*when she heard the news*]. Les filles ont commencé à pleurer [*quand elles ont entendu les nouvelles*].);

c) merge time and condition ([*Once I set up my goals*], I try to achieve them. [*Une fois que je me pose les objectifs*], je cherche à les atteindre.)

The notion of time and concession can also be expressed by the subordinator *when / quand*:

He was enjoying himself [*when he should have been sleeping*]. (..., [whereas he should have been sleeping].)

Il s'amusait [*quand il aurait fallu dormir*].

The analysis of different semantic roles performed by the subordinate clause of time in relation to the main clause demonstrates the range of functions the subordinators may have.

Clause of place

This clause type is mainly introduced by subordinators *where/où* (specific) and *wherever/partout où* (nonspecific*)* denoting position and direction:

a) I always want to return to the town [*where I grew up*].

J'ai toujours envie de retourner à la ville [*où j'ai grandi*].

However, there are cases when the subordinator *where/ où* expresses the notion of place and contrast at the same time:

b) [*Where I saw success*], she saw failure.

[*Où je voyais réussite*], elle voyait l'échec.

Usually the subordinate clause of place occupies the more neutral position within a sentence, i.e. it is positioned after the main clause (example *a*).

Clauses of condition, concession and contrast

According to Quirk et al. (1985: 1088) *conditional clauses* convey that the situation in the matrix clause is contingent on that in the subordinate clause; *clauses of contrast* merely convey a contrast between two situations; *concessive clauses* may also imply a contrast, but their main role is to reveal that the situation in the matrix clause is unexpected in the light of that in the concessive clause. In the adverbial clauses of condition, concession, and contrast the semantic roles of subordinators may overlap, e.g. *while* in English may convey

the meaning of concession as well as of contrast. The overlap is also indicated by the subordinators *if/si, whereas/tandis que, even if/même si*, as they may perform more than one role introducing the adverbial clause. Finally, in majority of cases conditional, concessive and contrast clauses occupy the initial position in a sentence.

Clause of condition

This clause type may be subdivided into *direct* and *indirect* conditional clauses. Direct conditional clauses 'convey that the situation in the matrix clause is directly contingent on that of the conditional clause' (ibid.: 1088) while indirect conditional clauses 'are open conditions that are dependent on an implicit speech act of utterance' (ibid.: 1095).

i) [*If you don't feed the dog*], he'll bark. (direct)

[*Si vous ne nourrissez pas le chien*], il va aboyer.

ii) We just want to be honest, [*if I may say so*]. (indirect)

Nous voulons juste être honnêtes, [*si je peux dire ainsi*].

In sentence (i) the *if*-clause is an adjunct, while in sentence (ii) it is a style disjunct. Thus, adverbial clauses of condition may be placed before or after the main clause.

There are quite a number of subordinators for conditional clauses (*as long as, assuming, given (that), in case, in the event that, on condition that, provided/providing (that), supposing (that) /dans la mesure où, à condition que, pourvu que, à moins que, pour peu que, en admettant que, en supposant que, au cas où, etc.*), but the most common are *if /si* and *unless/à moins que.* Several of them – *before, as long as, so long as, when, whenever, once / avant, aussi longtemps que, lorsque, à chaque fois que, une fois* – combine time with condition.

As for the indirect conditional clauses, there are certain conventional expressions to be used, such as *if I may say so, if I may be personal, if I may interrupt / si je peux dire, si je peux être personnelle, si je peux interrompre,* etc.

Adverbial clauses of *direct* condition may be subdivided into an *open condition* and a *hypothetical condition* (ibid.: 1091). 'Open conditions are neutral: they leave unresolved the question of the fulfilment or nonfulfilment of the condition, and hence also the truth of the proposition expressed by the matrix clause' (ibid.: 1091). For example: [*If I do not drive carefully*], I will have a car accident. [*Si je ne conduis pas attentivement*]*,* je vais avoir un accident de voiture.

A hypothetical condition 'conveys the speaker's belief that the condition will not be fulfilled [...], is not fulfilled [...], or was not fulfilled [...], and hence the probable or certain falsity of the proposition expressed by the matrix clause' (ibid.: 1091):

a) [*If he studied more*], he'd get better marks.

   [*S'il étudiait plus*], il aurait des meilleures notes.

   (It is very probable that he will not study more.)

b) They would be here by now [*if they had a city map*].

   Ils seraient ici maintenant [s'ils avaient un plan de la ville].

   (They presumably do not have this map.)

c) [*If she had left earlier*], she wouldn't have missed the train.

   [*Si elle était partie plus tôt*], elle n'aurait pas manqué le train.

   (She certainly did not leave earlier.)

Carter and McCarthy (2006) deciphers between 'real' conditional clauses (depending on whether the situation is true, has happened, generally happens or is likely to happen) and 'unreal' conditional clauses, in which the situation is untrue or imagined, has not happened and is only remotely to happen.

The distinction of conditional clauses in different types is of grammatical importance both in English and French because the verbs in the hypothetical (unreal) condition are moved backwards as in indirect speech. Moreover, there might be different time relations in the adverbial clauses of direct condition expressed by tense choice (see *Table 11*), which in many grammars, e.g. Cambridge Grammar of English (2006), are described as the first, second and third conditionals.

Table 11 **Verb forms with open and hypothetical conditions**

	Conditional clause	*Main clause*
Present and future reference for open condition (1st conditional)	**PRESENT / PRÉSENT** [*If you don't feed the dog*], [*Si vous ne nourrissez pas le chien*],	**FUTURE / FUTUR** he'll bark. Il va aboyer.
Present and future reference for hypothetical condition (2nd conditional)	**PAST / IMPARFAIT** [*If he studied more*], [*S'il étudiait plus*], [*If they had a city map*], [*S'ils avaient un plan de la ville*],	**PAST MODAL / CONDITIONNEL PRÉSENT** he'd get better marks. il obtiendrait des meilleures notes. they would be here by now. ils seraient ici maintenant.
Past reference for hypothetical condition (3rd conditional)	**PAST PERFECT / PLUS-QUE-PARFAIT** [*If she had left earlier*], [*Si elle avait quitté plus tôt*],	**PAST PERFECT MODAL / CONDITIONNEL PASSÉ** she wouldn't have missed the train. elle n'aurait pas manqué le train.

In English and French conditional clauses may contain the inversion with *had, should* and *were*. They may appear at the beginning or at the end of the main clause:

[*Should you lose one of your credit cards*], call our emergency helpline. (Carter and McCarthy, 2006: 562)

[*Serait-il arrivé à l'heure*], nous l'aurions vu. (Jones, 1996: 192)

However, this form is only present in formal registers; hence, the language learners graduating secondary school probably would not be familiar with this structure or would not be at ease using it in their written performance.

We can also decipher two types of open *rhetorical conditional* clauses (Quirk et al., 1985: 1094) - in the first one the assertion is derived from the conditional clause while in the second one it is derived from the main clause:

1) [*If they're Chinese*], I'm a genius.

[*S'ils sont chinois*], je suis un génie.

(Since I'm not a genius, they're certainly not Chinese.)

2) Your car must be worth a million, [*if it is worth a dollar*].

Votre voiture doit valoir un million, [*si elle vaut un dollar*].

(The car must certainly be worth a million.)

Clause of concession

'Concessive clauses indicate that the situation in the matrix clause is contrary to expectation in the light of what is said in the concessive clause' (Quirk et al., 1985: 1098). In other words, concessive clause contrasts the situation in the main clause to the one expressed in the subordinate clause. This clause type is mainly introduced by the subordinator *although* (or more informal *though*) in English and *bien que, quoique, encore que, malgré que* in French. It should be added that *although/bien que* and *though/quoique* may also relate the main and the subordinate clause with similar situations, e.g. [*Although my mum had bought me a Christmas present*], my dad bought one too / [*Bien que ma mère m'avait acheté un cadeau de Noël*], mon père a acheté un aussi. Certainly, there are some more subordinators to be used with concessive clauses: *if, even if, even though, whatever, when, whereas, wherever, while, whilst, whoever/ si, même si, que ce soit, quand, alors que, dans la mesure, tandis que, quiconque*:

She used my hair dryer [*even though I told her not to use*].

Elle a utilisé mon sèche-cheveux [*même si je lui ai dit de ne pas l'utiliser*].

Clause of contrast

Adverbial clauses of contrast are usually introduced by the subordinators *whereas* or *while/whilst* in English and *alors que, tandis que* in French, e.g. He likes chocolate, [*while she prefers biscuits*] / Il aime le chocolat, [*alors qu'elle préfère les biscuits*]. The same subordinators are used in the adverbial clause of concession, thus, they might be mixed up. Moreover, the adverbial clause of contrast resembles the clause with the coordinating conjunction *but/mais*. However, the contrast may be emphasized by such conjuncts as *in contrast/au contraire* and *by contrast/par contre* when the adverbial clause is at the beginning of the sentence.

Clause of reason

Quirk et al. (1985: 1103-1104) distinguish several types of reason clauses that are similar with respect to the main clause, and the situation in the subordinate clause precedes in time the situation in the main clause. Thus, reason clauses are subdivided into:

1) cause and effect:

He's happy [*because he passed well the entrance exams*].

Il est heureux [*parce qu'il a réussi aux examens d'entrée*].

(The cause for him being happy is that he passed well the entrance exams. The reason that he is happy is that he passed well the entrance exams.)

2) reason and consequence:

They washed the windows [*because they were dirty*].

Ils ont lavé les fenêtres [*parce qu'elles n'étaient pas propres*].

(The reason that they washed the windows was that they were dirty.)

3) motivation and result:

I washed the windows [*because my mother told me to do so*].

J'ai lavé les fenêtres [*parce que ma mère m'avait dit de le faire*].

(My motivation for washing the windows was that my mother told me to do so. The reason that I washed the windows was that my mother told me to do so.)

4) circumstances and consequence:

[*Since the rain has ceased*], the football game will be continued.

[*Puisqu'il a cessé de pleuvoir*], le match de football sera poursuivi.

(Considering the fact that the rain has ceased, the football game will be continued. The reason that the football game will be continued is that the rain has ceased.)

Reason clauses are typically introduced by the subordinators *because/parce que* and *since/puisque*. Of course, there are some other subordinators such as: *for* (formal style), *as, in case, in that, seeing (that), with* (informal style) in English and *vu que, attendu que, étant donné que* (formal style), *dès lors que, comme, du moment que* in French. However, their frequency of use is not that high. As for the position of reason clauses within a sentence, they may appear before (sentence *a*) or after the main clause (sentence *b*):

a) [*As I was very tired*], I went to bed early.

[*Comme j'étais très fatigué*], je suis allé me coucher tôt.

b) I took the map [*since I was afraid to get lost*].

J'ai pris la carte [*comme j'avais peur de me perdre*].

Clause of purpose and result

Clauses of purpose are more often infinitival than finite. Finite clauses are mainly introduced by the subordinator *so that*, *so, so as to* or *in order to/that* in English and *pour que, afin que, de peur que (ne), de crainte que (ne)* in French, which require the use of subjunctive. The finite purpose clause may occupy the position after the main clause:

He has to study hard [*so that he could pass the exams*].

Il doit étudier sérieusement [*afin qu'il puisse passer les examens*].

One should be very careful in deciphering between the clause of result and the clause of purpose as the expressed meaning is similar. The only difference is that in the result clause the result is already achieved, whereas in the adverbial clause of purpose the result is yet to be achieved. Moreover, in English they both have the same subordinators *so that* and *so* (*so that* is more often used for purpose and *so* to express result). In cases when *that* is omitted in English, the result clause might be regarded as the coordination. The conjunction *and* is to be inserted to avoid miscomprehension. Accordingly, in French the conjunctions *donc*, *alors* and *ainsi* differentiate between coordination and subordination.

She received the first salary, *and so she could buy a new dress.*

Elle a reçu le premier salaire, *donc elle a pu acheter une nouvelle robe.*

What is more, the result clause may only appear after the main clause and is separated by a comma.

Clauses of similarity and comparison

The clauses of *similarity* are introduced by *as* and *like* in English, and by *comme* in French. When used in a sentence they are transformed by *just/juste* and *exactly/exactement*, which

appear before the subordinators, e.g. She sings [*just as her brother did*] / Elle chante [*juste comme son frère faisait*].

The clauses of *comparison* are introduced by *as if*, *as though* and *like* in English, and by *comme si* in French, e.g. She looks [*as if she is going to fall*] / Elle a l'air [*comme si elle va tomber*].

If the verb in the clause is dynamic, the clause conveys the meaning of a manner. The clause may also express hypothesis, e.g. She behaves *as if she were a kid* / Elle se comporte *comme si elle était un enfant*.

One has to be careful not to confuse adverbial clauses of comparison with the comparative clauses, which express the second part of a comparison, e.g. The house was not *as big* [as I thought] / La maison n'était pas *aussi grande* [que je pensais].

Clause of comment

Comment clauses express: 1) the speaker's comment on what is given in the main clause, thus, performing the function of a content disjunct; 2) the speaker's view on the way he/she is speaking, thus, being a style disjunct. Quirk et al. (1985: 1112) distinguish different types of comment clauses:

    (1) like the matrix clause of a main clause:

        There were no tickets, *I believe*, for that concert.

        Il n'y avait pas de billets, *je crois*, à ce concert.

    (2) like an adverbial finite clause (introduced by *as/comme*):

        I'm going to France next week, *as you know*.

        Je vais en France la semaine prochaine, *comme vous le savez.*

    (3) like a nominal relative clause, which must be initial:

        What was the most surprising, he won the gold medal.

        Ce qui était le plus surprenant, il a remporté la médaille d'or.

Comment clauses contain a subject and a verb, thus, resembling main clauses. What is more important, they do not have any subordinators. Comment clauses are not 'independent clauses, since they are defective syntactically: the verb or adjective lacks its normally obligatory complementation' (ibid.: 1114). In relation to the main clause, they may take the initial (3), medial (1) or final position (2) in a sentence, and they are decipherable by the changes in the voice, i.e. by the tone.

The analysis of adverbial clauses indicates different types of semantic relations between the main clause and the subordinate clause. Moreover, it highlights the role of

subordinators in deciphering between different semantic roles as well as allows us to distinguish coordination from subordination, which is irrelevant when determining the clause type in the compiled learner corpora.

2.3.2.3 Adjectival Clause

Another type of complex sentences that we distinguish according to the categorical principle of clause classification is the *adjectival* clause (in many grammars, e.g. Longman Grammar of Spoken and Written English (1999), Cambridge Grammar of English (2006), Nouvelle Grammaire du Français (2004), etc., this clause type is referred to as *relative clause* according to more common functional approach to clause classification). In the Concise Oxford Dictionary of Linguistics (1997) these two concepts are equalled, as their role in a noun phrase is the same, i.e. to modify a noun. Hence, within the framework of this chapter as well as for the purposes of the qualitative analysis of the empirical data the term 'relative clause' will be utilised to avoid ambiguity.

Relative clause has been defined by Matthews (1997) as the one which 'modifies the head of a noun phrase and typically includes a pronoun or other element whose reference is linked to it' (Matthews, 1997: 341). Jones (1996) postulates that it is 'a clause which is appended to an NP and which describes a property of the entity to which the NP refers' (1996: 498). Hawkins and Towell (2001) remark that it is a clause 'within a sentence which modify noun phrases or pronouns' (2001: 349). Thus, the main function of relative clause is that of a modifier of the noun. Moreover, it consists of a head and a restricting clause. Furthermore, relative clause may be sub-divided into adnominal, nominal and sentential relative clause.
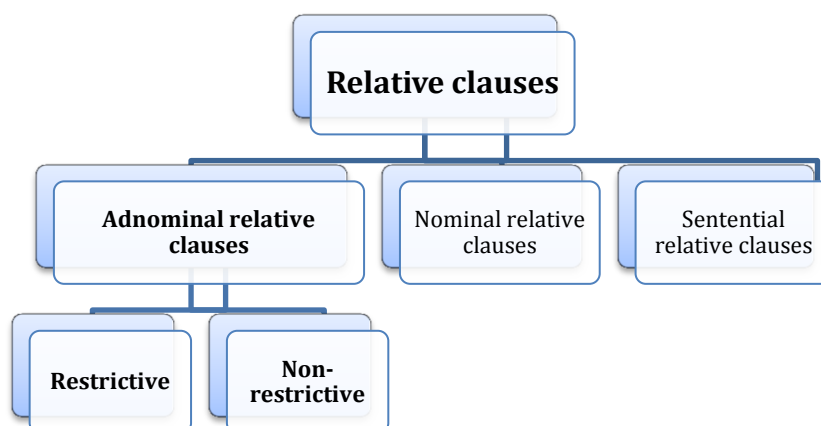
Figure 20 **Types of relative clauses**

According to Quirk et al. (1985), the main type of relative clause is called *adnominal* relative clause (many grammarians do not make this distinction but refer to it simply as 'relative clause'), which is further subdivided into *restrictive* and *non-restrictive* relative clauses based on the clause semantic relations in a sentence.

In a sentence *restrictive* relative clauses are directly linked to their antecedent (noun head) and 'denote a limitation on the reference of the antecedent' (Quirk et al., 1985: 366), e.g.:

This is not the most important thing [*that I would like to focus on*].

Ce n'est pas la chose la plus importante [*que je voudrais mettre en relief*].

It should be noted that restrictive relative clauses do not specify proper nouns or disjunctive pronouns. Moreover, this clause type is not separated by commas from the rest of the sentence.

Non-restrictive relative clauses are 'parenthetic comments which usually describe, but do not further define, the antecedent' (ibid.), e.g.:

The decision was made by the politicians, [*who do not regard things seriously*].

La décision a été prise par les politiciens, [*qui ne prennent pas les choses au sérieux*].

In written language commas always separate non-restrictive relative clauses from other parts of a sentence, which is not the case in the spoken language where intonation and pauses help to differentiate between the two clause types. As for the frequency of use of the two clauses, the restrictive relative clause is more common than the non-restrictive one.

Relative clauses can be further subdivided into *nominal* relative clauses (see chapter on noun clauses), which function as a noun phrase, and *sentential* relative clauses, in which the antecedent consists of a clause:

They can perform different stunts, [*which is unbelievable*].

Ils peuvent faire des trucs différents, [*ce qui est incroyable*].

In the above examples the sentential relative clause [*which is unbelievable*] / [*ce qui est incroyable*] postmodifies [*They can perform different stunts*] / [*Ils peuvent faire des trucs différents*] by expressing comments about the ideas given in the main clause, and they are always introduced by the relative pronoun *which* in English and *ce* in French.

Relative clauses are being distinguished from other types of clauses by the *relative pronoun*, which always appears at the beginning of the clause. Relative pronouns decipher: 1) between restrictive and non-restrictive relative clauses (especially in English); 2) determine the gender of the antecendent (personal or non-personal, or also neuter in

French). Relative pronouns can function as subject, object, prepositional complement and possessive determiner in the relative clause (see *Table 12*).

Table 12 **Relative pronouns in English and French**

Functions of relative pronouns ↓	Restrictive relative clause					Non-restrictive relative clause					
	personal		non-personal		neuter	personal		non-personal		neuter	
Subject	who that	*qui*	which that	*qui*		who	*qui*				
Object	who(m) that zero	*que*	which that zero	*que*		whom	*que*	which	*que*		
Prepositional complement or adverbial	who(m) that zero	*qui* *lequel*	which that zero	*lequel*	*quoi*	which	*qui*	which	*lequel*	*quoi*	
Possessive determiner	whose / *dont*										

In English relative pronouns might be divided into two groups: 1) *wh*-pronouns (*who, whom, whose, which*); 2) *that* and *zero* pronoun. Neither of the two groups has number or person. Nonetheless, *wh*-pronouns decipher genders (*who* − personal; *which* − non-personal) and cases (*who* − subjective; *whom* − objective; *whose* − genitive). The pronoun *whose* can have both personal and non-personal reference, and the pronoun *whom* is most frequently used in formal style (especially writing). Thus, it could be avoided in informal style and substituted by *who*, *that* or *zero*. In English the usage of relative pronouns *who, which, that* is determined by whether they refer to a person or an object.

In French the main relative pronouns are *qui*, *que* and *dont*. The subject relative pronoun *qui* and the object relative pronoun *que* do not decipher genders. Moreover, pronoun *que* is not to be omitted, as it is the case with the English relative pronoun *that*, which can be excluded in restrictive relative clauses. Thus, the relative pronoun *que* is more valid in French than the pronoun *that* in English. The pronoun *dont* can refer either to personal or non-personal antecedent, as it is also in English with the possessive pronoun *whose*. However, in French *dont* is used also when 'the noun phrase or pronoun heading a relative clause is the implied object of *de* in that relative clause' (Hawkins and Towell, 2001: 354), e.g. Le livre [*dont j'ai besoin*]… (J'ai besoin *de* ce livre).

When having the function of prepositional complement, in which the relative pronoun is introduced by a preposition, pronouns *that, which* and *zero* in English denote both personal and non-personal antecedent, e.g. The other <u>girl</u> [***that I told you about***] lives just across the street *or* The department <u>store</u> [***that I told you about***] is just over there. In

French the pronoun *qui* refers only to a personal antecedent, e.g. Le <u>garçon</u> [***à qui** j'ai parlé*] vient du Canada, whereas *lequel* with its variants: *laquelle, lesquels, lesquelles* (*lequel* has to comply in number and gender with the noun it modifies) are used to refer to a non-personal antecedent (e.g. Elle lui a dit le <u>nom</u> du magazine [***pour lequel** elle travaillait*]). When preceded by prepositions *à* or *de,* the component *le-* has to be modified according to the number and gender of the antecedent it describes, e.g. Le <u>livre</u> [***auquel** il fait reference*] est sur la table. Moreover, *lequel* is applied to avoid ambiguity when referring back to one of the several antecedents as it designates both gender and number (see e.g. sentence *b*):

> (a) Voici le fils de ma mère, [*dont je t'ai parlé*].
>
> (b) Voici le <u>fils</u> de ma mère, [***duquel** je t'ai parlé*].
>
> This is the son of my mother, [*whom I spoke to you about*].

Sentence (a) is ambiguous since it is not clear whether the speaker means *the son (le fils)* or *the mother (ma mère),* while in sentence (b) the relative pronoun *duquel* clearly indicates that the speaker refers to *the son (le fils)* as *duquel* in French denotes masculine gender.

The neuter relative pronoun *quoi* is mainly used in written French, and appears after prepositions to modify such non-personal items as *ce, rien* and *quelque chose*, e.g. Il n'y a <u>rien</u> [*à quoi je puisse penser*]. In non-restrictive relative clauses *quoi* modifies a clause, e.g. L'avion était en retard, [*sans quoi ma mère serait arrivée à l'heure*].

Sometimes relative pronoun *that* is used instead of *which* and *who,* and analysed as a subordinating conjunction in declarative clauses, e.g. She said [<u>*that*</u> *she was happy*]. Thus, in the sentence *I didn't like the man* [*that met us*] 'that' will not be a pronoun but simply a conjunction, which means that the relative clause will not have a subject. In French this phenomenon can be attributed to *que*, e.g. *le chapeau que j'ai vu* (*que* is more like a conjunction than a relative pronoun).

If we look at the structure of the main clause and the relative clause introduced by *that / que*, it is obvious that relative clauses are structurally incomplete:

> (a) that Ø went away
>
> (b) that he had written Ø
>
> (c) qu'il avait écrit Ø

Example (a) does not have a subject while examples (b, c) do not have any object. It should be stated that in French the structural incompleteness refers only to the relative clause introduced by *que* functioning as object as in (c). Moreover, in English the

subordinating conjunction *that* could be omitted, which is not the case in French. Another aspect worthy of attention is that the antecedent determines the person and number of the verb:

      (a) the <u>one</u> that *wakes* up early

             <u>celui</u> qui *se réveille* tôt

      (b) the <u>ones</u> that *wake* up early

             <u>ceux</u> qui *se réveillent* tôt

In French, if the antecedent is personal, the relative pronoun *qui* has to be used, whereas English allows utilizing both *who* and *that* to refer to the personal antecedent.

If we compare the relative clause *whom I met last week* (*que j'ai rencontré la semaine dernière*) of *the girl whom I met last week* (*la fille que j'ai rencontré la semaine dernière*) with an independent sentence *I met the girl last week* (*j'ai rencontré la fille la semaine dernière*), then it is obvious that the word order is not the same. In the independent clause the direct object *the girl (la fille)* follows the main verb *met (ai rancontré).* In the relative clause the relative pronoun *whom (que)* has been moved to the initial position of the clause, which is the main principle of relative clause formation. Comrie (1989) has elaborated a cross-linguistic Accessibility Hierarchy for relativization in a language (see also Chapter 1.1):

      Subject > Direct object > Non-direct object > Possessor

Thus, 'if a language can form relative clauses on a given position on the hierarchy, then it can also form relative clauses on all positions higher (to the left) on the hierarchy' (Comrie, 1989: 156). English and French are the languages which can have relativization on all four positions, which will be proved by the contrastive analysis of restrictive and non-restrictive relative clauses in both languages.

Restrictive relative clauses

When looking at the syntactic differences between restrictive and non-restrictive relative clauses, then fundamentally they have the same properties in both languages, i.e. the relative pronoun at the beginning of the relative clause modifies a personal or non-personal antecedent, which is a noun. However, the main difference lies in the function of the relative clause – whether it defines or describes the antecedent. Thus, *restrictive* relative clauses are the ones that *define* the preceding noun head.

In restrictive relative clauses in English *wh*-pronouns (*who, whom, which, whose*) and *that* or *zero (   )* pronoun are used while in French *qui, que, lequel, quoi, dont* are

utilised. In both languages (see *Table 12*) the relative pronouns perform the function of a subject, object, complement, adverbial (or in a preposiotional phrase it has a function of a complement) and possessive determiner of the relative clause which contains either personal or non-personal antecedents (in French also neuter antecedent when functioning as prepositional complement):

1) subject:

[1]   We are happy about

the person ⌈ who / *that* / *( ) ⌉ has won the award.

the play ⌈ which / that / *( ) ⌉ has been staged in the theatre.

In English in case the antecedent is personal (*the person*) and the pronoun is the subject of the relative clause, *who* is preferred instead of *that*. *Zero* pronoun cannot replace the subject, which is marked by *, in both cases – whether the antecedent is personal or non-personal.

[2]   Nous admirons ⌈ la personne *qui* a remporté le prix.
                        ⌊ la pièce *qui* a été mise en scène au théâtre.

In French the subject relative *qui* refers both to personal (*la personne*) and non-personal (*la pièce*) antecedent.

2) object:

[3]   We are happy about

the person ⌈ who(m) / that / ( ) ⌉ we have chosen for the post.

the play ⌈ which / that / ( ) ⌉ we have staged in the theatre.

When the antecedent is personal (*the person*) and the relative pronoun is used as object or prepositional complement, *that* and *zero* pronouns are preferred to *who(m)*.

[4]   Nous sommes contents ⌈ de la personne *que* nous avons choisie à ce poste.
                              ⌊ de la pièce *que* nous avons mise en scène au théâtre.

In French there is only one pronoun *que* used as object both for personal and non-personal antecedent.

3) prepositional complement :

[5]   He is the man ⎧ at whom   the little girl laughed.
                     ⎪ who(m)
                     ⎨ that   the little girl laughed *at*.
                     ⎩ (   )

[6]   This is the car ⎧ at which I threw stones.
                       ⎨ that ⌐ I threw stones (at).
                       ⎩ (   )

In English relative pronouns *who(m)*, *that* and *zero* functioning as prepositional complement are used to denote a personal noun, whereas *which*, *that* and *zero* pronouns specify a non-personal noun. If the relative pronoun is the complement of the preposition, there is a possibility to place the preposition either before the *wh*-pronoun (more often in formal English) or at the end of the sentence, which is not the case with *that* and *zero* pronouns.

In French when the relative clause is introduced by a preposition [7], the relative pronoun *qui* is used solely to refer to a personal antecedent:

[7]   C'est un homme *à qui* cette petite fille s'adresse.

For a non-personal antecedent the relative pronoun *lequel* is utilized, which has to agree in number and gender with its antecedent:

[8]   Voici la poupée *avec laquelle j'ai joué toute la journée*.

There are cases when the relative pronoun *lequel* with its variants can be used with personal antecedent, especially in plural cases:

[9]   Ce sont les élèves *parmi lesquels je dois trouver le meilleur*.

Finally, the neuter relative pronoun *quoi* modifies non-personal antecedents (ce, rien, quelque chose) which do not contain a head noun:

[10]   C'est quelque chose *sur quoi je peux m'appuyer*.

As for the position of the preposition in French, it is always positioned before the pronoun unlike English where it can be either before the relative pronoun or at the end of the relative clause (examples [5] and [6]).

   4) adverbial:

In English and French relative pronouns may function as adverbials. In such a case the preposition + pronoun can replace the adverbial expressions of time, place and cause:

[11]   This happened at the time ⎧ during which ⎫ she stayed with her parents.
                                 ⎩ when   ⎭

[12]   Elle a décidé de venir le jour *où* j'ai dû partir.

In [12] the relative adverb *où* denotes time (in majority of cases *où* denotes place) and is used instead of the adverb *quand*, which does not have the function of a relative adverb. What is more, the relative adverb *où* can be utilized only if the antecedent is 'an implied time adverbial in that relative clause, and is definite' (Hawkins and Towell, 2001: 356). Hence, *où* cannot be replaced by *quand* unlike in English in which *when* performs the function of both adverb and relative adverb. Furthermore, if the antecedent is indefinite, the relative pronoun *que* should be used, e.g. Un jour *que je suis rentre a la maison…*

When denoting a place, we can use the adverb *where* in English and *où* in French instead of the preposition + relative pronoun.

[14]   That's the place ⎡ in which ⎤ I like to return.
                         ⎣ where   ⎦

[15]   C'est la place ⎡ dans laquelle ⎤ je voudrais être.
                       ⎣ où   ⎦

The adverbial *why* in English and *pourquoi* in French can also be used as adverbials when specifying the cause.

[16]   That's the reason ⎡ for which ⎤ many people leave Latvia.
                          ⎣ why   ⎦

[17]   C'est la raison ⎡ pour laquelle ⎤ beaucoup de gens quittent la Lettonie.
                        ⎣ pourquoi   ⎦

Sometimes it is preferable to utilise nominal relative clauses instead of the adnominal relative clauses in which *wh*-forms in English and the adverbs *quand, où, pourquoi* in French express time, place and cause to avoid repetition:

[17]   This happened *when she stayed with her parents*.

        Ça s'est passé *quand elle est restée avec ses parents*.

[18]   That's *where I like to return*.

        C'est *où je voudrais revenir*.

[19]   That's *why many people leave Latvia*.

        C'est *pourquoi beaucoup de gens quitter la Lettonie*.

There is no relative pronoun *how* in English and *comment* in French to express manner with an antecedent noun, therefore we can only have sentences like [20] and [21]:

[20]   That's ⌈ how ⌉ she learnt English.
         ⌊ the way (*that*) ⌋

[21]   Elle parle ⌈ comment ⌉ sa sœur aînée fait.
                    ⌊ de même façon *dont* ⌋

Thus, in English the relative pronoun *that* or *zero* pronoun replaces *how,* while in French instead of *comment* we use the form *dont* in relative clauses. It is very important to understand the different meanings and functions of relative pronouns since they can be used in various syntactic constructions.

    5) complement:

When used as complement, pronoun *which* is preferred to: 1) *who* and *that* if the antecedent is personal (e.g. *wife* in [22]); 2) *tha*t and *zero* pronouns if the antecedent is non-personal (e.g. *car* in [24]). In French there is only the distinction between *qui* (if the antecedent is personal – sentence [23]) and *que* (if the antecedent is non-personal – sentence [25]):

[22]   She is an excellent wife ⌈ which ⌉ the previous one was not.
                                  | *who |
                                  ⌊ *that ⌋

[23]   Elle est une excellente femme *qui* la précédente n'était pas.

[24]   This is the most expensive car ⌈ which ⌉ his own is not.
                                        | *that |
                                        ⌊ *( ) ⌋

[25]   C'est la voiture la plus chère *que* la sienne n'est pas.

    6) possessive determiner:

When used in the possessive case, the relative pronoun determines the noun it follows:

[26]   The lady *whose daughter* you like is the wife of the president. ['The lady is the wife of the president; you like her daughter.']

[27]   La dame *dont* vous aimez *la fille* est la femme du président. ['La dame est la femme du président; vous aimez sa fille.']

Syntactically the restrictive relative clause may be:

    1) subordinated to another clause:

       I visited <u>the house</u> [*in which my grandfather had lived*].

       J'ai visité <u>la maison</u> [*dans laquelle mon grand-père avait vécu*].

2) embedded in the main clause:

The project [*that we have just started*] will be a success.

Le projet [*qu'on vient de lancer*] sera un succés.

In both cases the relative pronoun is not separated from its antecedent.

If we examine the word order in the subordinate relative clause, the subject preceeds the verb both in English and French. However, in French when the subject is a noun and the verb does not have a complement, there might be *inversion*. Thus, we can have:

1) SV ➔ La peinture que cet artiste (S) fait (V) est extraordinaire.

2) VS ➔ La peinture que fait (V) cet artiste (S) est extraordinaire.

Non-restrictive relative clauses

Another type of relative clause is called non-restrictive relative clause, which provides additional information about the antecedent (both personal and non-personal), i.e. describes it. In non-restrictive clauses the proper noun as well as other expressions can be modified, which was not the case in restrictive relative clauses. In majority of cases the non-restrictive relative clause is marked by *wh*-pronouns in English, i.e. *that* and *zero* pronouns are excluded, whereas in French the same relative pronouns are utilized also in non-restrictive clauses. In the sentence relative pronouns may perform the same functions as they do in restrictive relative clauses, i.e. they can be subject, object, complement or adverbial:

1) subject:

[1]   I asked my mother, *who did not want to answer*.

J'ai demandé à ma mère, *qui ne voulait pas répondre*.

[2]   This new film, *which has been criticized*, is my favourite one.

Ce nouveau film, *qui a été critiqué*, est mon préféré.

2) object:

[3]   I asked my mother, *who(m) I met at the supermarket*.

J'ai demandé à ma mère, *que j'ai rencontrée au supermarché*.

[4]   This new film, *which Dorothy has just seen*, is my favourite one.

Ce nouveau film, *que Dorothy vient de voir*, est mon préféré.

3) complement:

[5]   He is Spanish, *which no one else in our class is*.

94

Il est espagnol, *ce que personne d'autre dans notre classe n'est*.

[6]   She wants a new hat, *which this red one in the window is*.

Elle veut un nouveau chapeau, *ce que celui-ci rouge dans la vitrine est*.

4) prepositional complement:

[7]   This is a new perfume, *about which there have been so many commercials*.

This is a new perfume, *which there have been so many commercials about*.

C'est un nouveau parfum, *dont il y a eu autant de publicités*.

From the examples above it is obvious that the number of pronouns used in non-restrictive relative clauses is limited to *who(m)* and *which* in English, and *qui, que, dont* in French. When relative pronoun functions as complement, only *which* is chosen for both personal and non-personal non-restricitve relative clauses in English, whereas in French *que* is being used for personal and *dont* for non-personal non-restrictive relative clause. *Zero* pronoun cannot appear, while *that* is applicable very seldom. In writing non-restrictive relative clause is always marked out by commas. Moreover, semantically non-restrictive relative clauses [8] resemble coordination [9] or adverbial subordination [10]:

[8]   My sister, *who has lived in France since childhood*, speaks fluent Latvian.

Ma sœur, *qui a vécu en France depuis l'enfance*, parle couramment letton.

[9]   My sister speaks fluent Latvian, and she has lived in France since childhood.

Ma sœur parle couramment letton, et elle a vécu en France depuis l'enfance.

[10]   My sister speaks fluent Latvian although she has lived in France since childhood.

Ma sœur parle couramment letton bien qu'elle a vécu en France depuis l'enfance.

As for the sentence structure, de Villiers et al. (1979) elicit two variables when describing the structure of sentences containing relative clause. The first one is the position of the relative clause in a sentence, i.e. its *embeddedness*. In this case *the NP* in the main clause is modified by the relative clause. If the *subject* is modified, then it is called '*centre-embedded*'; if the object is modified, the term '*right-branching*' has been applied. The second variable is 'the role that the head noun plays in the relative clause, called its *focus*' (de Villiers et al., 1979: 500). Thus, these two variables specify four types of relative clauses:

Embeddedness (role of complex NP)	Focus (role of head noun)		Example
Subject	Subject	(SS)	The cat [*that bit the dog*] chased the mouse. Le chat [*qui a mourdu le chien*] chassait la souris.
Subject	Object	(SO)	The cat [*that the dog bit*] chased the mouse. Le chat [*que le chien a mordu*] chassait la souris.
Object	Subject	(OS)	The cat bit the dog [*that chased the mouse*]. Le chat a mourdu le chien [*qui chassait la souris*].
Object	Object	(OO)	The cat bit the dog [that the mouse chased]. Le chat a mourdu le chien [*que la souris chassait*].

The subject relative clauses (SS and SO) are center-embedded whereas the object relative clauses (OS and OO) are right-branching. Non-restrictive relative clauses have the same internal syntax as restrictive relative clause, though the function of the non-restrictive clause is to describe not to define.

Conclusion

Having analysed the syntactic complexity of both English and French it should be concluded that syntax is a very complex system consisting of words grouped into phrases that build up sentences. No level in the hierarchy should be omitted when examining different types of sentences – simple, compound and complex, as it allows to comprehend not only the rules that underlie sentence generation but also the function of various subordinate clause types – noun, adverbial and adjectival - in relation to the main clause. Moreover, the contrastive analysis of complex sentences proved that overall English and French have: 1) the same sentence structure, i.e. the subordinate clause occupies the same position (initial, medial or final position) within a sentence in either language, which depends not only on the constituent order but also on the semantic function of the embedded clause; 2) subordinate clauses perform the same syntactic functions, though there might be variations, e.g. in the use of relative pronouns having the function of adverbial.

Consequently, the aforementioned aspects substantiate the irrelevance of theoretical grounding for carrying out the empirical part of the present research, which will be based upon the authentic data retrieved from the compiled written learner corpora and will serve as evidential basis for elaborating the syntactic criterial features applicable when assessing written performance - the main objective of the present research.

CHAPTER 3 LEARNER CORPUS IN RESEARCH ON SYNTACTIC FEATURES

3.1 Research Methodology and Procedure

The present chapter uncovers the research methodology that has been applied to reach the investigation goals as well as lists the research procedure details.

In order to investigate the syntactic patterns produced by the learners of English and French, the author of the present research has applied *a corpus-based method,* which has also been used extensively by such corpus linguists as Leech (1992), Biber (1993), Sinclair (1996), McEnery et al. (2006), etc. According to McEnery et al. (2006), a corpus-based approach towards linguistic studies has prospered considerably since the 1980s and nowadays occupies the key position in empirical studies of language, as linguists aim at studying the actual language in the texts occurring naturally by applying several research techniques and strategies to analyse and describe the findings. In addition, linguistic corpora have been widely used in language teaching and testing, since they facilitate the understanding the linguistic development of language learners.

Hence, the research on syntactic structures is based on the data retrieved from the compiled written text corpus, which consists of two sub-corpora: the sub-corpus of essays in English (containing 191 texts composed of 2,839 sentences, which is 44,387 tokens) and the sub-corpus of essays in French (containing 104 texts composed of 1,240 sentences = 27,653 tokens) (see *Figures 21 and 22*).
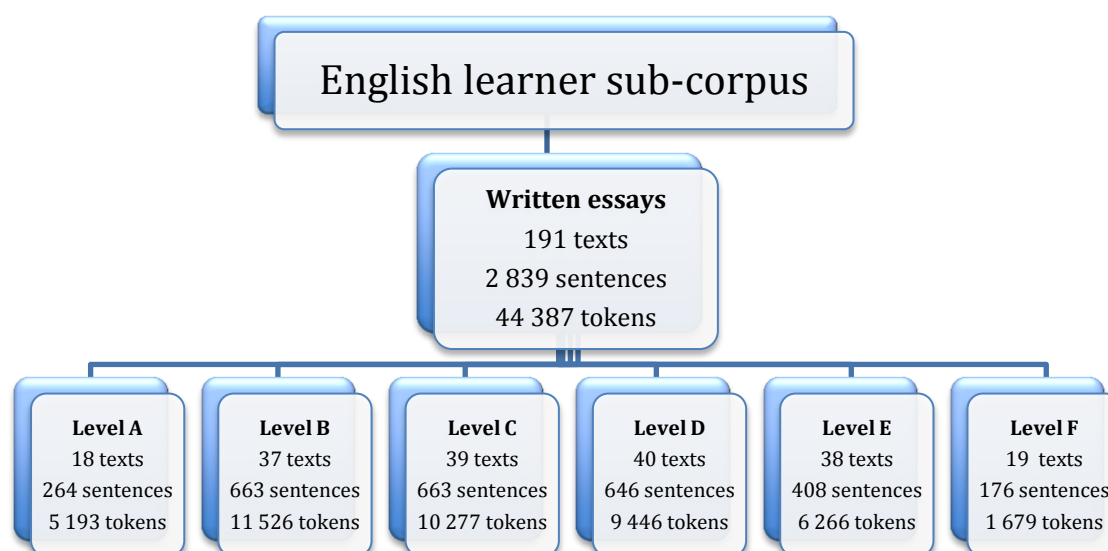
Figure 21 **The compiled English learner sub-corpus**
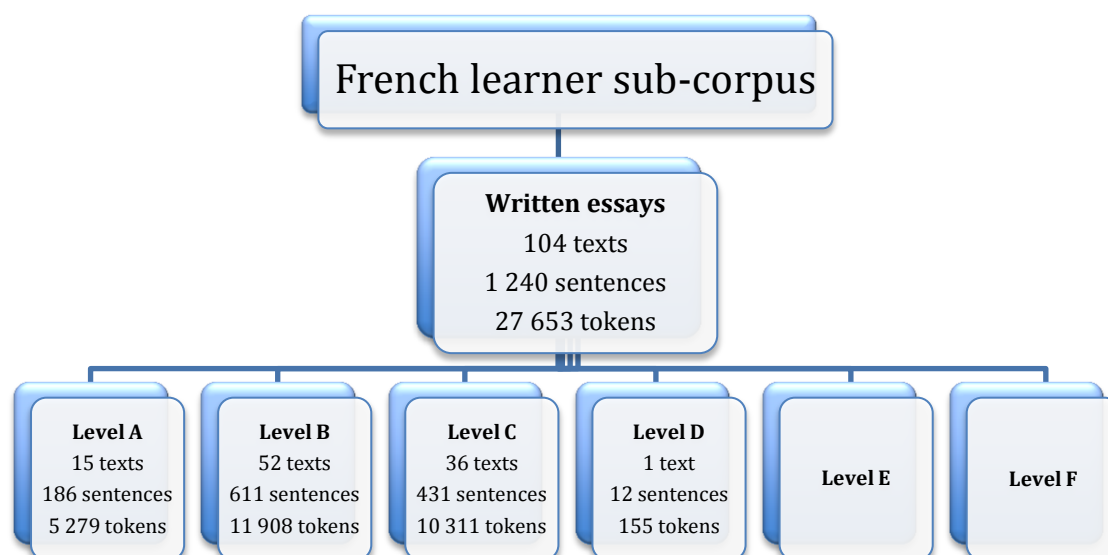
Figure 22 **The compiled French learner sub-corpus**

The methodology of sub-corpora creation was based upon the researcher theories on corpora classification, which may differ as each corpus is compiled with a particular aim for various research purposes. For example, McEnery et al. (2006: 59) classify corpora depending on their potential use into: general vs. specialised corpus, written vs. spoken corpus, synchronic vs. diachronic corpus, learner corpus and monitor corpus. Granger's (2003: 21) typology of corpora distinguishes between monolingual and multilingual corpora. Monolingual corpora contain texts in one and the same language and are further sub-divided into 1) original and translated texts or 2) native and learner texts, whereas in multilingual corpora the distinction is made between a) translation corpora (known also as parallel corpora) and b) comparable corpora containing either original texts or translated texts. Moreover, the above-considered corpus classification approaches reveal that there might be an overlap (one and the same corpus might be described in different ways). Nonetheless, it has been proved that they are a valuable source in 'pursuing various research agenda' (Meyer, 2002: 11), i.e. they can be used in various ways 'to validate, exemplify or build up a language theory' (Tognini-Bonelli, 2001: 65). Consequently, the compiled corpus for the present research (see *Figures 21* and *22*) is defined as:

- *bilingual* corpus as we compare original written essays produced by Latvian and/or Russian learners of English and French who took the national centralised examination in a foreign language in 2009;
- *specialized* corpus as the compiled texts are of a particular type, i.e. they represent a particular genre, namely student essays (Hunston, 2002: 14);

- *comparable* corpus because it consists of original texts of naturally occurring languages (Kenning, 2010) designed according to the same sampling frame (see e.g. Tognini-Bonelli, 2001; Kenning, 2010; McEnery, 2003; McEnery and XIAO, 2007; Hunston, 2002), which means that they match the same criteria, i.e. time of production, purpose, genre, intended audience, etc. (see *Table 14*). Moroever, they 'can […] form a valid basis of comparison' (Johansson and Hasselgård (1999: 146) by highlighting the similarities and differences across languages;
- *written* corpus - the samples are student essays, which were digitalised as the corpus was compiled;
- *learner* corpus since it can be used to recognise patterns characteristic of students (Tognini-Bonelli, 2001) learning a foreign language (in this particular case the learner corpus contains written essays produced by the learners of English or French as a foreign language).

These specific features of English and French Learner Written Text Corpus (EFLWTC) enabled its quantitative and qualitative analysis.

The *quantitative research method* has been applied in examining the frequency of simple, compound and complex sentences along with noun, adverbial and adjectival clauses retrieved from the EFLWTC by utilising the WordSmith Tools 5.0. This method has been used extensively by such researchers as Biber et al., 1998; McEnery, 2003; Leech, 2011, who acknowledge that with the advances in modern technology the computerized corpora allow accurate meta-data processing by counting frequencies and percentages which guarantee a reliable result. Hence, by applying *quantitative* research techniques it is possible to gather various quantitative data from a corpus ranging from simple words to complex syntactic structures. Moreover, the quantitative analysis is directly related to the classification and the *descriptive statistics* (see McEnery et al., 2006; McEnery and Hardie, 2012) of the studied phenomena.

The quantitative research is based on its main measure - descriptive statistics, which is a frequency count defined as 'a simple tallying of the number of instances of something that occurs in a corpus' (McEnery and Hardie, 2012: 49). Accordingly, in frequency count the central focus is on repeated elements rather than on single occurrences. Researchers (e.g. Leech, 2011 and Lindquist, 2009) distinguish three frequency types. The first type '*raw frequency*', which is a count of how many instances of some linguistic phenomenon X occur in some corpus, text or collection of texts, has been

considered by Leech (2011: 7-8) to be 'of little or no use in itself'. However, raw frequency has been used in the present research to count the total number of different types of sentences (simple, compound and complex) as well as different subordinate clauses (noun, adverbial and adjectival).

Another type of frequency *'normalized frequency'* or *'relative frequency'*, which is the frequency relative to a standard yardstick (e.g. tokens per million words) (Leech, 2011: 7-8), has been widely recognised and applied by researchers (e.g. Tono, 2002; Cosme, 2006; Hawkins and Filipović, 2012). Normalized frequency has been used in the present study to calculate the number of particular sentence and clause types relative to the total number of sentences or the total number of subordinate clauses in each sub-corpus. The obtained numbers have been expressed in percentages (see Appendix 7) to execute a reliable comparison of the data since the compiled sub-corpora differ in size.

The third type of frequency *'ordinal frequency',* which calculates the frequency of X compared with the frequencies of Y, of Z, etc. (Leech, 2011: 7-8), is considered to be most useful in language learning, e.g. a rank frequency list of different clause types may provide the data about which clause is more frequent than another one. Within the scope of the present research the linguistic data have been compared by applying ordinal frequency not only across different language proficiency levels but also cross-linguistically.

Another method applied in the present research is *the qualitative descriptive analysis* of various syntactic patterns extracted from the corpora and their functions by applying the Stanford NLP Parser 3.3.0. The qualitative analysis, as indicated by Svartvik (1992), has a significant role in corpus-based research since mere figures do not provide enough data. This view is supported also by McEnery and Wilson (2001), who state that 'it is not necessary to shoehorn the data into a finite number of classifications' (2001: 76). Hence, the aim of the qualitative analysis within the framework of the present research is to ensure a complete and detailed description of the studied features (e.g. the syntactic structure and the function of the relative clause in relation to the main clause) and provide richer and more precise information by focusing also on rare occurrences. What is more, Schmied (1993) reveals that in majority of cases the qualitative analysis precedes the quantitative analysis, as the linguistic phenomena have to be identified first and only then classified and counted, which is the case of the present research as well.

Thus, the quantitative analysis of the present research envisages collecting and analysing data by applying statistical methods (i.e. raw, normalized and ordinal frequency), whereas the qualitative research utilises the descriptive or interpretive analysis of different

clause types. However, a part of this study requires the use of *mixed methods*, i.e. both quantitative and qualitative, which, according to Heigham and Goker (2009), depends on the 'aims and context of the individual project and the nature of the research question' (2009: 5). Therefore, the multi-method approach, in which quantitative and qualitative method is used interchangeably, has been applied when analysing different clause types, thus complicating the systematic unveiling of the present research procedure and findings.

Finally, *the contrastive analysis method* has been applied to compare the quantitative and qualitative data across English and French in order to find the similarities between the syntactic structures produced by the test-takers in both languages. With the development of corpus linguistics, the contrastive analysis has proved to be 'a useful heuristic tool capable of throwing valuable light on the characteristic features of the languages contrasted' (Firbas, 1992:13). Furthermore, it envisages 'the linguistic comparison of (normally) two languages' (Hasselgård and Johansson, 2011) – the focus of the present research.

3.1.1 Text Sampling and Corpus Structure

In order to effectuate the study on syntactic features in the authentic texts produced by the test-takers of English and French, the EFLWTC (see *Figures 21* and *22*) was designed according to the methodology of the corpus compilation based upon the researchers definition of 'corpus' and the theories of its representativeness in the terms of the size.

In Latin the term 'corpus' means 'body'; hence, it would be logical to define *a corpus* as 'a body of naturally occurring language' (McEnery et al., 2006: 4) since 'in principle, any collection of more than one text can be called a corpus' (McEnery and Wilson, 2010: 29) However, in modern linguistics *corpus* has obtained a more elaborated connotation. According to Sinclair (1996), a corpus is a collection of pieces of language that are selected and ordered according to explicit linguistic criteria in order to be used as a sample of the language. The developers of Cambridge Learner Corpus emphasize also the size and the storage of the data; thus, 'a corpus is a large collection of samples of a language held on computer. The samples can come from anywhere the language is used in speech and writing' (Online 2). Leech (1992) complements by stating that 'computer corpora [...] are generally assembled with particular purposes in mind, and are often assembled to be (informally speaking) *representative* of some language or text type' (1992: 116). Consequently, the main features of a corpus in modern linguistics are: 1) it is

computer-readable; 2) contains authentic texts of finite size; 3) is representative of a particular language.

Thus, in order to obtain reliable results in the present research a corpus consisting of two sub-corpora was designed according to Biber et al., 1998 and McEnery et al., 2006 corpus creation methodology. It should be stated that EFLWTC creation was a time consuming and labour intensive process, as it comprised a range of corpus-creation steps.

The first step in corpus creation methodology is *sampling*, which ensures the representativeness of a corpus. Biber (1993) refers to *representativeness* as 'the extent to which a sample includes the full range of variability in a population' (1993: 243), i.e. how the texts representing each genre are chosen so that we arrive at representativeness. Manning and Schütze (1999) assume that a sample is representative 'if what we find for the sample also holds for the general population' (1999: 119). It means that the findings from the sample should reveal the characteristics of the population as much as possible. However, the author of the present research finds Leech's (1991) interpretation of 'representativeness' more appropriate as it states that 'a corpus is thought to be representative of the language variety it is supposed to represent if the findings based on its contents can be generalized to the said language variety' (1991: 27).

Within this context, before initiating the research a *sampling unit* was defined relying upon the criteria proposed by Tono (2002) for designing a new learner corpus, i.e. language-related (e.g. mode, medium, genre, topic, etc.), task-related (e.g. longitudinal vs cross-sectional; spontaneous vs prepared) and learner related (e.g. EFL or ESL, age, sex, mother tongue, etc.) criteria were taken into consideration. Thus, the written essays produced during the year 2009 centralised examination in a foreign language of secondary school graduates were chosen as a sampling unit. In English the test-takers had to write an essay about *'Reasons for Leaving Latvia'*:

One of the main reasons why people left Latvia during the last few years is that they say they are better paid in other countries. Add two other reasons and discuss all of them in an argumentative essay, giving your own opinion.

The following was the theme of the essay in French:

Pensez vous qu'il soit encore utile d'apprendre des langues étrangères alors que l'anglais est actuellement la langue de communication mondiale (échanges commerciaux, économiques, politiques...)? Présentez votre réflection de façon argumentée.

(*Do you think that it is still useful to learn foreign languages as nowadays English is the language of communication (in business, economics, politics…) in the world? Give your point of view by providing arguments.*)

Consequently, the written texts in both languages were opinion essays covering the topic related to the social domain (see *Table 14).* It should be noted that as the aim of the present research is to examine syntactic patterns across English and French, the focus has not been attributed to discourse analysis.

As for the learner-related criteria, the essays were written by the students coming from different L1 backgrounds (mostly Latvian or Russian). However, there were no data available to appoint the test-takers to one particular group. Neither do we have the age and gender of the test-takers, as such information was not to be disclosed by the NCE. Hence, the only data available related to the learners (see *Table 14*) is information about the test-taker's L2 (either English or French) and the level of language proficiency (A; B; C; D; E; F) awarded at the examination.

Table 14 **Criteria for a sampling unit**

language - related	task - related	learner - related

mode
  [written]
	data collection
	  [cross-sectional]
		internal-affective
		  [high stakes examination]

genre
  [essay]
	elicitation
	  [spontaneous]
		L1 background
		  [Latvian / Russian]

style
  [opinion]
	time limitation
	  [fixed]
		L2 background
		  [EFL / FFL]

topic
  [social]
		L2 proficiency
		  [levels A – F in EFL/ levels A– D in FFL]

Overall, a sampling unit in both languages met the same criteria, which provided a comparable basis for the present research.

Another aspect to consider when compiling a corpus is the *sample size*. First of all, it was decided how many texts should be included in the corpus, and secondly, whether the whole text or document has to be chosen or only some part of it. Bowker and Pearson (2002) hold the view that a corpus should contain 'a greater number of texts than you would be able to easily collect and read in printed form' (2002: 10). However, the corpus

that is developed, for example, for quantitative studies of grammatical features and needs to be annotated manually, has to be 'necessarily small' (McEnery et al., 2006: 72). What is more, Biber (1990) acknowledged that ten texts for each category in LOB (Lancaster-Oslo-Bergen) corpus were enough for studying grammatical features. Hence, if these are syntactic features to be studied, then the size of a corpus is not that weighty at all, since the number of syntactic variations is rather restricted or, as Hakulinen et al. (1980: 104) state, 'the syntactic freezing point is fairly low.' Furthermore, it has been proved that syntactic structures are usually more difficult to extract from a corpus than simple words (Guilquin, 2002: 183). Though the author of the present research agrees with Sinclair's viewpoint that 'if within the dimensions of a small corpus, using corpus techniques, you can get results that you wish to get, then your methodology is above reproach' (Sinclair, 2004: 189), we are to remember that 'the numbers of samples across text categories should be proportional to their frequencies and/or weights in the target population' (McEnery et al., 2006: 20).

Thus, when creating EFLWTC within the framework of the present research its size was determined by: 1) the total number of students who took the foreign language examination since it differed greatly with respect to the chosen foreign language; 2) the number of texts per each language proficiency level.

Table 15 **Number of test-takers per year**

	2009	2010	2011	2012	2013
English	23,652	22,638	22,174	19,160	16,968
French	104	71	77	49	53

*Data available from http://visc.gov.lv/vispizglitiba/eksameni/statistika

In 2009 (see *Table 15*) the examination in English as a foreign language (EFL) was taken by 23,652 students, which drew up almost the whole population of school graduates, while the examination in French as a foreign language (FFL) was taken only by 104 students (mainly students from language schools), which comprised only 0.43% of the test-takers in English.

Since the English examination was taken by the biggest part of the school graduate population, the obtained results (see *Table 16*) were dispersed across all levels of language proficiency A to F (A – being the highest, and F – the lowest), whereas in French the examination results ranged from levels A to D, as this examination was taken only by very motivated language school graduates. Hence, in French there were no texts of lower levels

E and F to be sampled (*Table 16*) illustrates the distribution of test-takers across levels in both languages).

Table 16 **Number of test-takers and the obtained level at the examination in 2009**

Level	A	B	C	D	E	F
English	1029 (4.35%)	4259 (18.01%)	6318 (26.71%)	6646 (28.10%)	4268 (18.04%)	1132 (4.79%)
French	15 (14.42%)	52 (50%)	36 (34.62%)	1 (0.96%)		

*Data available from http://visc.gov.lv/vispizglitiba/eksameni/statistika

Furthermore, the percentage of the students who were awarded each particular level of language proficiency was not similar; in English it ranged from 4.35% at level A to 28.10% at level D, while in French it comprised 14.42% at level A; 50% at level B; 34.62% at level C and only 0.96% at level D. Therefore, in order to ensure the balance when designing the corpora, a proportional number of texts representing each level of language proficiency were chosen. Moreover, the texts were selected by applying the stratified random sampling technique, i.e. they were selected at random to represent each level of language proficiency (in French all the texts were included in the corpus because of the limited number of test-takers per year). Consequently, all the samples were arranged according to the proficiency level obtained at the examination.

When aiming for a balanced corpus, Sinclair (1991) and Biber (1993) suggest including only segments of texts as larger samples are required if the texts contain rare features. Nonetheless, there are no criteria to measure how balanced a corpus is as it merely relies on the corpus-builder intuition. What is more, Atkins et al. (1992) postulate that sometimes we may judge whether the corpus is balanced or not only after having designed the entire corpus.

Therefore, for the purpose of the present research it was decided to include the whole texts produced by the test-takers, since their size was not that bulky (it varied from 404 tokens to 13 tokens per text depending on the level of the student language proficiency). Though the number of texts was chosen to represent proportionally each language proficiency level, their number per level varied depending on the total amount of students who were awarded the particular level at the examination. What is more, in French all the texts were included in the corpus since the number of test takers comprised 104 students. However, there was only one text for level D and no texts to represent the lower levels of language proficiency (E and F) in French as the examination was taken mainly by the students from language schools. Thus, the compiled corpus consists of two sub-corpora: 1)

the English sub-corpus (see *Figure 21*) contains overall 191 texts (1240 sentences or 27653 tokens) sub-divided accordingly to different levels of language development (level A comprises 18 texts (264 sentences / 5193 tokens); level B – 37 texts (663 sentences / 11526 tokens); level C – 39 texts (663 sentences / 10277 tokens); level D – 40 texts (646 sentences / 9446 tokens); level E – 38 texts (408 sentences / 6266 tokens) and level F – 19 texts (176 sentences / 1679 tokens)); 2) the French sub-corpus (see *Figure 22*) consists of 104 texts composed of 1240 sentences or 27653 tokens: level A contains 15 texts (186 sentences / 5279 tokens); level B – 52 texts (611 sentences / 11908 tokens); level C - 36 texts (431 sentences / 10311 tokens) and level D – 1 text (12 sentences / 155 tokens). Overall, the compiled corpus comprises 295 texts (4079 sentences or 72040 tokens).

The following step incorporated the creation of the electronic version of the target texts, i.e. the texts had to be *digitalised,* which was rather a labour intensive and time-consuming process. Withal, the corpora had to be created very carefully since they were to provide authentic data for linguistic analysis, whose results had to be valid. According to Meyer (2002), if a corpus is 'haphazardly created, with little thought put into its composition, then any analysis based on the corpus will be severely compromised' (2002: xiv). Moreover, a digitalised corpus enables the retrieval of linguistic data for research purposes by applying different corpus techniques. Finally, the electronic texts were classified in different file directories according to the language level of proficiency, and each electronic sample was saved in a separate file both in the document format and the ASCII format (as suggested by Meyer, 2002), so that it could be accessed easily whenever incumbent.
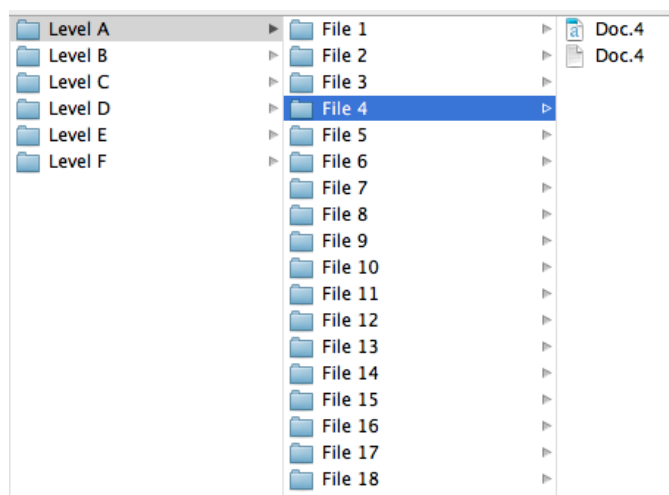
Figure 23 **Sample of electronic text classification**

3.1.2 Corpus Mark-up

The next stage in the corpus creation procedure is its *marking*. According to McEnery et al., 'corpus mark-up is a system of standard codes inserted into a document stored in electronic form to provide information *about* the text itself and govern formatting, printing or other processing' (2006: 22). Several metadata mark-up schemes have been elaborated, which differ in their complexity, e.g. the Corpus Encoding Standard or the Text Encoding Initiative (TEI) (see McEnery et al., 2006; McEnery and Wilson, 2010). For the purposes of the present research the sample header patterns of the structural mark-up according to the TEI standards were explored and the template for the corpus files was selected (see *Figure 24*).

```
<teiHeader>
 <fileDesc>
 <titleStmt>
 <title> Opinion essay: digitalised</title>
 </titleStmt>
 <publicationStmt>
 <p>Published for the research on syntactic structures
</publicationStmt>
 <sourceDesc>
 <p> Year 2009 centralised examination in English; code
05130200-006
 </sourceDesc>
 </fileDesc>
 </teiHeader>
```

Figure 24 **Template for the file structural mark-up**

Thus, each template contains the information about the text type (opinion essay: digitalised), the reason for publication (published for the research on syntactic structures) and the source description (year 2009 centralised examination in English/French; code attributed to the test-taker). Finally, the created template was copied in each text allowing the researchers to operate easily with the required data.

3.1.3 Corpus Annotation

In order to carry out the linguistic data analysis, the compiled corpus was *annotated* by adding 'interpretative, linguistic information' (Leech, 1997: 2) to the linguistic data. Leech (1997) and McEnery (2003) hold the view that an annotated corpus is useful for several reasons: 1) being annotated it allows the researcher to retrieve the necessary information from them; 2) it can be reused for other linguistic enquiries; 3) it is multifunctional since it can be used for a completely different purpose compared to the initial one; 4) it preserves

the linguistic analysis unequivocally. Thus, an annotated corpus serves as a 'stable base of linguistic analyses' (McEnery et al., 2006: 30).

Researchers distinguish several methods (see e.g. McEnery et al., 2006; Lindquist, 2009; McEnery and Wilson, 2010) in annotating a corpus – *automatic* (i.e. computer does the annotation using special programmes, e.g. in the case of part-of-speech (POS) tagging); *computer-assisted* (when the outcome of the automated approach is not reliable and the human correction in needed afterwards, e.g. when 'parsing' is applied) and *manual* annotation, which is expensive and time-consuming, thus applicable only when annotating small corpora. Furthermore, corpus annotation may be performed on different linguistic levels (phonological, morphological, lexical, syntactic and discourse level) by applying special annotation tools for each particular level, e.g. POS tagging is applied when annotating parts-of-speech.

As the compiled learner corpus contains authentic texts with errors, no automatic tagger could be utilised. Moreover, the present research objectives required the analysis only on syntactic level. Therefore, the author applied *problem-oriented tagging* (de Hann, 1984), which differs from all the aforementioned types that can be used for a wide-range of research by employing generally accepted criteria. McEnery and Wilson (2010) emphasize two major aspects that differentiate the problem-oriented annotation from the other types: 1) problem-oriented tagging is not 'exhaustive' since 'only the phenomena directly relevant to the research are annotated' (2010: 69). For example, the annotation of only the main types of sentences within a corpus; 2) problem-oriented tagging uses 'an annotation scheme which is selected not for its broad coverage and consensus-based theory-neutrality' but for its relevance to the particular research question (ibid.). Thus, the researcher is the one to decide which annotation scheme to utilise to study the specific phenomenon. Consequently, the syntactic marking was accomplished manually by assigning a corresponding grammatical tag-set for each sentence type (<SS> for simple sentences; <CS> for compound sentences and <CLS> for complex sentences).

```
C'est sûr que l'anglais est actuellement la langue de communication
mondiale mais je crois qu'il est nécessaire d'apprendre des langues étrangères.
Il me semble qu'il soit encore très utile de savoir parler quelques langues pour
avoir le succès dans la vie moderne.<CLS>
       Premièrement, si on parle français, on peut voyager et communiquer avec
les Français ou les Canadiens sans aucuns problèms.<CLS> Je le trove très
pratique et ça m'ouvre beaucoup d'opportunitès.<CS>
       A mon avis, il y a beaucoup de gens qui ne parlent que la langue maternelle
et ça me paraît très insuffisant parce que l'on ne peut pas utiliser quelques
appareils chez nous, par example, l'ordinnateur, Internet et d'autres choses.
       En plus, je pense que l'apprentissage des langues étrangères est très
interessant et amusant parce que ça nous donne une chance à se rendre avec les
gens et faire connaissance.<CLS> Je crois que l'interaction culturele fait une
grande part du développment mondial.<CLS>
       Je parle russe, letttone, anglais, français et polonais.<SS> Ça me plaît
beaucoup et je vais apprendre l'espagnol et peut-être le chinois.<CS> Je suis
très heureux de communiquer avec tous les gens parce que c'est une experience
inoubliable.<CLS>
```

Figure 25 **Sample of text annotation**

When undertaking the procedure of syntactic tagging, the errors produced by the students sometimes encumbered the identification of a sentence type and assigning an appropriate tag. Therefore, the expert viewpoint was of great importance.

3.1.4 Concordancer-based Data Extraction

Having accomplished the compiled text annotation, the quantitative analysis of the required data could be undertaken, which was effected by using the concordancing tool *WordSmith Tools 5.0* developed by Lexical Analysis Software Ltd. and Oxford University Press (Scott: 2005). The WordSmith Tools enabled to execute the frequency count of simple, compound and complex sentences. Before downloading the ASCII format sample to the programme, all the texts representing a particular level of language proficiency, e.g. level A, were gathered in one file. Afterwards, the file was downloaded into the programme 'Concord', which enabled to search the specified tag, e.g. <CLS>, through the text. The obtained list of the particular clause type with its frequency of occurrence was displayed on the screen (see the sample below).

Figure 26 **Sample of concordancing on tags**

The same procedure was repeated for each sentence type at each level of language proficiency in both languages, which provided the data of raw frequency of different sentence types across levels as well as languages.

3.1.5 Quantitative and Qualitative Data Contrastive Analysis

Subsequently, complex sentences of each language proficiency level were separated from simple and compound sentences to be subdivided into finite and non-finite clauses according to Dik's taxonomy of embedded constructions (see Chapter 2). This procedure was accomplished manually, as well as the frequency count of complex sentences containing finite clauses. Furthermore, the obtained data were compared across levels in English and French.

Furthermore, only the finite clauses were selected for the advanced research, since they were the main object of the present study. Before initiating the research on different clause types, the number of complex sentences containing one subordinate clause were compared with the ones comprising more than one subordination. The comparison was effected across levels in both languages. Moreover, the total number of subordinate clauses per level was counted by applying the formula:

$$\frac{N^o \text{ of mono sub-clauses} + N^o \text{ of multi sub-clauses}}{\text{Total } N^o \text{ of complex sentences}} \times 100\%$$

110

In order to obtain the information about the average number of sub-clauses per level, the mean number of sub-clauses was calculated by using the formula:

$$\frac{N^o \text{ of mono sub-clauses} + N^o \text{ of multi sub-clauses}}{\text{Total } N^o \text{ of complex sentences}}$$

At the next stage the complex sentences were tag-stripped by applying the PC software 'Complete Lexical Tutor v.6.2' (developed by Tom Cobb (1999) of the University of Quebec at Montreal) and divided into two groups: mono and multi sub-clauses. Then both types were manually tagged again, but this time a grammatical tag of different subordinate clause types, such as <N> for noun clause; <ADV> for adverbial clause and <ADJ> for adjectival clause, was assigned to each mono sub-clause. The same procedure was accomplished for multi sub-clauses by assigning a more complex tag, e.g. <ADJ, ADV, ADV>. Afterwards, the ordinal frequency count of each clause type across levels in both languages was effected (see Appendix 7).

In order to perform the qualitative analysis of different mono sub-clauses, noun, adverbial and adjectival clauses were subdivided into various types: noun clauses - into declarative, interrogative, exclamative, nominal relative clauses; adverbial clauses – into time, place, manner, comparison, circumstantial and comment clauses; adjectival clauses – into restrictive and non-restrictive relative clauses. Before initiating the analysis of various functions of the aforementioned sub-clauses, the ordinal frequency of each type was counted to have the clear picture of their distribution across levels.

Subsequently, the qualitative and contrastive analysis of different functions of sub-clauses in relation to the main clause was accomplished across languages. The example sentences were chosen to represent each syntactic feature typical of various clause types produced by the test-takers at different levels of language proficiency. It should be specified that the sentences were not error free since it was crucial to reveal the test-taker different levels of linguistic competence.

In order to effectuate a more profound analysis of various subordinate clauses, the author employed *the Stanford NLP Parser v.3.3.0* developed by the Stanford Natural Language Processing Group in 2006, 2010. The Stanford NLP Parser is based on the phrase-structure grammar, and distinguishes between different clause relations (e.g. coordination or subordination) in a sentence, thus ensuring full parsing of the sentence (represented as a tree diagramme) to be analysed. It should be specified that the parser was

used for the qualitative analysis of a limited number of subordinate clause types to exemplify the range of syntactic features. Moreover, in cases when the example sentences contained errors, they had to be /corrected, as the Stanford NLP Parser would generate the tree diagramme inappropriately.

Before commencing the qualitative analysis of the multi sub-clauses, it was decided to examine only the most frequent types of multi sub-clauses, i.e. the ones whose number of occurrence exceeded fifteen across all levels of language proficiency (see *Table 27*). The qualitative analysis of multi sub-clauses was also effectuated by applying the Stanford NLP parser.

Furthermore, the syntactic structures extracted from the test-taker corpora were compared with the grammatical criterial features for each language proficiency level (A1 – C1) provided by the English Profile (2011) as well as with the level descriptors (see Appendix 8) elaborated by the test administrators of NCE to assess student written performance cross-linguistically. This enabled the author to develop the syntactic criterial features in both languages for assessing the learner linguistic competence at each level of language development. The criterial features were organised according to different clause types, and each criterial feature was provided an example sentence (see *Tables 28 and 29*).

Finally, the elaborated criterial features were empirically validated by the English and French teachers, assessors and test developers, and the conclusions about the research results were drawn.

3.3 Research Findings and Discussions

The quantitative research results show (*Figures 27 and 28*) that *simple sentences* have been used most extensively at all levels of language proficiency. In the ELWT corpus their number ranges from 32% at level B; 36% at level A to 51% at level F. In the FLWT corpus the frequency of simple sentences also increases towards the lowest level of language proficiency attaining 67% at level D. Thus, the obtained data fully agree with the CEFR level key distinguishing features in which simple sentences are the discriminatory clause type at level A2/D.

The number of *compound sentences* used by the test-takers both in English and French is rather low. In English their number varies from 10% at levels A, D, F to 13% at level E. In French the situation is rather similar – 7% at level C; 10% at level B and 12% at level A (there is a tendency for a slight increase towards the highest levels of language

proficiency in French). If compared to the total number of sentences in English and French learner corpora, then the average number of coordinate clauses constitutes 11% in English and 9% in French, which is considerably low.
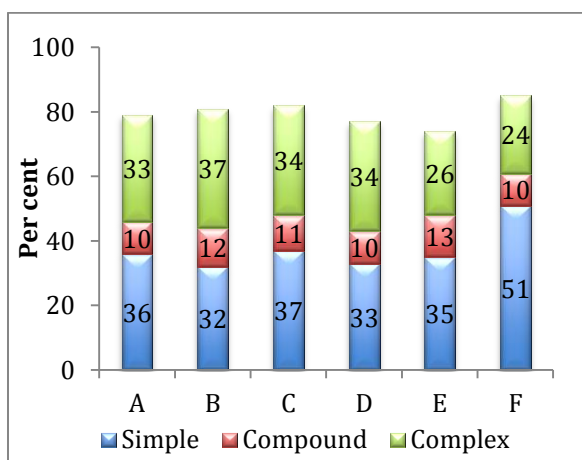
Figure 27 **The frequency of clause types in ELWTC**
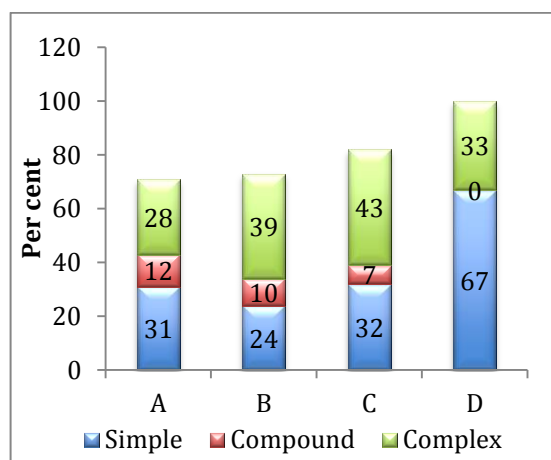
Figure 28 **The frequency of clause types in FLWTC**

Figures 27 and 28 provide evidence that *complex sentences* are the ones that might serve as a discriminatory element when attributing a certain level at the examination as their number steadily increases from the lower levels to the higher levels of language proficiency, which is also indicated in the CEFR by considering complex sentence to be a distinguishing feature already at level B1/C. Moreover, it complies with Pienemann's Processability theory (Chapter 1.1), according to which, subordinate clauses are processed in the final stage of language development. However, the research data show that the frequency of complex sentences in both languages is different. In English they constitute 24% at level F, then gradually rise to 34% at levels C and D attaining their peak at level B (37%). In French the complex sentences are unevenly distributed. The majority of them appear at levels C (43%) and B (39%). At level A their number is not that high, which might be explained by the tendency of students at higher levels of language proficiency to use rather many complex-compound sentences, thus producing fairly long and complicated syntactic structures. On average, the number of complex sentences used by the test-takers of English constitutes 31.33%, whereas in French - 35.75%. Hence, the contrastive analysis of different clause types proves the postulation of Chuquet and Paillard (1989: 151) that French uses subordination more intensely than English.

Further on, complex sentences were classified using Dik's (1997) taxonomy of embedded constructions (Chapter 2) as it could be attributed cross-linguistically. Thus, the complex sentences were sub-divided into the finite and non-finite clauses. It should be

specified that the present research focalizes on finite embedded clauses, as they are the ones that make a subordinate clause (further referred to as finite sub-clauses). Moreover, it is assumed that non-finite clauses (especially *–ing* clauses) are typical of English.
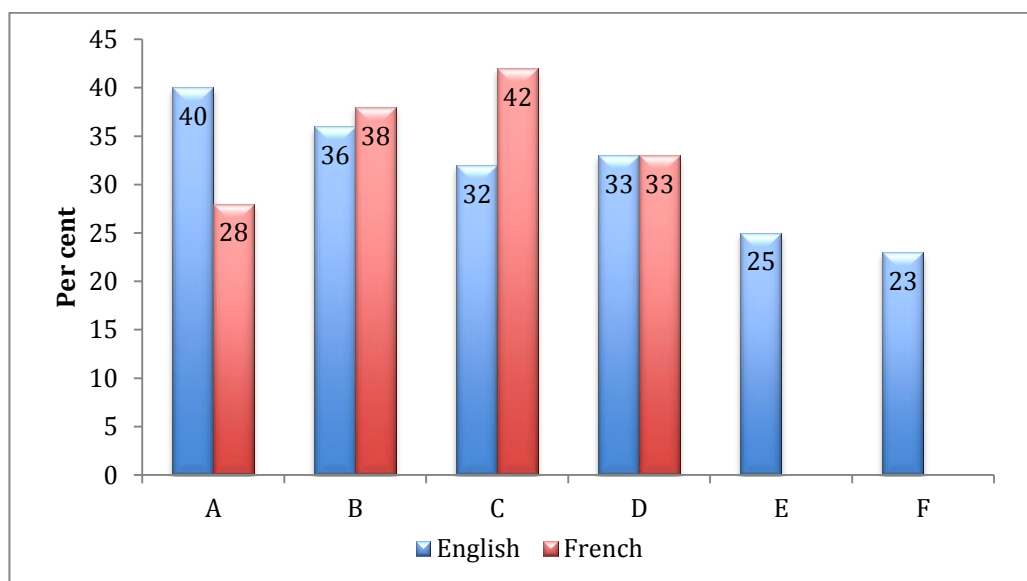
Figure 29 **Frequency of complex sentences containing finite sub-clauses**

According to statistics, we can observe a considerable increase of finite sub-clauses in English towards the higher levels of language proficiency, i.e. from 23% at level F to 40% at level A. In French the distribution of sub-clauses across levels is not similar. At level D the column reaches the same number as in English (33%) while at level C there is a rapid increase as the numbers attain 42%, which is the highest level on the diagramme (in English the number of sub-clauses at this level is considerably lower - 32%). However, at levels B and A there is a serious decline in numbers falling down to 28% at level A, which is 12% less than in English at the same level of language proficiency. Even so, the average number of finite sub-clauses across levels A – D (there are no sentences at level E and F in French to compare with) is similar in both languages – 35%. Thus, it could be claimed that the learners of English use as many finite sub-clauses as the learners of French though their number differs across the levels of language proficiency, which rebuts the assumption that non-finite clauses are more frequent in English than in French. Consequently, French should not be considered as more hierarchical than English.

As complex sentences may contain more than one subordinate clause, and to verify the supposition that at higher levels of language proficiency students produce more than one subordinate clause, the sentences were divided into mono and multi subordinate clauses, i.e. complex sentences containing only one subordination and complex sentences consisting of two and more subordinate clauses either hypotactic or embedded ones.
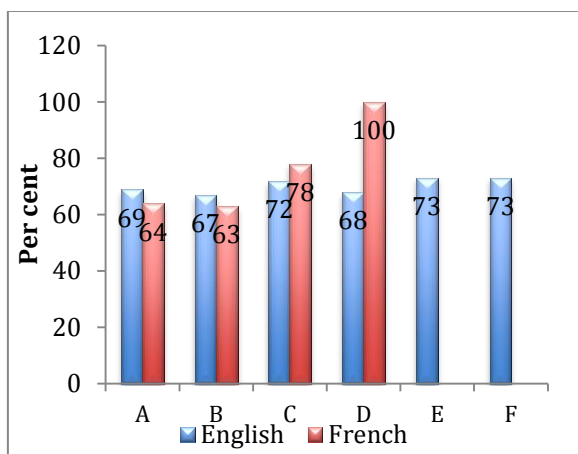
Figure 30 **Ordinal frequency of mono sub-clauses**
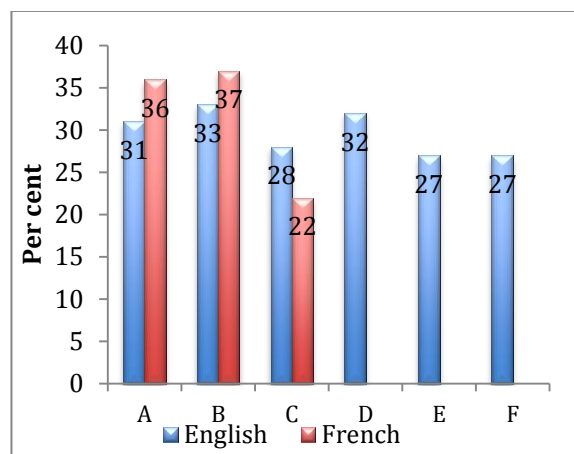
Figure 31 **Ordinal frequency of multi sub-clauses**

The data show that the number of mono sub-clauses at all levels of language proficiency in both languages is distributed rather evenly (*Figure 30*) ranging from 68% to 73% in English and 63% to 100% in French (at level D among all the sentences marked there is only one subordinate clause, which constitutes 100%). However, slight variations can be observed, e.g. in English at level A the column reaches 69%, but in French – 64%; at level B the number of mono sub-clauses in English comprises 67% while in French – 63%. Nevertheless, at level C the percentage of mono sub-clauses in French surpasses the one in English attaining 78%. The distribution of multi sub-clauses (*Figure 31*) is of another kind, i.e. in French the frequency of complex sentences containing several subordinate clauses at levels A and B outnumbers that of English, which is just the opposite when counting the number of mono sub-clauses or complex sentences in general. Thus, in French the multi sub-clauses constitute 36% at level A; 37% at level B and then drop considerably at level C attaining only 22%. In English the use of multi sub-clauses across all levels is more even – ranging from 27% at levels E and F to 31% at level A and 33% at level B. Overall, in both languages the test-takers give preference to more simple sentence structures containing only one subordinate clause, though the ordinal frequency of multi sub-clauses increases along with the language proficiency level.

In order to have a more profound analysis, the total number of finite subordinate clauses in complex sentences at each level of proficiency was counted and compared cross-linguistically.
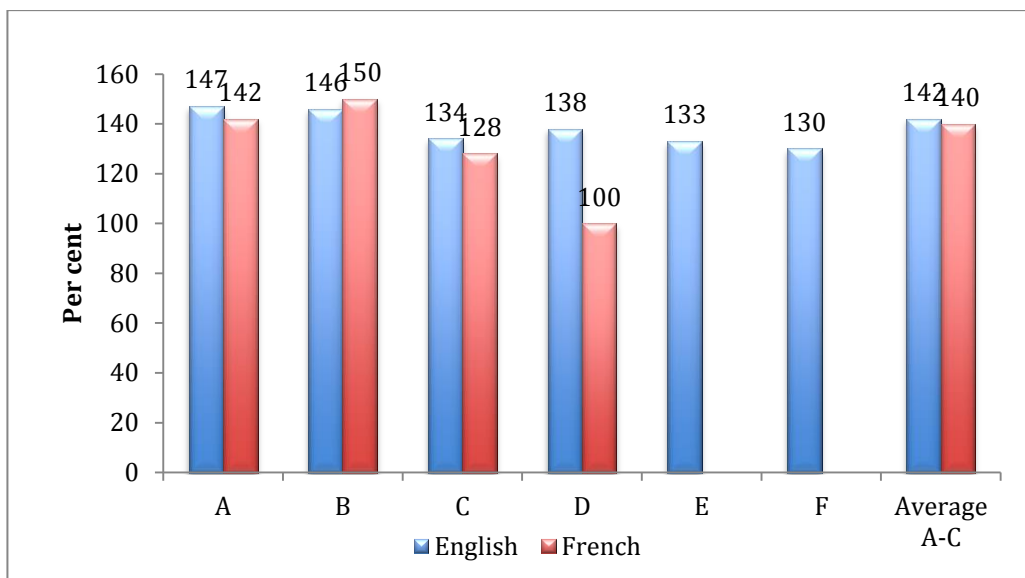
Figure 32 **Ordinal frequency of subordinate clauses**

The data reveal (*Figure 32*) that the correlation in English and French learner corpora is roughly similar. In English, the number of sub-clauses per level steadily increases from the lowest level F (130%) to the highest level A (147%); however, there is a slight drop of 4% at level C in comparison with level D. In French there is a rapid growth of 50% from level D to level B; yet, we can observe a drop of 8% at level A. The average number of sub-clauses across the higher levels of language proficiency (A - C) in both languages is almost equal (142% in English and 140% in French). Thus, the discrepancy between the number of sub-clauses used by English and French learners at higher levels of language proficiency (A – C) is not that noteworthy – only 2%.

If we look at the mean number of subordinate clauses per sentence (SUB/SEN), the results show (*Table 17*) that on average complex sentences comprise 1.37 sub-clauses per sentence in the ELWT corpus while in the FLWT corpus the mean number of sub-clauses is 1.29, which is almost equal in both languages.

Table 17 **The mean number of subordinate clauses per sentence**

Levels obtained at the examination	A	B	C	D	E	F	Average (A-F)
English	1.46	1.46	1.34	1.38	1.32	1.3	**1.37**
French	1.41	1.49	1.27	1			**1.29**

Thus, we may conclude that the learners of English and French produce almost the same number of sub-clauses. Moreover, the data substantiate the allegation that along with the

116

linguistic development there is an increase in syntactic complexity (proved also by Cosme, 2006) as the mean number of sub-clauses increases at levels A and B in both languages.

3.3.1 Complex Sentences Containing One or Mono Subordinate Clause

For further research mono subordinate clauses were classified into three groups: noun, adverbial and adjectival clauses (see Appendix 10) according to the categorical principal (see Chapter 2).
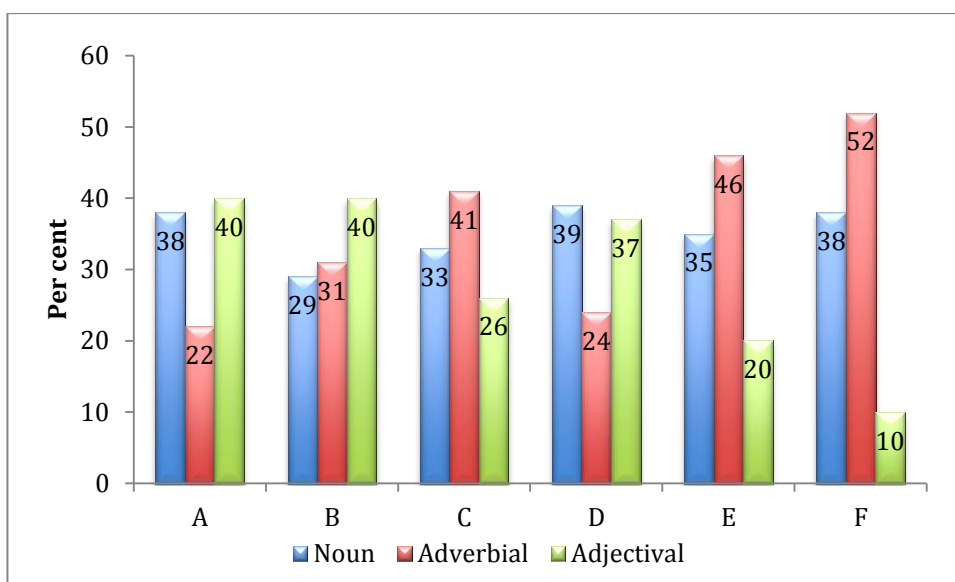
Figure 33 **Ordinal frequency of mono sub-clauses in ELWT sub-corpus**
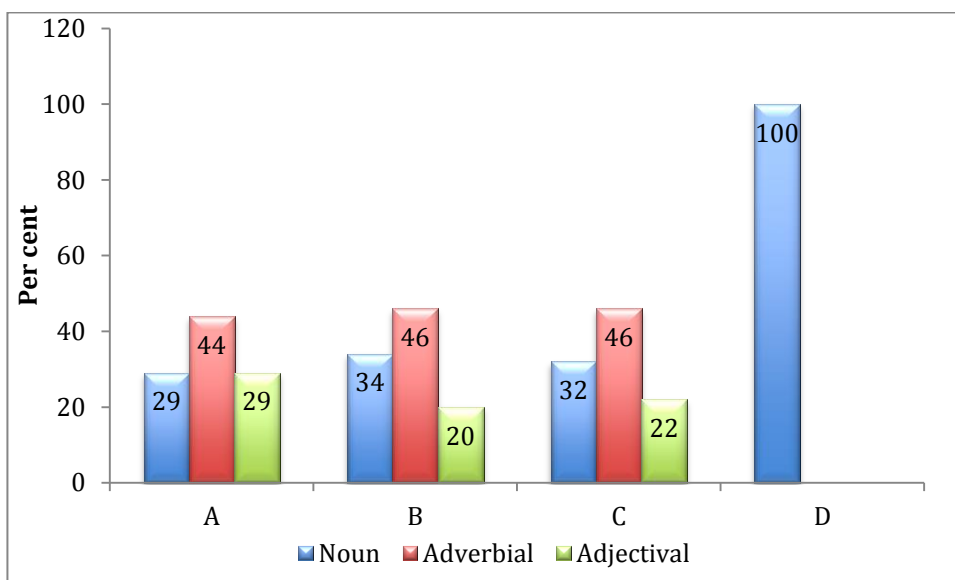
Figure 34 **Ordinal frequency of mono sub-clauses in FLWT sub-corpus**

Figures 33 and *34* reflect the distribution of different sub-clause types across language proficiency levels. As for *noun clauses*, their numbers range from 39% at level D to 29% at level B. On the whole, we can observe only slight fluctuations between the

levels, with the exception of level B where the number of noun clauses produced by the test-takers comprises 29%. In French (*Figure 34*), the raw frequency of noun clauses is rather similar to English (it varies from 29% at level A to 32% at level C and 34% at level B). At level D the column amounts to 100% because at this level only 4 complex sentences were marked, all of them containing only noun clause.

Adverbial clauses have been used most frequently both in English and French. However, in English their frequency increases towards the lowest levels of language proficiency, attaining the highest number at level F – 52%. Withal, there is a considerable fall at level D, where the column amounts only to 24%. In French, the situation is rather stable at all levels comprising on average 45%.

The distribution of *adjectival clauses* across levels is diametrically opposed to the distribution of adverbial clauses. In both languages the test-takers have used them most efficiently at higher levels of language proficiency. In English, the number of adjectival clauses constitutes 40% at levels A and B, then considerably falls at level C reaching only 26%. The numbers do not vary greatly from level C to level E. Then again we can observe a noticeable decrease at level F, where the numbers reach only 10%. In French the number of adjectival clauses per level differs from 20% to 29%.

The obtained data on the ordinal frequency of different clause types allow us to conclude that there is a certain order in acquiring syntactic structures, according to which noun and adverbial clauses come first whereas adjectival clauses appear at the highest level of linguistic complexity, which should be relied upon when discriminating between different levels of language proficiency.

In order to have a more profound analysis each clause type has been examined apart focusing on its different functions and positions in relation to the main clause in a sentence. The sentences used as examples represent different levels of language proficiency awarded by the markers at the centralized examination in English and French, thus it is obvious that they might contain mistakes. However, it does not hinder the author of the present research to determine the most distinguishing syntactic features criterial for each level of the test-taker linguistic development.

Noun clause

Noun clause is one type of subordination that has been used extensively by the students at all levels of language proficiency in both languages (*Figures 33 and 34*). We might distinguish different types of noun clauses (subordinate declarative clauses, subordinate

interrogative clauses, exclamative clauses, nominal relative clauses), which perform various functions in relations to the main clause. Therefore, the further research will be contributed to the analysis of different noun clauses retrieved from the compiled EFLWT corpus.

The statistic data demonstrate that subordinate declarative clauses have been produced most effectively in both languages at all levels of proficiency as their numbers range from 73% at level F in English to 94% at level C in French.
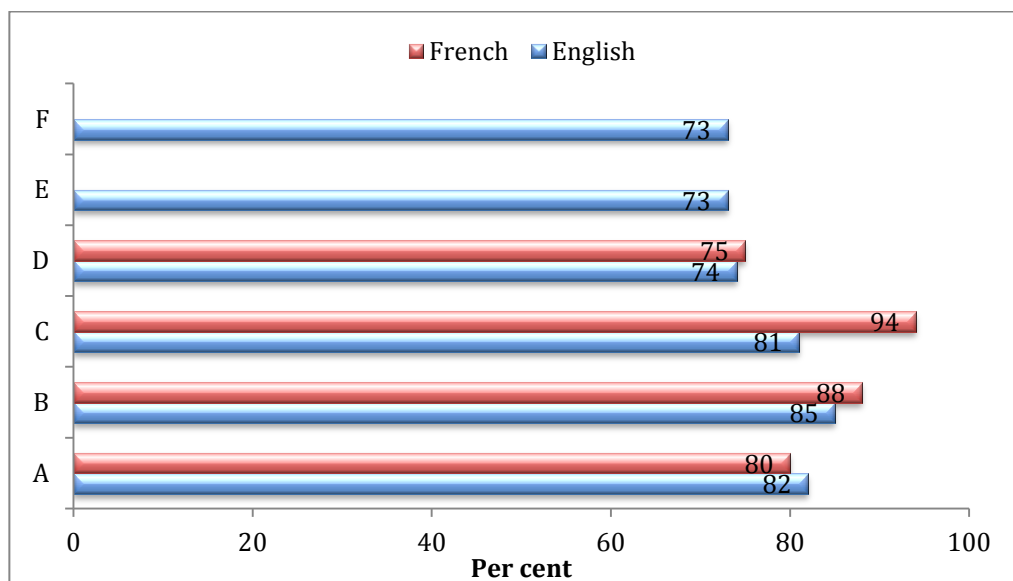
Figure 35 **Distribution of subordinate declarative clauses across levels**

If we examine various functions of *subordinate declarative clauses (Table 18)* in relations to the main clause, then we see that the majority of the examined sentences at all levels of language proficiency contain the subordinate declarative clauses introduced by the conjunction *that* in English and *que* in French which function as *direct object* in a sentence.

Table 18 **Functions of subordinate declarative clauses**

Level	A		B		C		D		E		F	
Language	*En*	*Fr*	*En*	*Fr*	*En*	*Fr*	*En*	*Fr*	*En*	*Fr*	*En*	*Fr*
Subject	3	2	6	5	3	3	1					
Direct object	17	6	25	36	30	38	35	3	17		8	
Subject complement	3		9		7	1	4		2			
Adjectival complement				3	2	2	1					

In addition, *Figure 35* reveals that this clause type has been utilised most extensively by the lower level students as their numbers gradually increase towards the lower levels of language proficiency, i.e. from 62.5% at level B to 100% at level F in English, and from 70% at level A to 86% at level C in French.
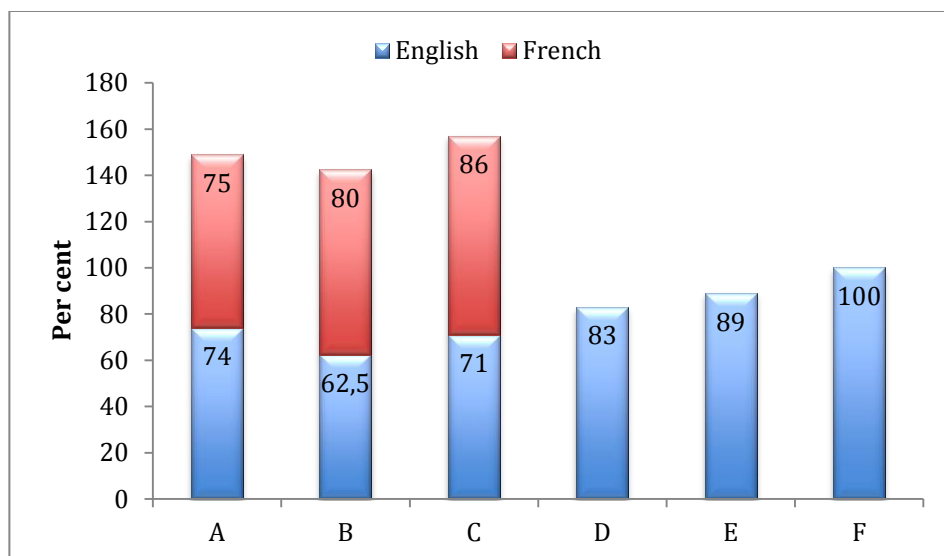
Figure 36 **Ordinal frequency of subordinate declarative clauses functioning as direct object**

Structurally, the subordinate declarative clause functioning as direct object appears at the end of the sentence, i.e. it is subordinated to the main clause:

Level A

>I feel [*that my country doesn't need me*].

>Je crois [*que l'interaction culturele fait une grande part du devéloppment mondial*].

Level B

>I do not think [*that leaving Latvia is bad*].

>Je pense [*que c'est necessaire d'apprendre de langues*].

Level C

>I think [*that this number will be smaller in future*].

>Je pense [*que tout le monde comprent ça*].

Level D

>I think [*that I can't live long time in other contry*].

>Je pense [*que on peut parler differentes langues*].

Level E

>I think [*that it is well country*].

Level F

>I think [*one problem is crisis*].

[1]
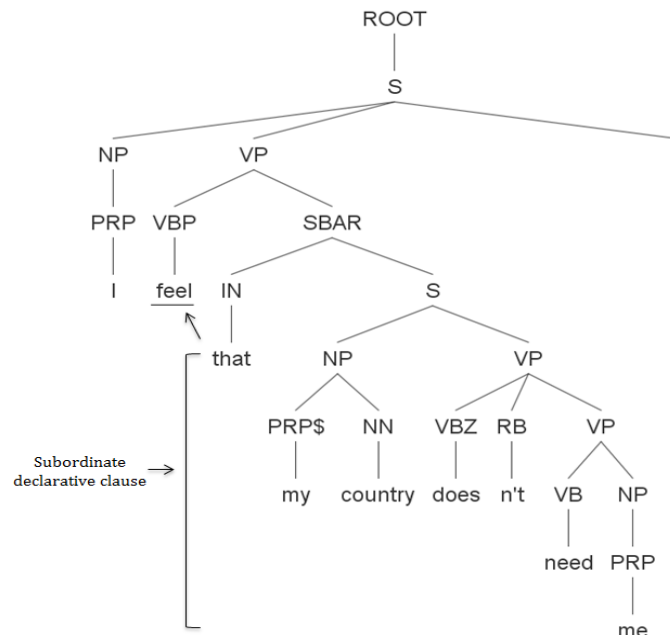
Figure 37 **Subordinate declarative clause functioning as direct object (in English)**

[2]
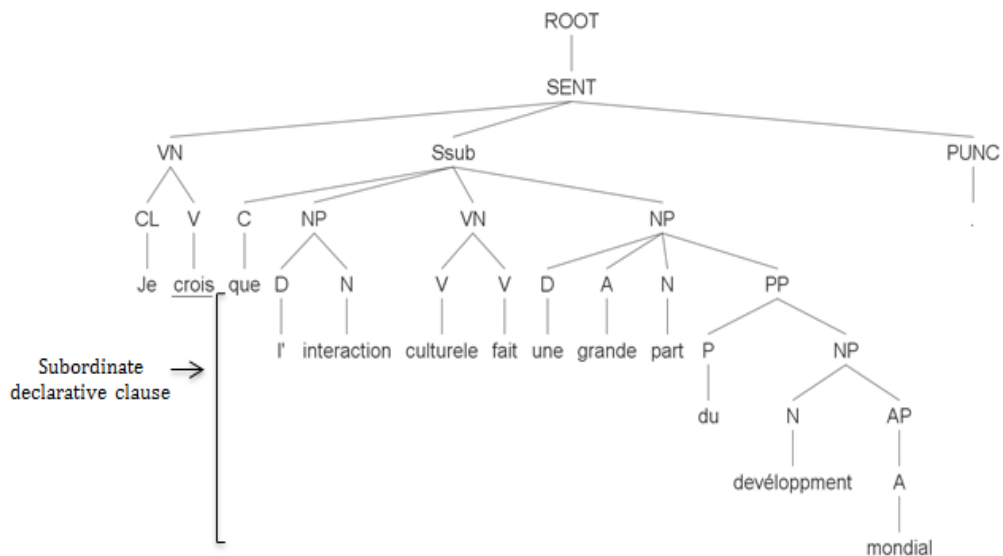
Figure 38 **Subordinate declarative clause functioning as direct object (in French)**

In the sentence represented in *Figure 37* the subordinate clause [*that my country doesn't need me*] functions as direct object and follows directly the predicate of the main clause [feel]. The same could be observed in French (*Figure 38*) where the main clause [Je crois] is followed by the subordinate clause [*que l'interaction culturele fait une grande part du devéloppment mondial*] functioning as the object of the verb [crois].

As stated in Chapter *2, that*-clause may follow directly the verbs of perception, emotion, evidence, communication and mental activity. However, the scope of verbs

chosen by the students is rather limited. The most common verbs extracted from the produced sentences are*: think, feel, know, assume, say, mean, believe, understand, show, suppose, hope, realise, tell* (in English) and *savoir, montrer, croire, penser, souhaiter, voire, considerer, dire, trouver, imaginer, douter, espérer, être content* (in French). At lower levels of language proficiency *think* (*penser)* predominates. If compared with the data from the Longman Spoken and Written English Corpus (the LSWE corpus), which contains four main registers of British and American English – conversation, fiction, newspaper language, and academic prose, the results are rather similar, i.e. the most frequently used verbs are: think, say, know, see, find, feel and believe (Biber et al., 1999: 662).

Another aspect worth mentioning is that in several sentences retrieved from the corpora the students have omitted the conjunction *that*, which is possible in English if the *that*-clause functions as direct object, e.g. *I think ^ a person should think a lot before making the final decision*. It should be specified that this occurrence is more common when the clause is short and simple. Moreover, the omission of *that* is like a tendency at lower levels of language proficiency as students avoid using the conjunction *that* at all, e.g. *I think ^ one problem is crisis. I think ^ second problem is, better life.* It might be explained by the restricted use of various syntactic structures in general.

In French the conjunction *que* cannot be omitted from the sentence, as it is the case in English. What students have to bare in mind in French is the problem of the mood as there are a number of verbs which demand the subjunctive mood after the conjunction *que*, for example, in the sentence Je doute [qu'on *puisse* être dans l'harmonie et heureux en utilisant seulement une langue] the verb *douter (to doubt)* in the main clause requires the verb in the subjunctive mood in the subordinate clause after the conjunction *que*. However, the number of sentences in which the verb has been used appropriately is rather low.

Overall, the main structural pattern for the verbs followed by the noun clause functioning as direct object is [verb + (that)/que-clause].

Another function performed by the subordinate declarative clauses is that of *subject, e.g.* It could be [*that the people there are more kinder*] *(Level B)* / Premièrement, c'est vrais [que l'anglais est actuellement la langue de communication mondiale] (level C). However, their number is rather low both in English and French (*Table 18*). The total number comprises 13 noun clauses in English (at level A – 3 clauses; at level B – 6 clauses; at level C – 3 clauses; at level D – 1 clause) and 10 in French (at level A – 2 clauses; at level B – 5 clauses; at level C – 3 clauses).

At lower levels E and F the test-takers have not used any *that*-clauses functioning as the *subject* of the noun clause. In English as well as in French there are no cases when *that*-clauses have been used at the beginning of the sentence. Nonetheless, there are sentences in which *that*-clause has been extraposed and its position marked by impersonal constructions introduced by *it* in English or *ce* and *il* in French,* thus having [subject + verb] and [subject + verb + complement] structures:

Level A

It is true [<u>that</u> a lot of people are leaving Latvia]. (type *SVC*)

[3]
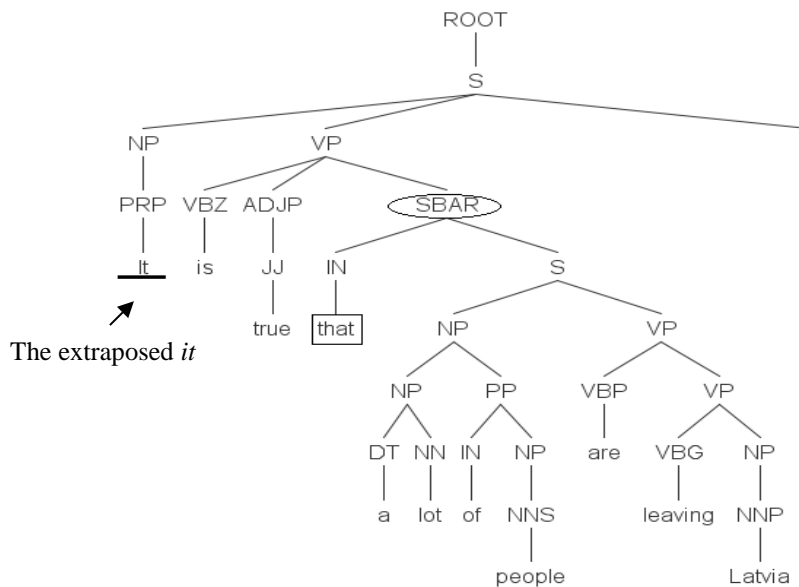
Figure 39 **Subordinate declarative clause functioning as subject (in English)**

Ce n'est pas vrai [<u>que</u> tout le monde parle l'anglais]. (type *SVC*)

[4]
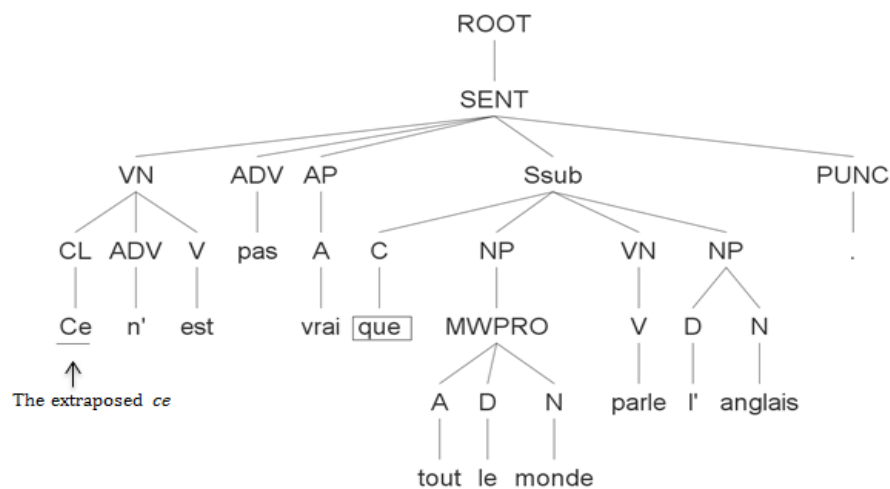
Figure 40 **Subordinate declarative clause functioning as subject (in French)**

Level B

 It is understandable [<u>that</u> people want to learn as much as possible for the work they do]. (type *SVC*)

 Deuxiémemnt, *c'est évident* [<u>qu'</u>on trouve cette langue a chaque lieu, toujours]. (type *SVC*)

 Mais parfois *il pourrait* [<u>qu'</u>il ne faut pas apprendre d'autre langue qu'anglais]. (type *SV*)

Level C

 It is not a secret [<u>that</u> abroad you will be able to get in Latvia]. (type *SVC*)

 Mais *c'est vrais* [<u>que</u> maintenent on apprend l'anglais dans toutes l'écoles]. (type *SVC*)

Level D

 The very bad is [<u>that</u> the government do nothing]. (type *SVC*)

In the example sentence of Level D the introductory construction is wrong; however, from the context it is obvious that the student has intended to use either *it is very bad* or *the worst is*. Thus, this sentence cannot serve as the evidence of the fact that students at lower levels of proficiency are familiar with such introductory constructions and are able to apply them in their written performance. Hence, the conclusion that only the students of the higher level of language proficiency use *that-clauses* as the subject of the noun clause having the structure [It + VP + *that*-clause] / [Ce/Il + VP + *que*-clause].

 There are cases when *that*-clauses have the function of the *subject complement*, although the frequency of them is not very high (*Table 18*). In English it ranges from 3 cases at level A to 9 at level B; 7 – at level C; 4 – at level D and 2 at level E. In French only one *that-clause* functioning as the subject complement has been marked at level C (Mon *avis* est <u>que</u> la connaissance de langues reflechit sur l'intelligence de gens.). In English in majority of cases *that*-clause explains the reason for doing something, since it was given in the task, e.g. The next *reason* is that it is difficult to find a job in Latvia. Nonetheless, there are other variations as well, such as:

Level B

 My *conclusion* is [that people don't leave Latvia without a good reason].

 My personal *opinion* is [that Latvia in the near future will become a bancrote].

Level D

 But *reality* is [that they do the same job for same paid].
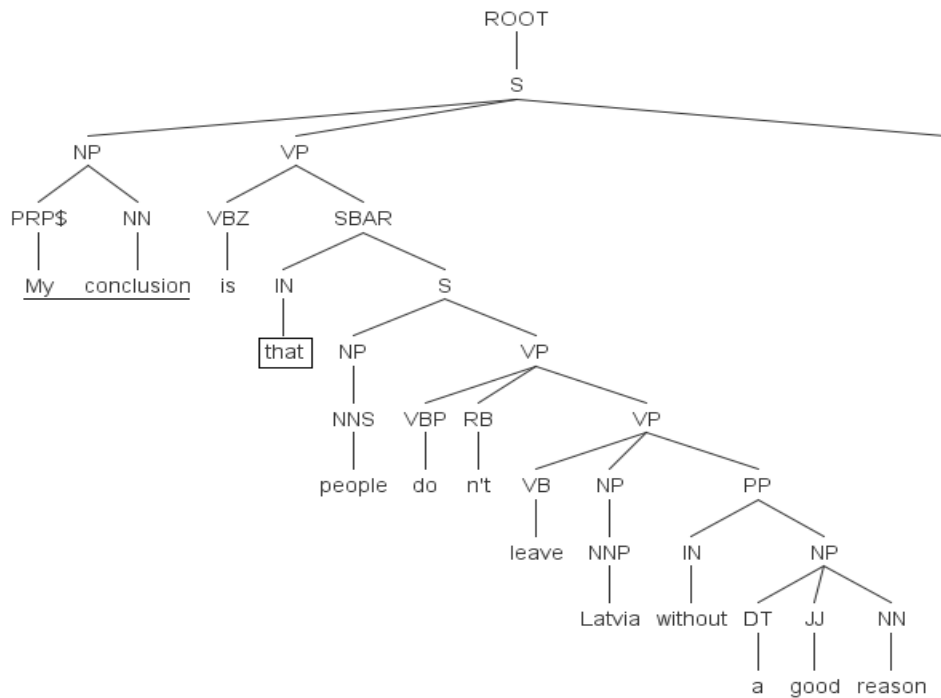
*Figure 41 **That**-clause functioning as subject complement (in English)*

In the above sentence, the NP [my conclusion] is the subject of the main clause, and *that-*clause [*that people don't leave Latvia without a good reason*] functions as the complement of the subject and appears after the linking word *is*.
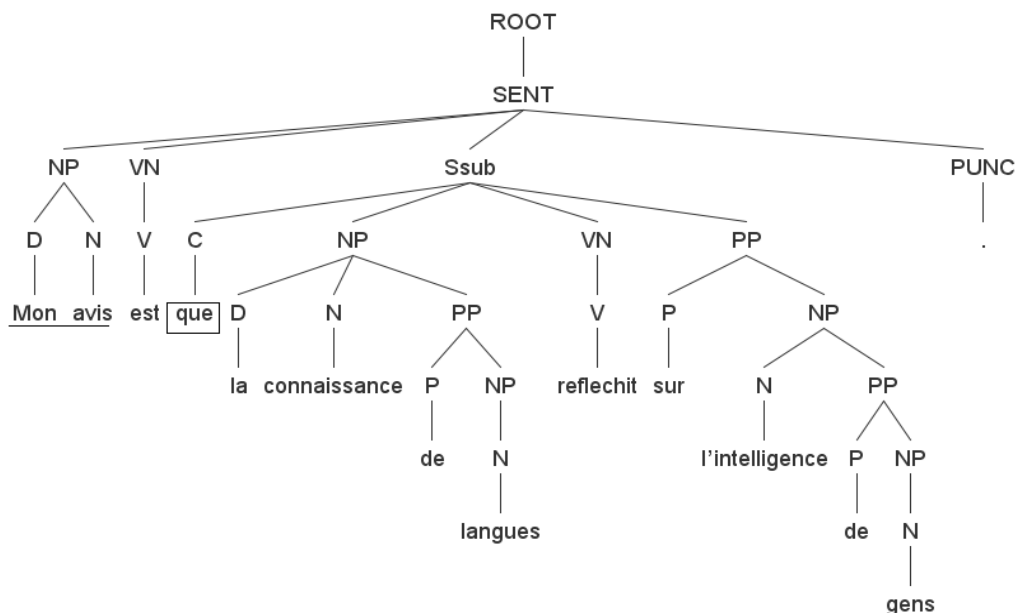
*Figure 42 **That**-clause functioning as subject (in French)*

In French, in the sentence above the NP [mon avis] is the subject and [*que la connaissance de langues reflechit sur l'intelligence de gens*] functions as the complement of the subject. Thus, syntactically *that-clauses* functioning as *the subject complements* have the same

structure in both languages, i.e. [NP + V + that/que-clause]. However, this clause type has not been used very frequently.

The cases when *that*-clause functions as *adjectival complement* are also very few: in English - 2 sentences at level C and 1 sentence at level D; in French − 3 sentences at level B and 2 sentences at level C (*Table 18*):

Level B

Je suis *sûr* [que ce serait important à apprendre des langues étrangères à l'avenir aussi].
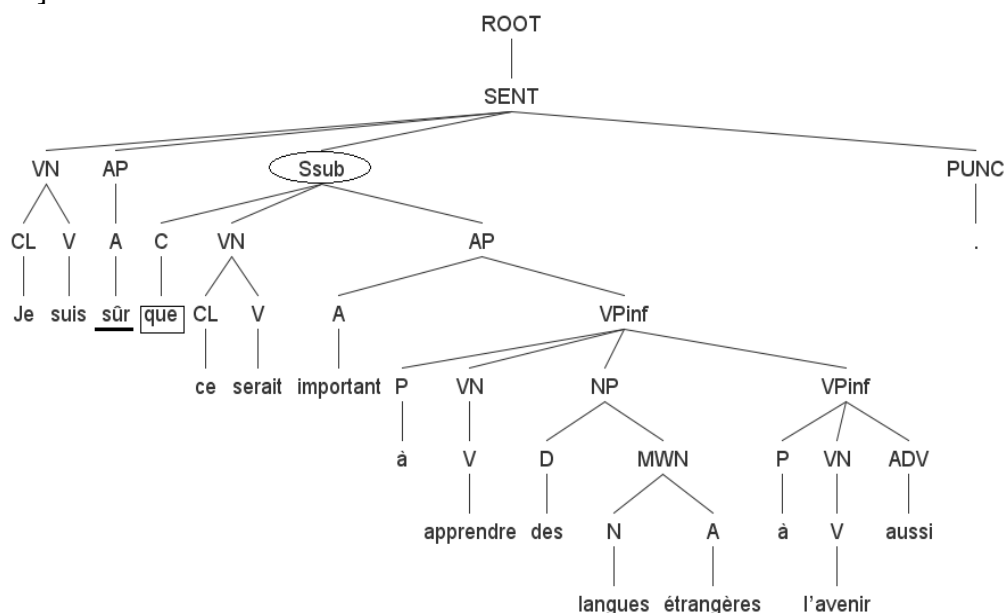
Figure 43 **That-clause functioning as adjectival complement (in French)**

From the tree diagramme in *Figure 43* it is obvious that the adjective [*sûr*] in the main clause [Je suis sûr] is followed directly by *that*-clause [*que ce serait important à apprendre des langues étrangères à l'avenir aussi*] being used as the complement of the adjective [*sûr*].

Level C

In my opinion people *proud* [that there are in abroad very huge facilities].

I am very *sure* [that I will never leave this country because of better job or something like that].

Çe ca *mieux* [que tu parles avec ils dans leur langue].

Level D

I'm *sure* [that it's very useful for our country].

126
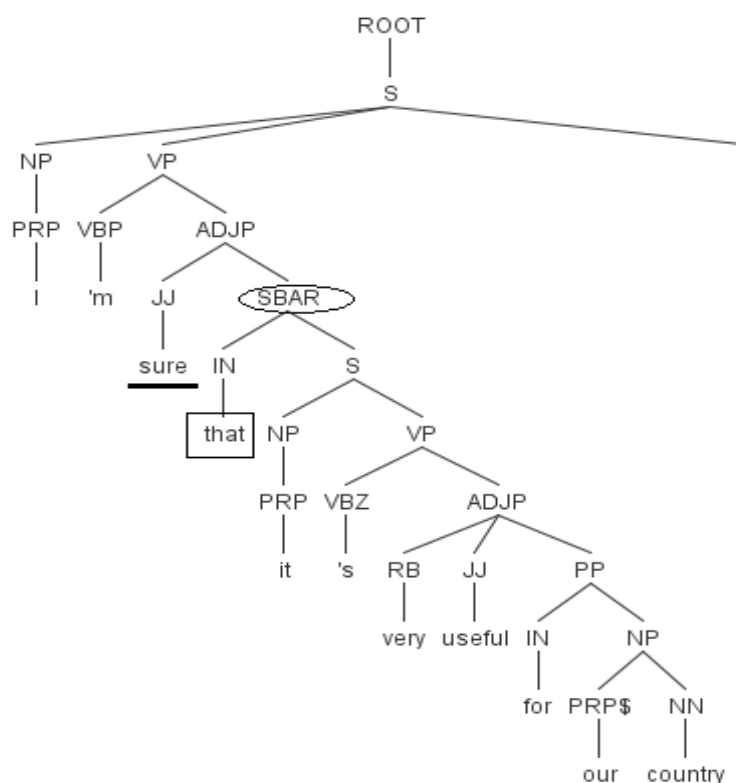
*Figure 44 **That-clause functioning as adjectival complement (in English)***

In the above sentence [*I am sure*] is the main clause in which [*sure*] is the adjective, hence [*that it's very useful for our country*] functions as the complement of the adjective [*sure*]. In majority of cases the adjectival predicate *be sure* appears before the *that*-clause. Moreover, the noun clause having the structure of [NP + VP + AdjP + that/que-clause] has not been produced at the highest level of language proficiency (i.e. at level A) in neither of the languages, which is rather surprising since it should be mastered at the early stages of linguistic development, i.e. at the Threshold level (B1).

Subordinate interrogative clauses

The subordinate interrogative clauses also have different functions in relations to the main clause. They may function as subject, direct object, subject complement, adjectival complementation and prepositional complement. However, in the compiled EFLWT corpus only some noun clauses containing the subordinate interrogative clause which functions as subject, direct object, subject complement and prepositional complement have been remarked (*Table 19*). Moreover, in English all the subordinate interrogative clauses appear at levels C and D (at level C − 10 clauses; at level D − 18 clauses), whereas in French only one alternative clause has been marked at level B.

Table 19 **Functions of subordinate interrogative clauses**

Level	A		B		C		D		E		F	
Language	*En*	*Fr*	*En*	*Fr*	*En*	*Fr*	*En*	*Fr*	*En*	*Fr*	*En*	*Fr*
Subject							1		1			
Subject complement	1											
Direct object					7		12				2	
Adjectival complementation			1									
Prepositional complement							4				1	
Yes-no and alternatives				1			1					

The subordinate interrogative clauses functioning as:

1) subject:

Level E

   [*Why people left Latvia*], that a big problem in Latvia.
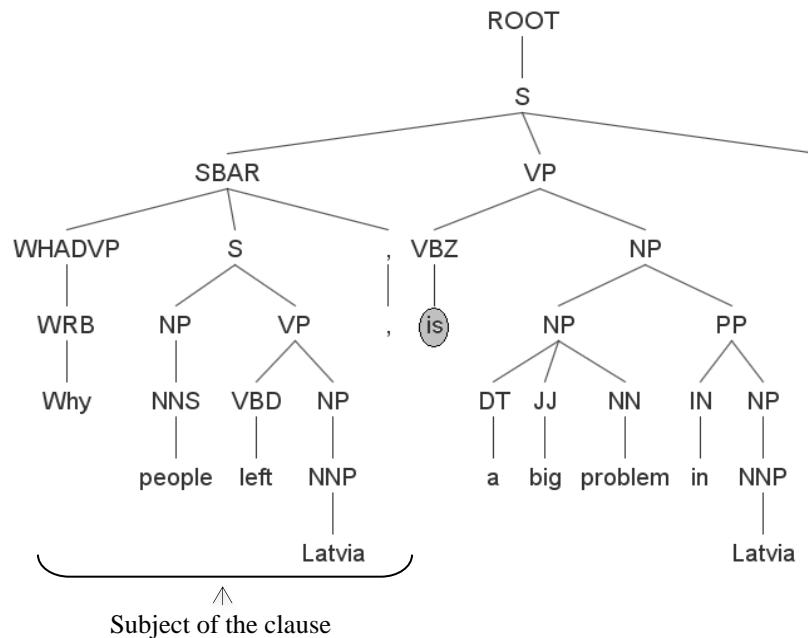
Figure 45 **Interrogative clause functioning as subject (in English)**

Although the above sentence is grammatically incorrect (instead of *that* the student should have used *is*), which has been changed in order to draw the tree diagramme, it exemplifies the subordinate interrogative clause which is introduced by the *wh*-word and followed by NP + VP, thus performing the function of a subject in the sentence.

2) subject complement:

Level A

   That is [*what drives young, powerful people into the unknown*].
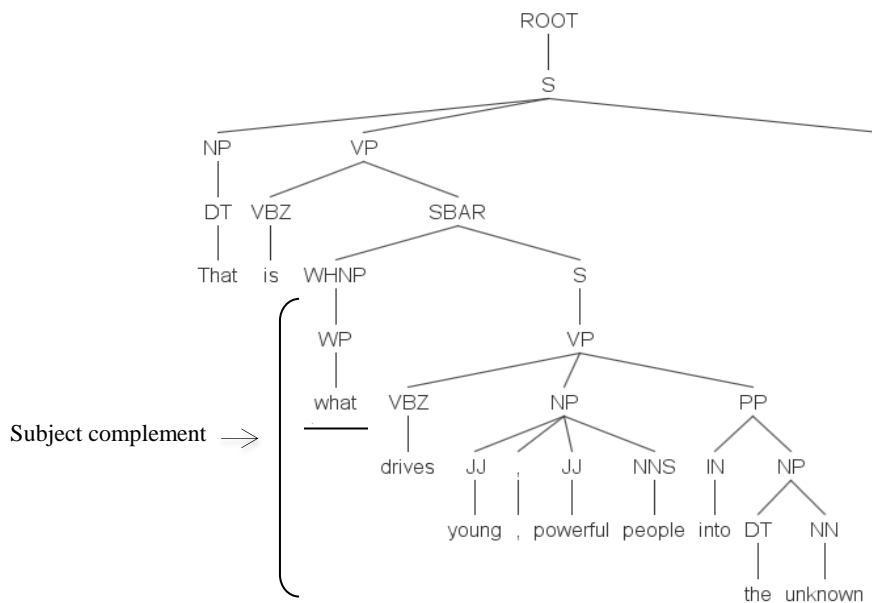
Figure 46 **Interrogative clause functioning as subject complement (in English)**

The tree diagramme of the above sentence clearly exemplifies the structure of the noun clause which contains the subordinate interrogative clause [what drives young, powerful people into the unknown]. The *wh*-interrogative clause functions as the complement of the subject [*that*] followed by the verb [*be*]. In this case the main clause contains only NP consisting of the determiner and VP, which could be represented as [NP + VP + *wh*-clause]. However, this clause type has been used only once by the student of English having the highest level of language proficiency.

3) direct object:

Level C

> Just remember [*how hard our parents, grand parents and so on tried for free Latvia*].

Level D

> Than you can seen [*how Latvia be a good, rich and beautiful place*].

Level F

> He dont now [*what them do*].
>
> I can't explain [*why people have left Latvia during the last few years*].

129
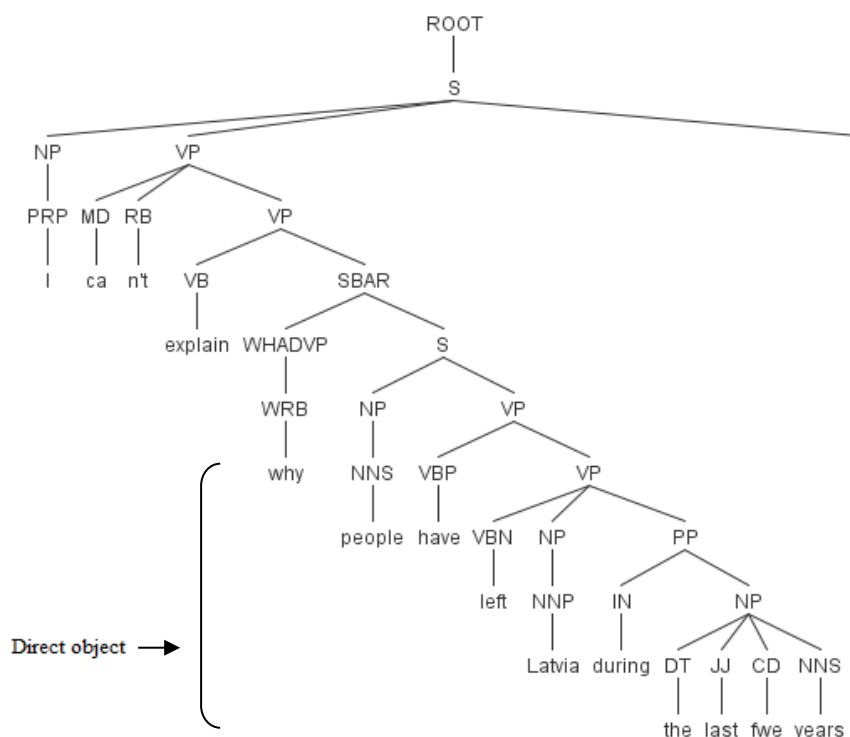
Figure 47 **Interrogative clause functioning as direct object (in English)**

From the above examples it is obvious that the interrogative subordinate clause [why people have left Latvia during the last fwe years] functions as direct object of the main clause [I can't explain] and follows directly the predicate of the main clause. Moreover, this clause type has been marked extensively at levels C and D in English, whereas no interrogative subordinate clause functioning as direct object has been used by the learners of French.

4) adjectival complementation:

Level B

It's very personal [*what are the reasons for leaving*].

There has been only one case when a student has used the interrogative subordinate clause that functions as the complement of the adjective in the main clause.

5) prepositional complement:

Level D

At the bigining they were a fraid – [*what will say others*]?

People are scared [*how they will live without a job*].

We all have to think [*how to improve this situation*].

Economic crazy in Latvia are guilty [*why Latvians people leave Latvia*].

Level F

The my opinion <u>about</u> [*why people have left Latvia during*] is negativ.
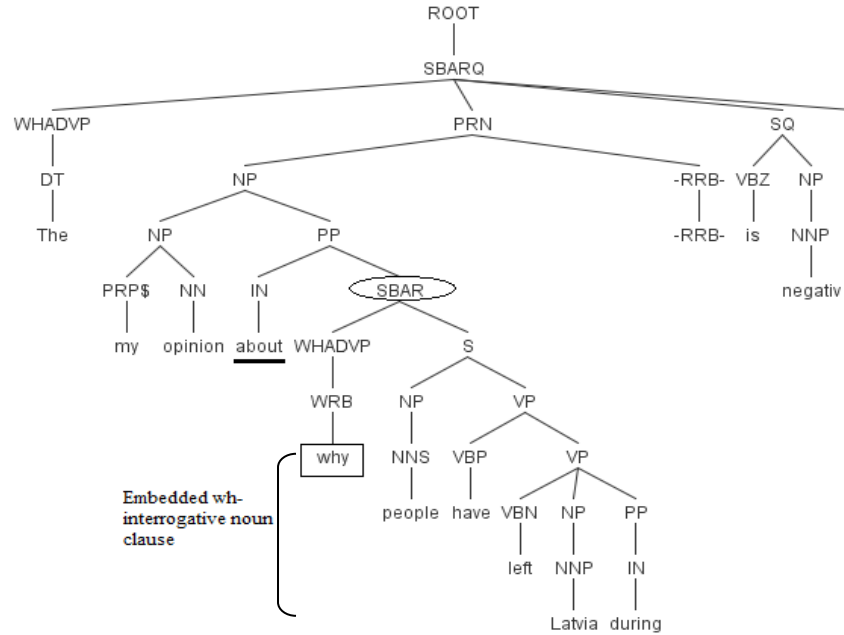
Figure 48 **Interrogative clause functioning as prepositional complement (in English)**

In the sample sentence (*Figure 49*) the interrogrative noun clause [why people have left Latvia during] is embedded in the main clause [The my opinion about is negativ].

All the sentences in which the subordinate interrogative clause functions as the complement of the preposition of the main clause have been noticed in English at the lower levels of language proficiency; hence, they contain a lot of grammar mistakes. If we are to examine this particular clause type, then we see that either prepositions are missing or wrong prepositions have been used, e.g. in the sentences at level D there is no preposition *of* after the words *afraid*, *scared, to think*, *guilty.* As for syntactic structures, in majority of cases the noun clause is subordinated to the main clause, i.e. [NP + VP + PP + *wh*-clause]; only in the sentence at level F the *wh*-interrogative noun clause is embedded in the main clause having the structure [NP + PP + *wh*-clause + V + NP].

Yes-no **and alternative interrogative clauses**

Yes-no and alternative interrogative clauses may perform all the aforementioned functions of the subordinate interrogative clauses; however, there has been only one case in English and one case in French when the students have used this subordinate clause type in their written performance (e.g. I dont know [*if I can do that....*]; Dans nos jours, nous pouvons seulement décider [*si l'apprentissage des langues est importante pour nous ou non*]). In

both examples the subordinate clause is introduced by the correlative *if* or *si,* and the subordinate clause functions as the direct object of the main clause. Therefore, two examples are not enough to draw any further conclusions.

Nominal relative clauses

The same could be referred to when analysing *nominal relative clauses* as their number is rather low. There is only one nominal relative clause at levels A, B and D in English, and 4 clauses at level B; 1 clause at levels C and D in French:

Level A

So I will start with the statement [*that other countries pay better*].

Level B

Generally speaking, I agree with statement [*that nowdays it is necesarry to leave Latvia*].

Dans ce cas ça me donne des doutes [*que l'anglais soit la langue mondiale*].

C'est quelque chose [*que personne peut te voler*].

La langue, c'est quelque chose [*qu'on aprend pour nous, pour participer dans la vie pour communiquer*].

La langue est l'une des identifications [*que nous sommes differents*].

Level C

Avec la langue anglais c'est dificile pour explique tout [*que nous voulons*].

Level D

From recent times people in Latvia lost a faith [*hat they will find job*].

Je suis d'accord avec opinion, [*que l'anglais est actuallement la langue de communication mondiale*].

In all the examples the nominal relative clause follows the noun phrase of the main clause and it is introduced by the pronoun *that* in English and *que* in French. When analysing this particular subordinate clause type, one has to be very considerate not to regard it as a relative clause.

Having examined different types of noun clauses as well as their function in relations to the main clause, the author of the present research concludes that noun clauses have been utilised by students at all levels of language proficiency in both languages, which was proved also by the quantitative data analysis. However, there is a difference in

the frequency of use of certain types of noun clauses. The dominant position is taken by the subordinate declarative clauses functioning as direct object. As for other functions, only the students at the higher level of language proficiency use *that-clauses* as *subject* of the noun clause; *that-clauses* functioning as *subject complements* are not being used very frequently in comparison with the one functioning as *direct object*. The subordinate declarative clauses performing the function of adjectival complement do not appear at the highest levels of language proficiency at all.

The subordinate interrogative noun clauses have been widely used in English, particularly at levels C and D, while the French learners entirely exclude this syntactic structure. The few *yes-no* and *alternative interrogative* clauses as well as *nominal relative* clauses signal that the learners are familiar with these syntactic structures; however, their usage is rather low.

If we compare the obtained results with the findings of Biber et al. (1999), who analysed syntactic structures in the Longman Spoken and Written English (LSWE) corpus, then we see that the most common clause to report people's emotional state or personal thoughts is by using *that*-clause introduced by the verbs (e.g. think, know, feel, hope, etc.), among which *think* is the most common in conversation, *say* occupies the second position in news and conversation, whereas mental verbs (e.g. mean, suppose, realize, etc.) are frequent in fiction. This might explain the extensive use of *that*-clause having the function of direct object because the task requires expressing one's point of view on the given themes in both languages. What is more, the omission of *that* is the norm in conversation while it has to be retained in academic prose (Biber et al., 1999: 680). This could justify the omission of *that* at lower levels of language proficiency. As for the *that*-clause having the function of adjectival complement, it is much less common (except for *be sure*). The other clause types appear rather rarely, e.g. the extraposed *that*-clauses and the *that*-clause having the function of subject complement are 'moderately common in news and academic prose (but rare in fiction and conversation)' (ibid.: 674); interrogative clauses following the verbs – *see, tell NP* and *wonder* - are more common in conversation, where it reports that the speaker does not know something, and are relatively rare in news and academic prose (ibid.: 688).

Hence, we might conclude that the most often used structural patterns are [verb + *that*-clause] and [verb + NP + *that*-clause], and that the frequency of the structural patterns produced by the test-takers in English and French correspond the ones used by the native speakers determined by the discourse type.

Adverbial clause

Adverbial clause is another clause type being examined within the scope of the present research. The quantitative data (*Figure 50*) demonstrate the distribution of adverbial clauses across levels in both languages. In English the number of adverbial clauses increases towards the lowest levels of language proficiency, except for level D, where the numbers plummet from 29% to 16.5%, which is almost identical to the initial position of level A (16%). Otherwise, the number of adverbial clauses gradually rises from 22% at level B to 29% at level C reaching the highest numbers at level E (32%) and level F (37.5%). In French, the distribution of adverbial clauses is more regular. The number of adverbial clauses at the higher levels of language proficiency is almost the same (at level A – 28%; at level B – 29%). The ordinal frequency of this clause type increases at level C attaining 36%. Even so, at level D no adverbial clauses have been used by the test-takers in French.
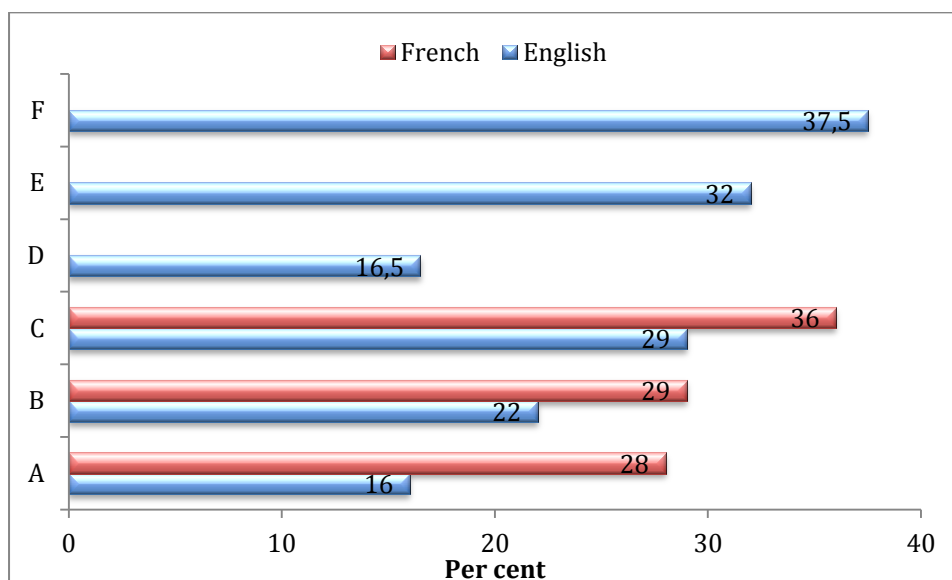
Figure 49 **Distribution of adverbial clauses across levels**

Overall, it is evident that this clause type might be criterial at the lower levels of linguistic knowledge both in English and French.

For a more detailed analysis the adverbial clauses marked in the student written essays were sub-divided into several groups according to their function as modifiers of the main clause. Thus, the adverbial clauses specify circumstances such as time, place, manner, comparison, condition, concession, etc. (*Table 20*).

Table 20 **Types of subordinate adverbial clauses**

Level	A		B		C		D		E		F	
Language	*En*	*Fr*	*En*	*Fr*	*En*	*Fr*	*En*	*Fr*	*En*	*Fr*	*En*	*Fr*
Time & place	4	1	8	8	12	7	5		5		1	
Manner & comparison					1							
Circumstantial												
Condition	2	12	15	43	26	37	9		7			
Concession	1		2									
Contrast			1									
Reason, cause	9	1	22	14	24	15	21		21		14	
Purpose			2	1								
Comment			1	1								

Figure 51 reveals the test-taker most frequently used clause types. According to the obtained data, the adverbial clause of reason and cause occupies the first place among the learners of English, comprising 52% of all clause types. The preference for this clause type could be explained by the fact that the assigned written task envisages giving reasons for leaving Latvia. The second most common clause type, which reaches 28%, is the one determining condition. The clause of time and place occupies the third position having the frequency of 16%. The learners of French have produced mainly the same clause types as the learners of English, though their proportions are different. The adverbial clause of condition is the most common clause type in the FLWT sub-corpus as the figures rise to 64%; the second most frequently used type is the adverbial clause of reason and cause (21%), while the adverbial clause denoting time and place reaches 12%, thus being the third most often used clause type. The number of other adverbial clauses ranges from 0.5% to 1.5%, which is significantly lower compared to the three most common clause types. Moreover, no adverbial clauses of contrast, concession and comparison have been marked in the FLWT sub-corpus.
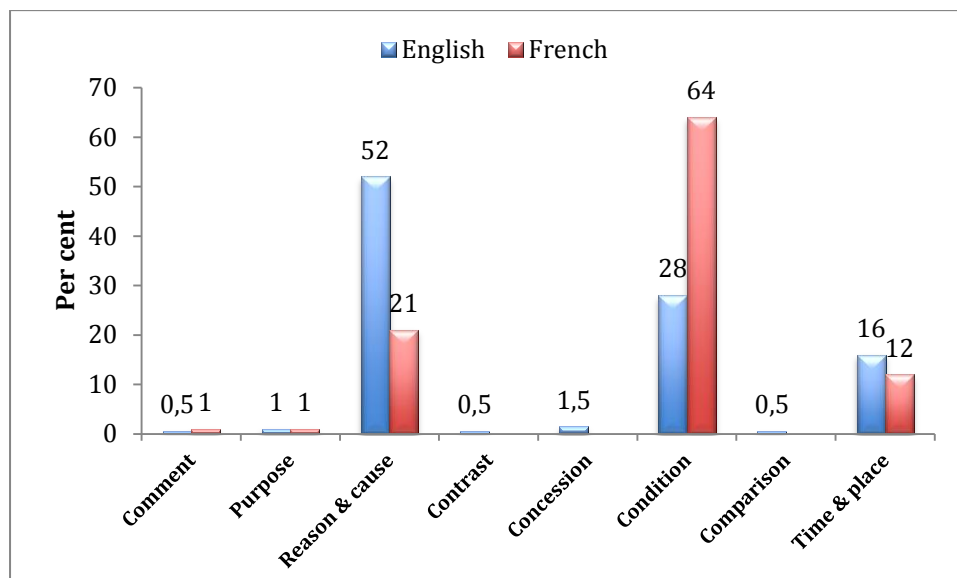
Figure 50 **Frequency of different types of adverbial clauses**

The obtained results are largely in line with the ones retrieved from the LSWE corpus, i.e. the use of adverbial clauses is similar to 'the relative distribution of circumstance adverbials generally' (Biber et al., 1999: 820), which depends on the register. The LSWE corpus findings show that the most frequent categories of adverbials are: place, time, process, and contingency (cause/reason, purpose and condition), though their preference across registers differs. For example, conditional clauses are most common in conversation, and moderately common in academic prose; reason/cause clauses are common only in conversation, and time clauses are very common in fiction and news (ibid.: 820). Thus, the data in *Figure 51* substantiate the LSWE corpus findings as the clauses of reason/cause, condition and time/place are among the most frequently produced adverbial clause types also by the test-takers of English and French.

For a more profound analysis the focus will be attributed to the three most frequently used clause types since they can be compared across languages.

Adverbial clauses of time and place

The quantitative analysis of adverbial clauses shows that the adverbial clause of time and place occupies the third place among the most frequently produced clauses by the students both in English and French. Nonetheless, *Figure 52* provides a clearer picture of the distribution of this clause type across proficiency levels. It is obvious that the distribution curve in English is uneven, i.e. it varies from 25% at level A to 7% at level F. In French there is not such a noticeable difference among levels as the number of clauses ranges from 7% at level A to 12% at levels B and C.

136
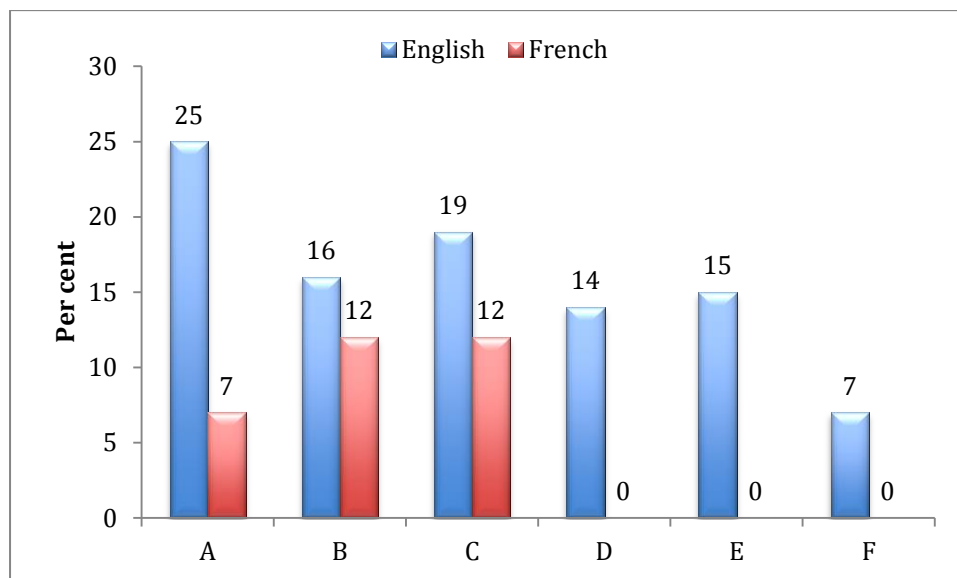
Figure 51 **Adverbial clauses of time and place**

Adverbial clauses of time (as stated in Chapter 2) may be divided into three groups according to the time relations between the main clause and the subordinate clause. Thus, we distinguish the adverbial clauses of *time before*, *same time*, and *time after*. The subordinate conjunctions are the ones that denote time relations in a sentence.

In English and French the students have widely utilized the subordinators *when, while* and *quand* to express *simultaneity:*

Level A

[*When they go abroad*]*,* they see a lot of places, hear different languages and see different cultures.

Level B

We need to be smarter now [*when we all are down*].

In Latvia [*while is a crisis*] lot of people loose their jobs.

[*When people graduate from university*] they are looking for a job they could have.

Et on peut les apprendre uniquement [*quand on a appris la langue de tel ou tel nationalité*].

Level C

[*Quand tu etudies la langue*] tu aussi etudies au culture de ce pay.

However, there are cases when the situation is dubious, i.e. the subordinator *when / quand* might be replaced by *if / si* or a complex preposition *on condition that / à condition que*, thus obtaining the meaning of a conditional clause:

Level B

Most of people are happy [*when you are happy*]. (La plupart de gens sont heureux [*quand vous êtes heureux*].) ↓

Most people are happy [*if (on condition that) you are happy*]. (La plupart de gens sont heureux [*si/ à condition que vous êtes heureux*].)

Few students have made use of the subordinator *till* and *until* to *express anteriority*

of the situation in the subordinate clause in relations to the situation in the main clause:

Level B

It probably will be [*until Latvia's economic system is going to be fixed*].

Level C

People keep leaving Latvia for some more years, [*till situation get fixed*].

Both clauses may express not only time, but also result.

Subordinators *since* and *after* in English have been used only once to express *anteriority:*

Level C

For 19 years, [*since Latvia has got an independence*], country has a social 'war' inside.

Level F

The Coast of Arms of Latvia was adapted four days [*after it gained city rights on 18 March*].

The marked sentence at level F cannot be considered as an example of the appropriate use

of the subordinator *after* because it is obvious that the sentence has been copied from some

source and does not comply with the task.

In French two sentences where students have used time expressions *'il y a dix ans que'* and '*il y a quelques mois que'* instead of the subordinator *depuis* have been marked.

Both sentences appear at level B:

[*Il y a dix ans que*] le monde est entré dans un nouveau millenaire.

[*Il y a quelques mois qu'un*] institut de langues étrangères à Paris a fini la prèmiere dictionnaire de mots intraductiables.
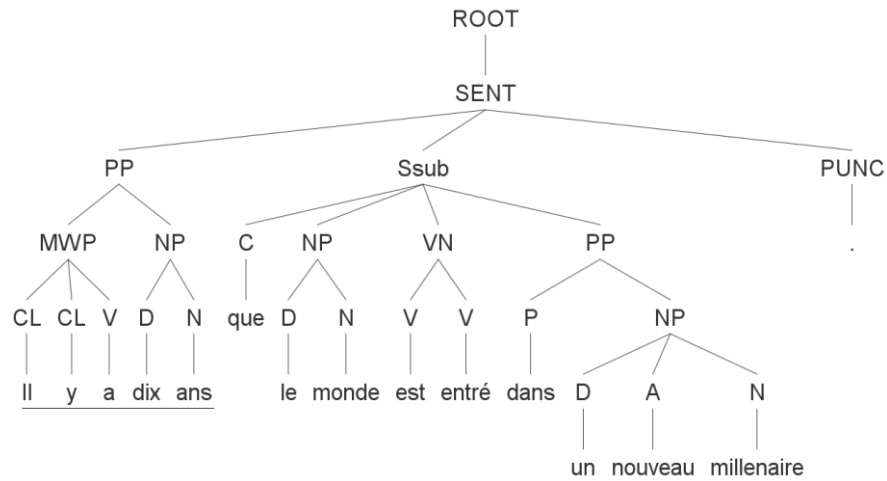
Figure 52 **Adverbial clause of time in French**

From the tree diagram we can see that the time expression *'il y a dix ans'* operates as the main clause of the sentence although it is a prepositional phrase; hence, there is neither subject nor predicate. Moreover, the subordinate clause is introduced by the conjunction *que*, which is typical for noun and relative clauses.

If we are to analyse the syntactic structures of all the adverbial clauses of time, then we will have the examples of all three positions in a sentence, i.e. the adverbial clauses of time appear before [1], in the middle [2] or after the main clause [3].

[1]
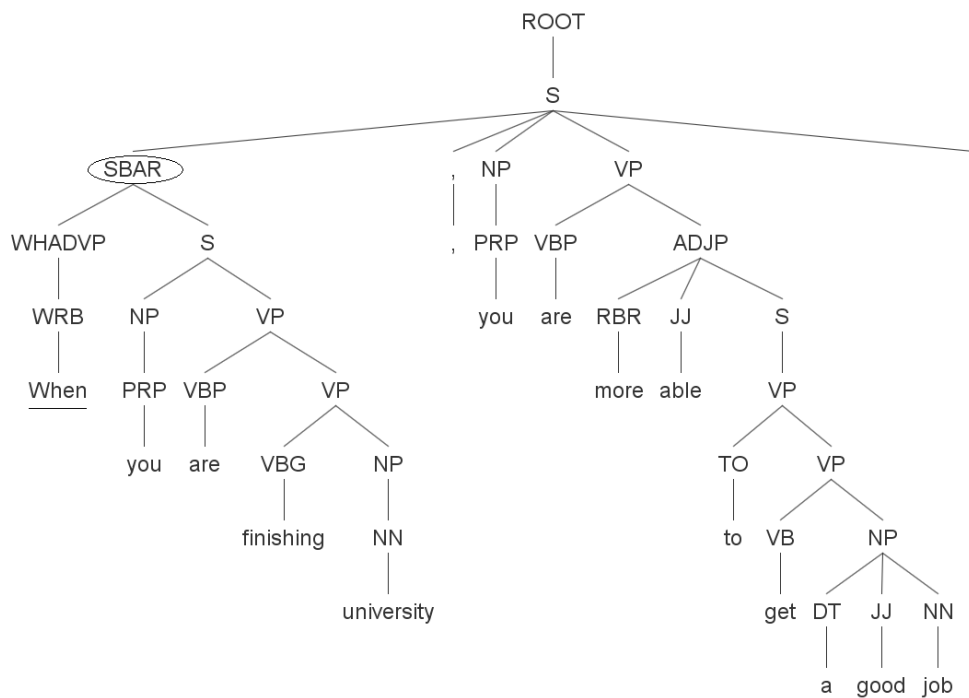
Figure 53 **Initial position of the adverbial clause of time in English**

[*When you are finishing the university*,] is the subordinate adverbial clause of time, which occupies the initial position in relation to the main clause [you are more able to get a good job].

[2]
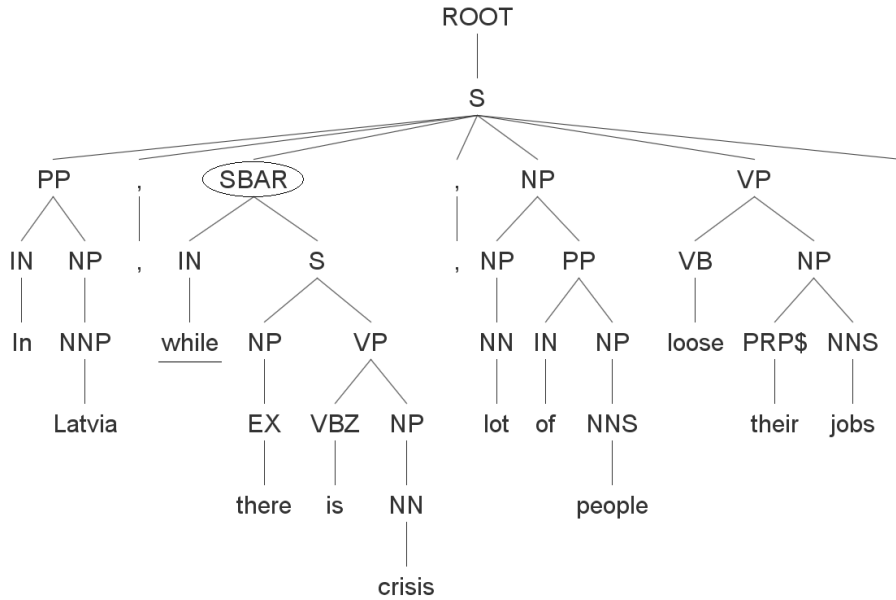
Figure 54 **Medial position of the adverbial clause of time in English**

The subordinate adverbial clause of time [*while is crises*] is embedded in the main clause [In Latvia lot of people loose their jobs], thus having the medial position in a sentence (when parsing this sentence, it was necessary to insert *there* so that the hierarchical representation of the subordinate clause is appropriate).

[3]
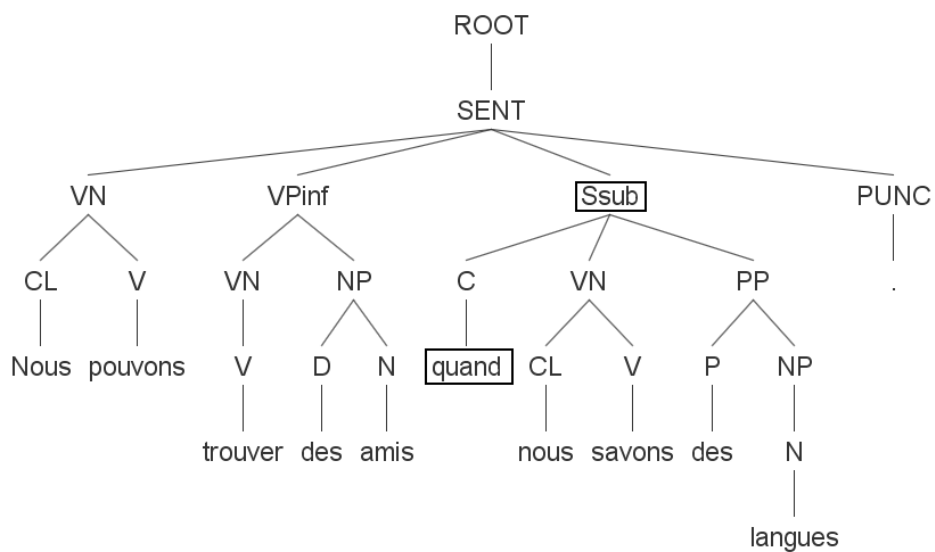
Figure 55 **Final position of the adverbial clause of time in French**

In French the sentence [Nous pouvons trouver des amis] is the main clause that is followed by the subordinate adverbial clause of time [*quand nous savons des langues*].

As regards the *adverbial clauses of place*, only one sentence in English and two sentences in French have been marked:

Level A

And why not to go abroad, [*where people with such education are needed*].

Level B

Tous [*où vous venez*] quelqu'un connait cette langue.

Level C

Avec sa education de langues personne peut aller [*où il elle voudra*].

These three examples are not enough to draw any conclusions whether students are familiar with this structure and whether they can use it appropriately because in English at level A, which is the highest level of language proficiency, the produced sentence is grammatically incorrect.

Having examined the adverbial clauses of time and place, it could be concluded that foreign language learners express: 1) simultaneity mainly by restricting themselves to the subordinators *when, while, quand* and 2) duration by using the subordinators *until, while,* and *depuis.* Moreover, the students have utilized very few sentences with subordinators denoting time after and before. As for the adverbial clause of place, more data are needed to draw any grounded conclusions.

The adverbial clause of time and place may appear either in the initial or final position within a sentence; however, there are some cases when it has been placed also in the medial position, i.e. being embedded in the main clause. However, such cases are very few, which has been proved also by analysing the LSWE corpus data, i.e. the occurrences of medial adverbial clauses were too rare to carry out meaningful quantitative analysis (Biber et al., 1999: 831).

Adverbial clauses of condition

Adverbial clauses of condition are the most widely used clause type among the test-takers of French (they constitute 64% of all the adverbial clauses), while in English the adverbial clauses of condition occupy the second place among the adverbial clauses attaining 28%.

Figure 57 reveals the distribution of conditional clauses across the levels of language proficiency. In English, only two sentences or 12.5% containing conditional clauses have been found at level A, which is surprisingly few. The number of conditional clauses increases at levels B and C, reaching 41% at level C, and then falls down to 26% at level D and then again rises to 47% at level E. At the lowest level F no conditional clauses have been used. In French, the students have used this clause type most effectively, which could be affirmed by the statistical data shown in *Figure 57*. Thus, the number of conditional clauses constitutes 75% at level A, which slightly descends remaining at 63% at level B and 62% at level C.
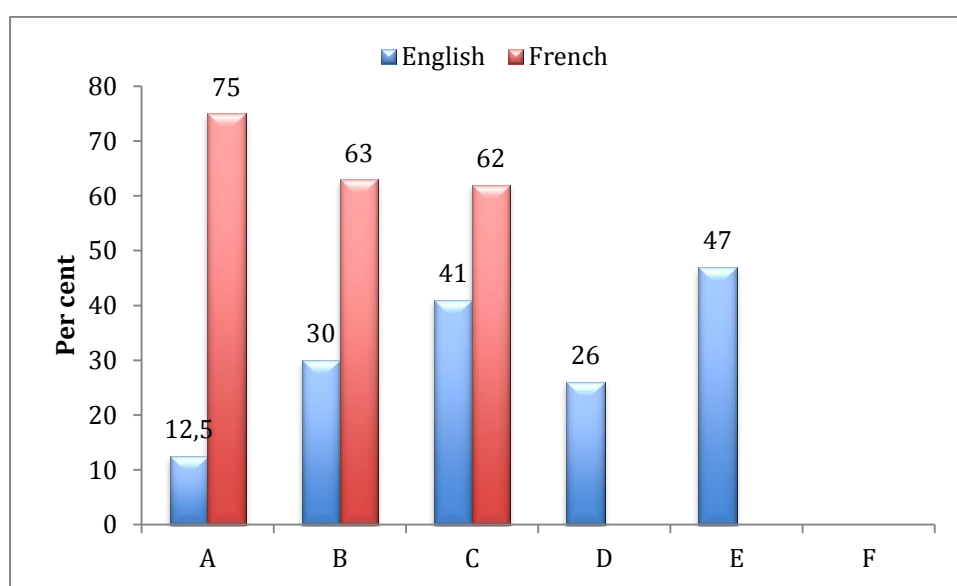
Figure 56 **Adverbial clauses of condition**

The adverbial clauses of condition (as stated in Chapter 2) can be subdivided into *direct* and *indirect c*onditional clauses. The majority of conditional clauses marked in the English and French learner corpora belong to the first group, i.e. direct conditional clauses, which perform the function of adjuncts:

Level B

> [*If I decide to leave*], my main reasons would be better paid job, better education possibilitys and chance to see the world.

> [*Si on veux faire certains travails*], les conaissances des plusieures langues sont obligatoires.

Furthermore, the adverbial clauses of *direct* condition are divided into *open* and *hypothetical* conditions, which is of grammatical importance as verb forms in the main and subordinate clause may be used in different tenses depending on the type of condition.

<p style="text-align:center;">*Table 21* **Different structures of conditional clauses**</p>

Level	Conditional clauses in English	Conditional clauses in French
	if + present / present *present / if + present*	*si + présent / présent*
A		A mon avis, *si* on <u>apprend</u> une langue, on <u>étudie</u> aussi une nouvelle culture.
B	*If* people <u>loose</u> their jobs they <u>do not have</u> money to pay for their homes.	*Si* la personne <u>parle</u> en plusieurs langues, elle <u>est</u> plus intelligente.
C	*If* you <u>have</u> much money and good knowledge in language you <u>can</u> go and study abroad.	*Si* le personne <u>sait</u> beaucoup de langues, il <u>a</u> beaucoup de possibilités.
D	*If* you <u>have</u> money, you <u>can</u> by this aducation and marks.	
E	Nowaday we <u>must</u> work hard *if* we <u>want</u> to live in beautiful and rich country.	
	if + present / future	*si + présent / futur*
B	If he get there new friends, nice job and have no feelings, no sentiment and no wish to come back, he will stay there.	*Si* on <u>sait</u> beaucoup de langues, on <u>pourra</u> trouver beaucoup d'information de l'autres pays et cultures.
C	If they want to be together, someone will have to choose to stay in a motherland or to go abroad.	*Si* tu <u>connais</u> plusieurs langues <u>va aider</u> au beaucoup de situation.
D	*If* they <u>countinuing</u> learning in other contry, they <u>will apper</u> opportunities to get better job.	
		conditionnel présent / si + imparfait
B		Ça <u>serait</u> une catastrophe *si* tout le monde <u>commançaient</u> parler anglais.
C		Premierement, *si* tout les personnes <u>parlaient</u> l'anglais toute le mond <u>serait</u> comme un grand Etats Unis.

Overall, the data in *Table 21* reveal that the test-takers in English are familiar with the *open* conditional sentences at all levels of language proficiency, except for level F (there is no example given for level A in English as the marked clause contains wrong tense forms). At levels B and E the construction <u>if + present / present</u> predominates; at levels C and D the students have used the constructions <u>if + present / present</u> and <u>if + present / future</u>. In French the majority of sentences at all levels of language proficiency contain the construction <u>si + présent / présent,</u> while the constructions <u>si + présent / futur</u> and <u>conditionnel présent / si + imparfait</u> appear at levels B and C. However, their number is

rather low (only two or three clauses have been marked). What is more, neither in English nor in French the students have produced the constructions with the past reference for the hypothetical condition.

As for the subordinators used by the students, it should be noted that their scope is rather restricted. Almost all conditional clauses are introduced by the subordinators *if* (in English) and *si* (in French). The subordinators *even if* and *as long as* have been used only once at level A, while the subordinator *même si* appears four times.

If we look at the hierarchical layout of the sentences, then we can have two variations. In the first case, the subordinate clause is placed at the beginning of the sentence, preceding the main clause (*Figures 58* and *59*):

[1]
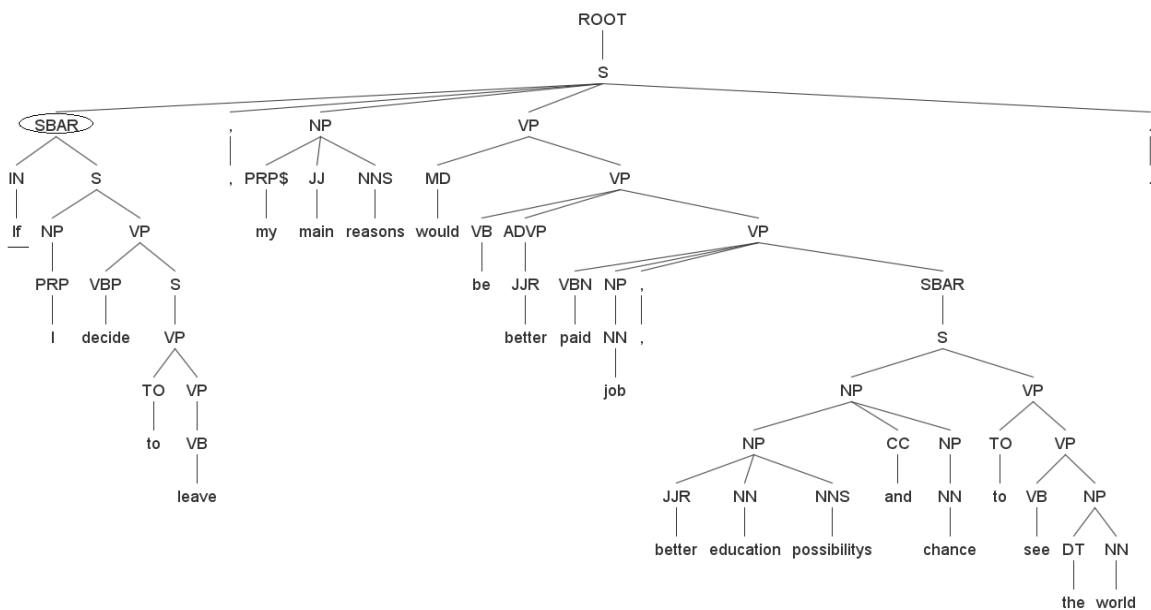
Figure 57 **Initial position of conditional clause in English**

[2]
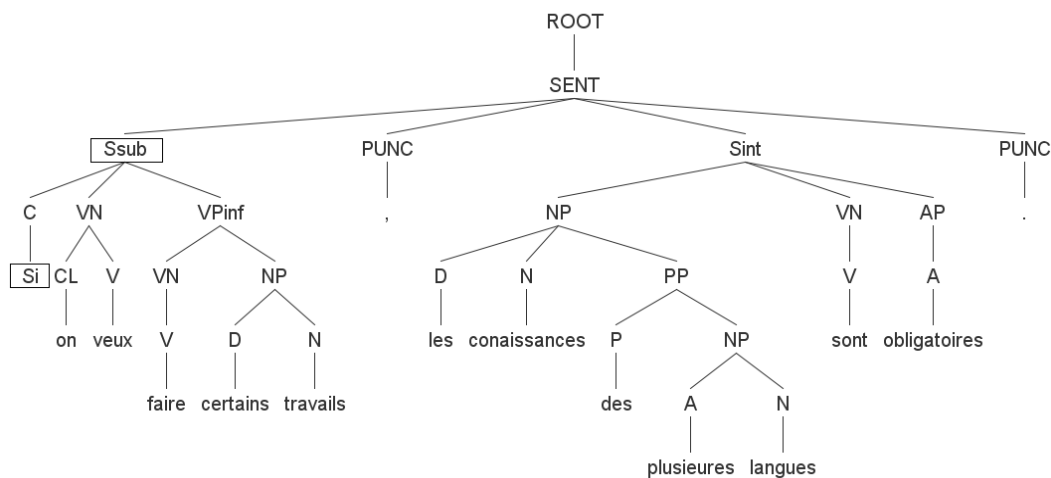
Figure 58 **Initial position of conditional clause in French**

The subordinate clauses [*if I decide to leave*] and [*si on veux faire certains travails*] appear at the initial position of the sentence, i.e. before the main clause. This structure is the one that predominates in the learner written performance.

In the second case (*Figures 60* and *61*), the conditional clause occupies the final position in the sentence:

[3]
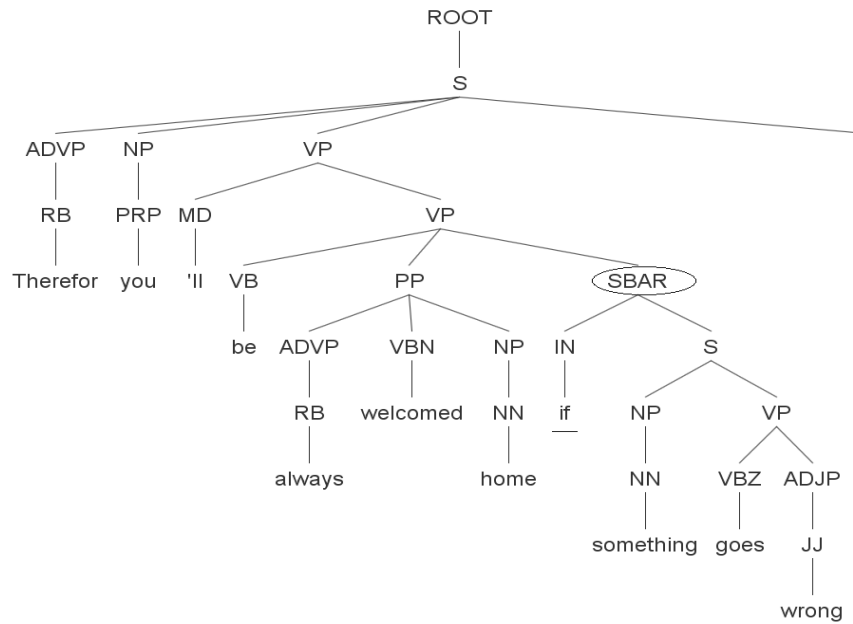
Figure 59 **Final position of conditional clause in English**

[4]
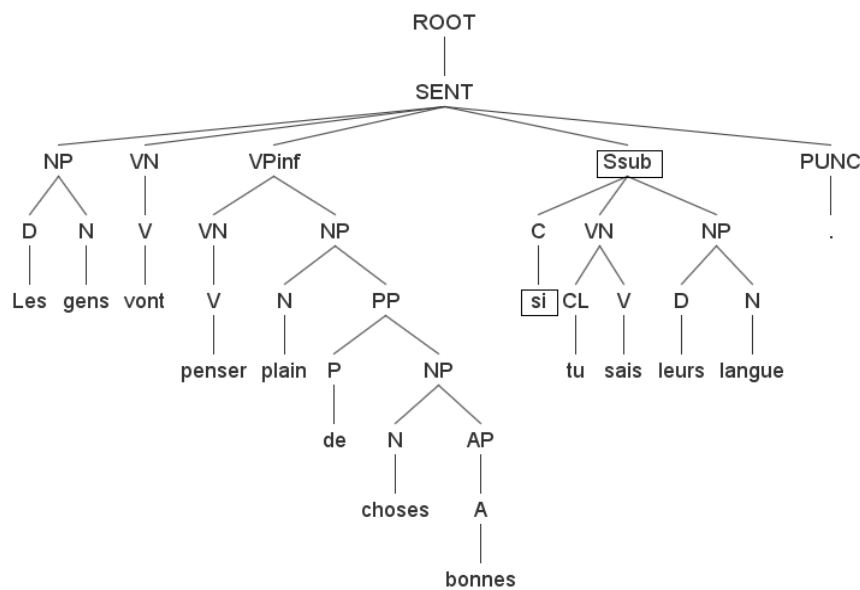
Figure 60 **Final position of conditional clause in French**

In the example sentences [3] and [4] the subordinate clause follows the main clause, i.e. the sub-clause [*if something goes wrong*] is subordinated to the main clause [Therefor you'll be always welcomed home] and [*si tu sais leurs langue*] is subordinated to the main clause [Les gens vont penser plain de choses bonnes]. However, the number of cases when the conditional clauses have been placed after the main clause are fewer than the cases when the conditional clauses are being used in the initial position (especially at lower levels of language proficiency). Moreover, if the conditional clause appears at the beginning of the sentence, it has to be separated by a comma from the main clause, which has not been done in rather many sentences at all levels of language proficiency.

Hence, we may conclude that the test-takers feel more secure with open conditional clauses in both languages. Although some students have attempted to use hypothetical conditional clauses as well, their number is rather low and only few of the clauses are grammatically correct. What is more, the number of conditional clauses is low at the higher levels of language proficiency in English, which might be explained by the fact that the sentences analysed contain only one subordinate clause (at higher levels students use more complex sentence structures). In majority of cases the conditional clauses are placed before the main clause introduced by the subordinator *if* in English and *si* in French, thus demanding a comma after the subordinate clause, which is rather often forgotten.

Adverbial clauses of reason and cause

Adverbial clauses of reason and cause are the third largest group marked in the test-taker essays. It comprises 52 % of all the subordinate clauses in English, while in French the number of this clause type is considerably lower attaining only 21% (in French the adverbial clauses of condition occupy the main position).

The data in *Figure 62* demonstrate the frequency distribution of this clause type. In English the adverbial clauses of reason and cause predominate at the lower levels of language proficiency ranging from 21 subordinate clause or 60% at level D to 14 subordinate clauses or 93% at level F. However, 56% at level A show that this clause type has also been used rather often by the students having the highest level of language proficiency. The percentage of this clause type is the lowest at levels B and C (44% at level B and 37.5% at level C). In French the situation is different. At level A the number of the adverbial clauses of reason and cause is the lowest amounting only to 7% or one clause. There is a gradual rise at levels B and C (at level B the numbers attain 21% and at level C

146

– 25%). Thus, we might say that there is a tendency for the increase towards the lowest levels of language proficiency. However, it is obvious that in English the students use this clause type more often than in French because of the aforementioned reason, i.e. the written task requirements.
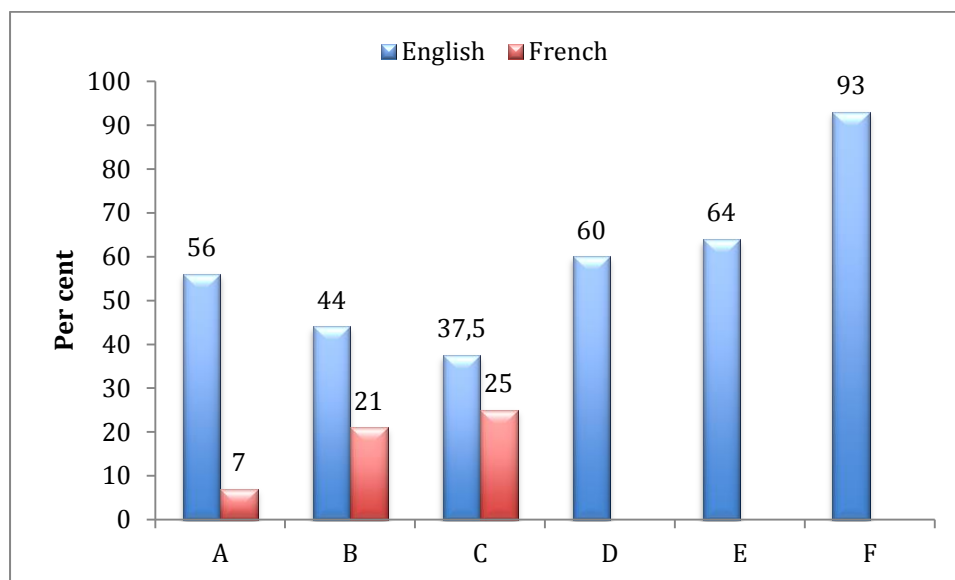
Figure 61 **Frequency of adverbial clauses of reason and cause**

Though, according to Quirk et al. (1985), the adverbial clauses of reason and cause might be sub-divided into several groups taking into consideration their semantic meaning (*Chapter 2*), it is rather complicated to interpret what notion, e.g. effect, motivation or result, is expressed by each subordinate clause since they have been retrieved from their context. Nonetheless, all the sentences marked in the student essays in both languages have the same syntactic structure, i.e. the main clause is subordinated by the adverbial clause of reason introduced by the subordinator *because* in English and *parce que* in French:

Level A

The job issue is pretty bad, [*because people can't get any good jobs or any jobs at all*].

Les gens devient plus ouverts [*parce que ils ne doivent pas parler anglais*].

Level B

Some are trying to fulfil their dreams abroad, [*because they see more opportunities there*].

Les gens doit étudier beaucoup les langues étrangères [*parce que la langue fait les gens plus inteligents*].

Level C

At the moment it is very dangerous to use the public transport [*because there are a lot of criminals*].

C'est tres important en Lettonie [*parce que notre pays n'est pas tres grand*].

Level D

Nowadays a lot of people go to the other countries [*because they have some reasons to do it*].

Level E

Some of pepoles go to other countries [*because they want spend they life at other countries*].

Level F

People have left Latvia during the last feve years [*becos Latvia go in the Europe*].
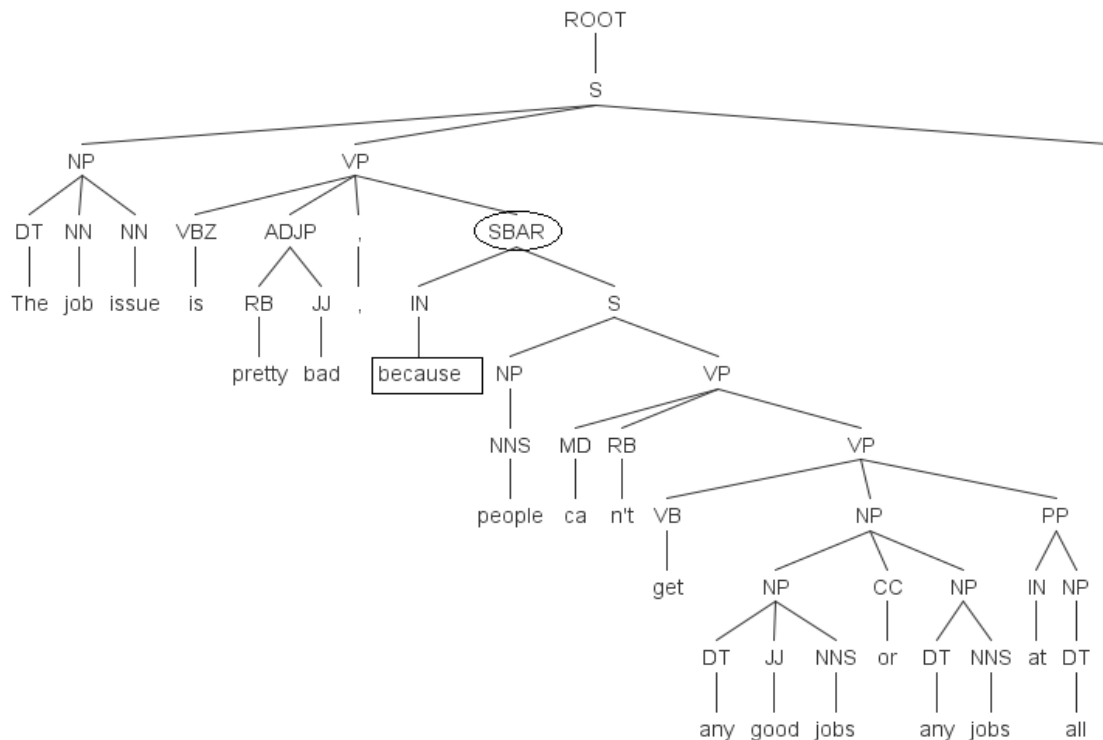
Figure 62 **Adverbial clause of reason in English**

The complex sentence in English (*Figure 62*) begins with the main clause [The job issue is pretty bad] and is subordinated by the adverbial clause of reason [*because people can't get any good jobs or any jobs at all*].

ROOT

SENT

NP VN AP Ssub PUNC

D N V ADV A MWC VN ADV VPinf

Les gens devient plus ouverts C C CL ADV V pas VN AP

parce que ils ne doivent V A
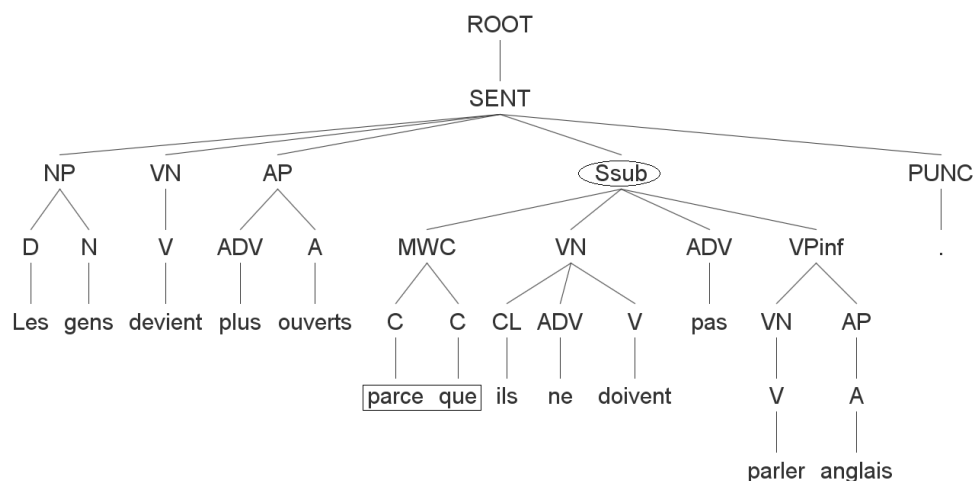
parler anglais

Figure 63 **Adverbial clause of reason in French**

In French (*Figure 64*) the sentence structure containing the adverbial clause of reason is similar to that of English, i.e. the adverbial clause [*parce que ils ne doivent pas parler anglais*] introduced by the subordinator *parce que* is subordinated to the main clause [Les gens deviant plus ouverts].

Overall, the research results show that the students, apart from the adverbial clause of time and place, and condition, use the adverbial clause of reason and cause quite effectively, especially at the lower levels of language proficiency. Consequently, they might be regarded as criterial for the lower levels of linguistic knowledge. As for the other types, their number is rather low, ranging from one to two sentences per level (if used at all), which is not enough for the qualitative analysis.

Adjectival clause (= relative clause)

Adjectival clauses are the third type of subordinate clauses examined in the EFLWT corpus. The quantitative data show (*Figure 65*) that the students learning English as a foreign language use this clause type more than the students of French. Moreover, their number increases towards the higher levels of language proficiency attaining 40% at levels A and B in English, while in French the columns reach 29% at level A and 20% at level B, which constitutes almost one half of the number of sentences used in English at the same levels. However, in English we can observe a rapid decrease at level C (the numbers plummet to 22%), whereas in French at level C the data remain stable – 22%. Then at level D there is a sharp rise again, and the numbers reach 37%, which is followed by a considerable drop at levels E (20%) and F (10%).
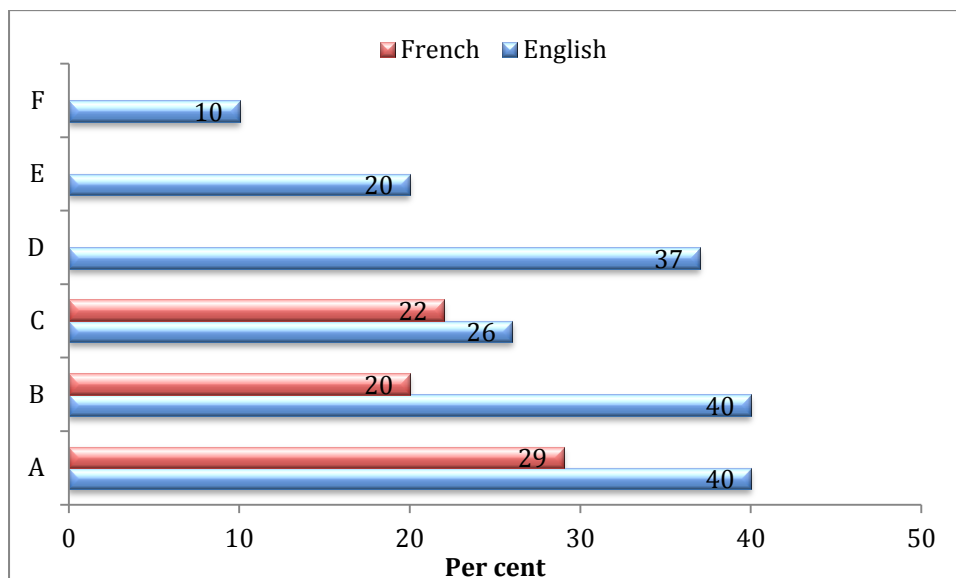
Figure 64 **Distribution of adjectival clauses across levels**

The situation is different if we do not count *why*–clauses produced by the students of English, as this clause type was given in the task stating that 'One of the main reasons why people have left Latvia during the last few years is that...' and asking them to add two other reasons. Naturally, students tend to reproduce the already given sentence structure, which does not reveal the veritable capability of producing the adjectival clauses of cause.
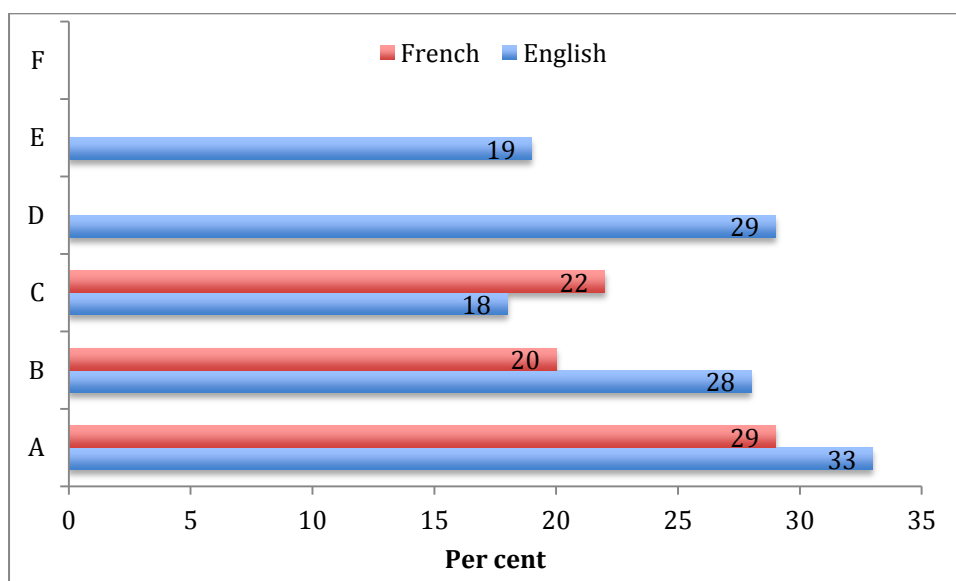
Figure 65 **Distribution of adjectival clauses across levels without why-clauses**

Thus, *Figure 66* reflects more objective numbers, although the data do not differ noticeably. Adjectival clauses are used most effectively at level A (33%) and at levels B (28%), and D (29%). Compared with the data in *Figure 65* the number of adjectival clauses at level C remains rather low (18%), almost the same as at level E (19%). What is

more, at level F the test-takers of English have been able to reproduce only *why*-clauses given in the task. Hence, there are no adjectival clauses used at level F.

Overall, students of English at the higher levels of language proficiency utilize more adjectival clauses in their written performance than the students of French, which is reflected in both diagrammes.

For a more profound analysis the adjectival clauses were further subdivided into restrictive and non-restrictive relative clauses. The diagrammes below clearly reveal the tendency of using restrictive relative clauses more often than non-restrictive relative clauses, which has also been proved by Biber et al. (1999) who stated that 'restrictive clauses are much more common than non-restrictive clauses (marked by a comma) in all written registers' (1999: 603). Moreover, both clause types are rare in conversation, which might explain the absence of relative clauses at lower levels of language proficiency (see *Figures 67 and 68*) since the students are not capable of producing more complex sentence structure that do not appear in every day speech.
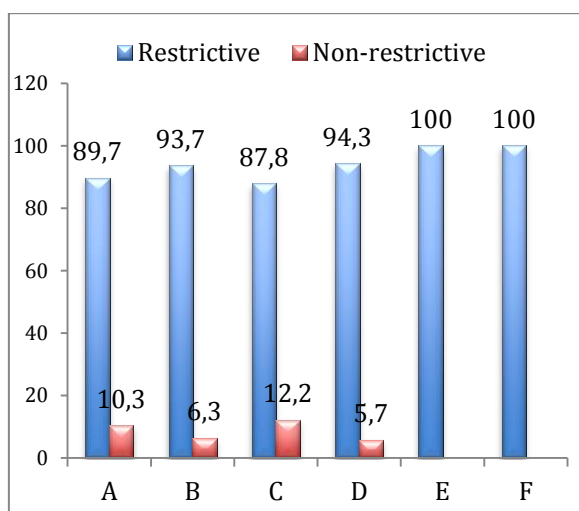
Figure 66 **Percentage of relative clauses in English**
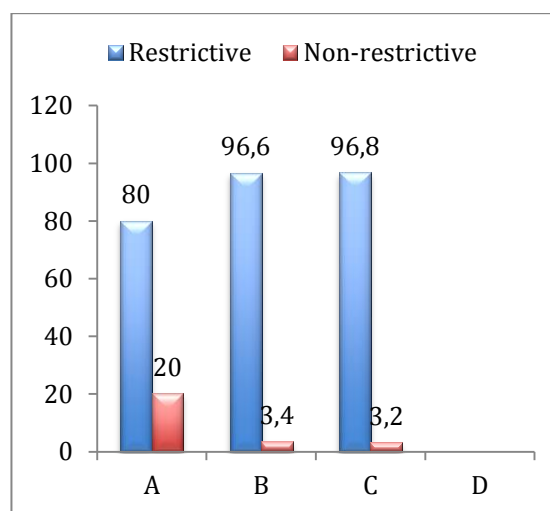
Figure 67 **Percentage of relative clauses in French**

The statistical data demonstrate that the frequency of non-restrictive relative clauses varies from 5.7% to 12.2% in English (at levels E and F no non-restrictive relative clause has been marked), whereas in French it amounts from 3.2% at level C to 20% at level A.

Relative clauses have been distinguished from other clauses by the use of relative pronoun, which not only deciphers between restrictive and non-restrictive relative clauses, but also can function as subject, object, complement and adverbial.

Table 22 **Functions (%) performed by relative pronouns in restrictive relative clauses**

Level	A		B		C		D	E	F
Language	En	Fr	En	Fr	En	Fr	En	En	En
Subject	34.61	22.22	38.33	64.28	38.88	53.33	40.0	33.33	
Object	19.23	22.22	6.66	32.14	5.55	16.66	6.00	20.0	
Prepositional complement	3.84	23.22				13.33			
Adverbial	42.30	44.44	55.0	3.57	55.55	16.66	54.0	46.00	100.0

Table 23 **Functions performed by relative pronouns in non-restrictive relative clauses**

Level	A		B		C		D	E	F
Language	En	Fr	En	Fr	En	Fr	En		
Subject	1	1	4	1	5	1	2		
Object							1		
Adverbial	2								

Figures in *Tables 22* and *23* reflect the frequency of different functions performed by relative pronouns in the marked sentences. The students at all levels of language proficiency (except at level E in English and level D in French) make an extensive use of restrictive relative clauses where the relative pronoun functions as subject. Relative pronouns functioning as adverbials occupy the second place in the frequency of usage, which is pursued by the relative pronoun operating as the object in the clause. As for the function of prepositional complement, only six cases have been marked in the English learner corpus.

If we compare the obtained data of different types of relative clauses with the findings of Paula Buttery (2010) (see *Table 24*) when working on criterial features for each CEFR level, then we might see that the subject relative clause has been the most widely used clause type among learners of English all over the world as their number ranges from 61% to 75%.

Table 24 **Usage of different types of relative clauses as percentage of total within each CEFR level (Hawkins and Buttery, 2010: 14)**

Level	A2	B1	B2	C1	C2
Subject RCs	67.74%	61.05%	71.11%	70.35%	74.80%
Object RCs	30.65%	37.33%	26.09%	25.02%	20.90%
Ind/obl RCs	1.61%	1.62%	2.80%	4.63%	4.30%

* Ind/obl RCs correspond prepositional complement clauses in *Table 22*.

The data (also retrieved by Buttery) from various subcorpora of the British National Corpus allow us to compare the learner data of the present research (*Tables 22 and 23*) with the data from the native speakers of English (*Table 25*), which demonstrate that the subject relative clause is the most frequent clause type among the others as its numbers outreach 80%.

Table 25 **Usage of different types of relative clauses as percentage of total within subcorpora of the BNC (Hawkins and Buttery, 2010: 14)**

	The Guardian	Mills and Boon	New Scientist	The Law Report
Subject RCs	77.20%	75.63%	80.67%	58.32%
Object RCs	17.02%	20.03%	12.44%	21.98%
Ind/obl RCs	5.76%	4.34%	6.89%	19.71%

Overall, all the aforementioned corpora provide similar information about the frequency of different relative clause types, in which the subject relative clause is the most extensively used clause type followed by the object relative clause.

The further research will be attributed to the structural analysis of relative clauses as well as to the more profound analysis of various functions performed by relative pronouns within a clause.

Subject relative pronoun in restrictive relative clause

The use of relative pronouns is determined by the antecedent, which can be either personal or non-personal. If the antecedent is personal, *who* is preferred instead of *that* in English, while in French *qui* is used in both cases. At all levels of language proficiency students of English and French prefer using relative clauses that contain personal antecedent:

Level A

During the last few years we have heard a lot of information about Latvians ***who*** *have left their country.* (personal)

There are many organizations [***that*** *offer this exchanging*]. (non-personal)

La question sur les langues étrangères a toujours été important surtout parmis les jeunes [***qui*** *pensent à leur profession future*]. (personal)

Il y a plusieurs facteurs [***qui*** *approuvent ça*]. (non-personal)

Level B

A lot of them are young people [***who*** *are getting education abroad*]. (personal)

Other reason for leaving Latvia is well developed technologies and infrustructures [***that*** *need bright heads to keep progressing*]. (non-personal)

Les gens [***qui*** *connaissent beaucoup de langues*] sont respectés par autres. (personal)

Par example ça faire des problémes de voir des films [***qui*** *sont à l'autre langue*]. (non-personal)

Level C

Peoples [***who*** *live in Latvia*] is very helpfull and frendly. (personal)

Secondly, there are a lot of countries, [***thet*** *are bettter paid*]. (non-personal)

In the second sentence the student has put a comma before the relative clause, which is wrong as the relative clause defines or specifies the noun *countries*.

Toujours il va être personnes [***qui*** *parle different langue que l'anglais*]. (personal)

L'internet démole les lignes [***qui*** *sont entre les payes et les gens*]. (non-personal)

Level D

I know many people [***who*** *want to study in other country*]. (personal)

Very importand problem for the young people is getting a good education [***that*** *can allow them to get a good proffesion and then good work*]. (non-personal)

Level E

I'm anderstand this peaples [***who*** *go to the ather contrys*].(personal)

In the examples the relative pronouns *who* and *qui* are related directly to the noun that is personal; the relative pronouns *that* and *qui* – to the noun that is non-personal. There are no sentences containing *zero* pronoun, which is typical of English.

There are two main sentence structures characteristic of restrictive relative clauses: 1) main clause + subordinate restrictive relative clause; 2) subordinate restrictive relative clause embedded in the main clause.
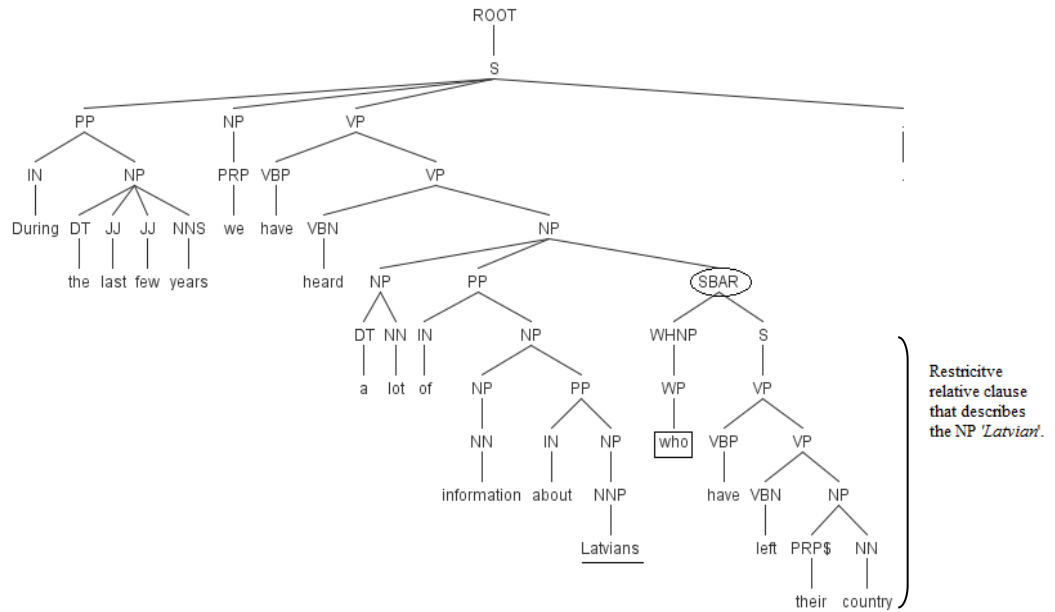
Figure 68 **Restrictive relative clause defining a personal antecedent in English**

In *Figure 69* the restrictive relative clause introduced by the subject relative pronoun *who* [*who have left their country*] is subordinated to the main clause [*During the last few years we have heard a lot of information about Latvians*] and defines the personal antecedent [*Latvians*].
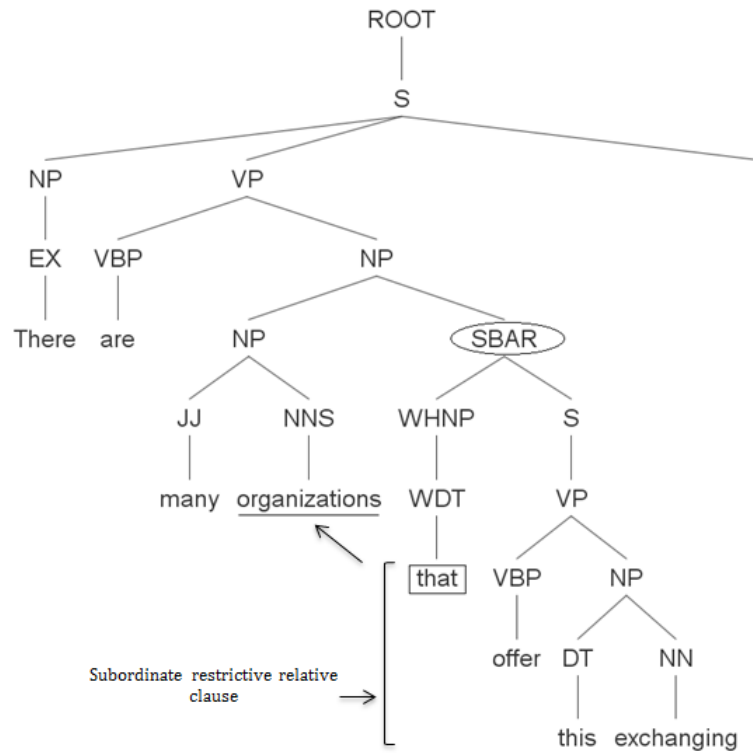
Figure 69 **Restrictive relative clause defining a non-personal antecedent in English**

In the example sentence (*Figure 70*) the relative pronoun *that* appears at the beginning of the relative clause [*that offer this exchanging*] defining the antecedent [*many organizations*] in the main clause, which is non-personal. Thus, the relative clause is subordinated to the main clause.
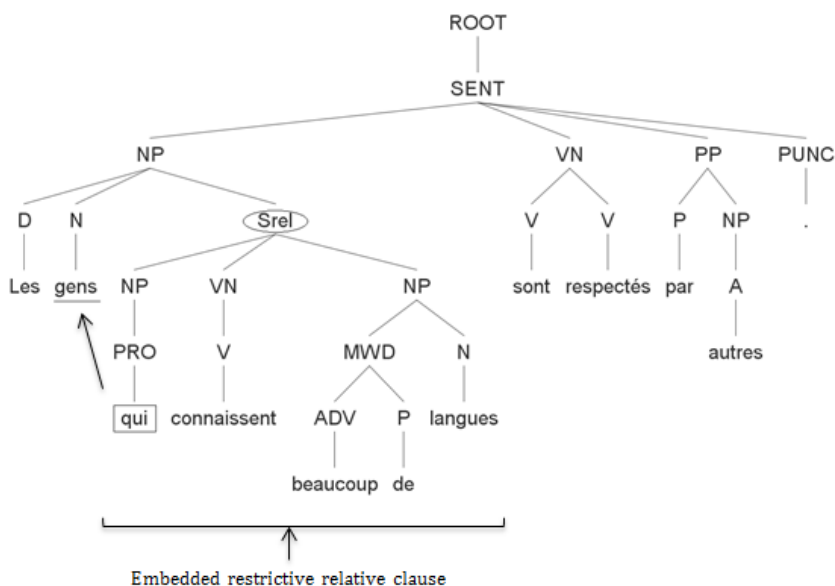
Figure 70 **Restrictive relative clause defining a personal antecedent in French**

In the sentence above we can see that the relative clause [*qui connaissent beaucoup de langues*] may be embedded in the main clause [*Les gens sont respectés par autres*], and the relative pronoun *qui* introduces the clause [*qui connaissent beaucoup de langues*] being directly linked to the antecedent [*les gens*].

Subject relative pronoun in non-restrictive relative clause

The number of non-restrictive relative clauses marked in the student essays in both languages is rather low, which is comprehensible, as restrictive relative clauses are more common both in oral and written performance. There are only 12 non-restrictive relative clauses containing subject relative pronoun in English and 3 sentences in French. All of them have been marked in the student essays of higher levels of language proficiency.

Level A

Finally, there are trouble makers and <u>adventure lovers</u>, [*who cannot stay for a long time at one place*]. (personal)
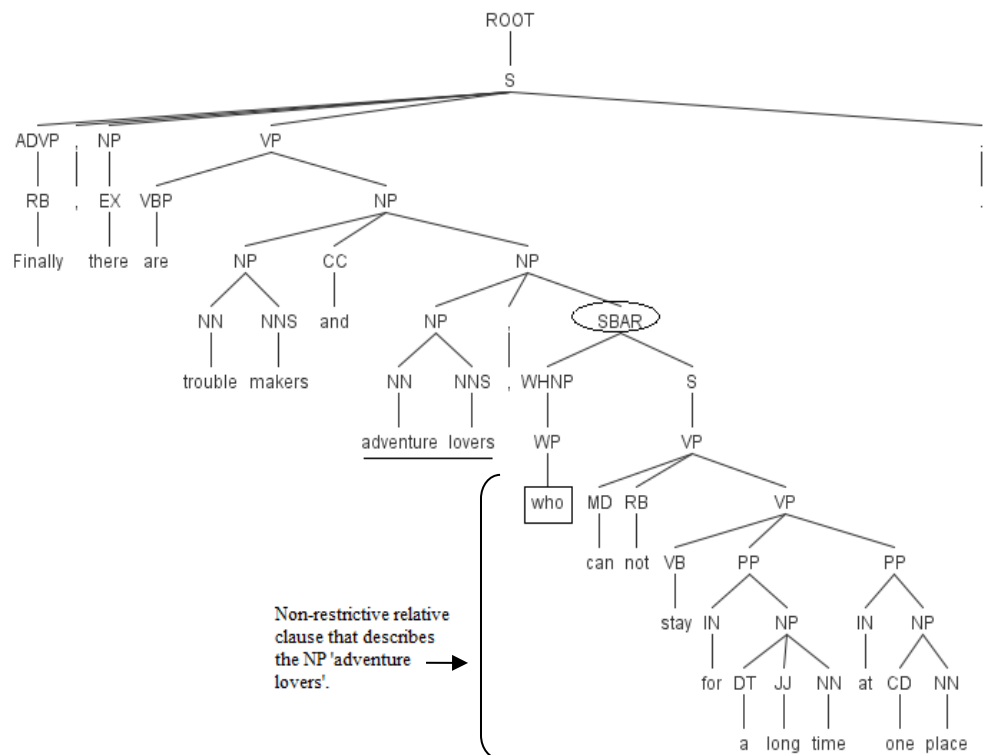
Figure 71 **Non-restrictive relative clause defining a personal antecedent in English**

Subject relative pronoun *who* introduces the non-restrictive relative clause [*who cannot stay for a long time at one place*] that is subordinated to the main clause [*Finally, there are trouble makers and adventure lovers*] and describes the NP [*adventure lovers*].

Par exemple, on doit comprendre l'espagnol pour comprendre le vrai beauté de flamenko _ [*qui est une partie importante dela culture espagnole*].

Level B

From all countries we have the most expressive history, hard working people, [*who are smart and can concurate with others in high level*]. (personal)

The Republic of Latvia is currently in the state of recession, [*which encourages people to seek a better place to live.* (non-personal)

Anglais est une de raisons de la globalisation _ [*qui se passe dans le monde*].

157
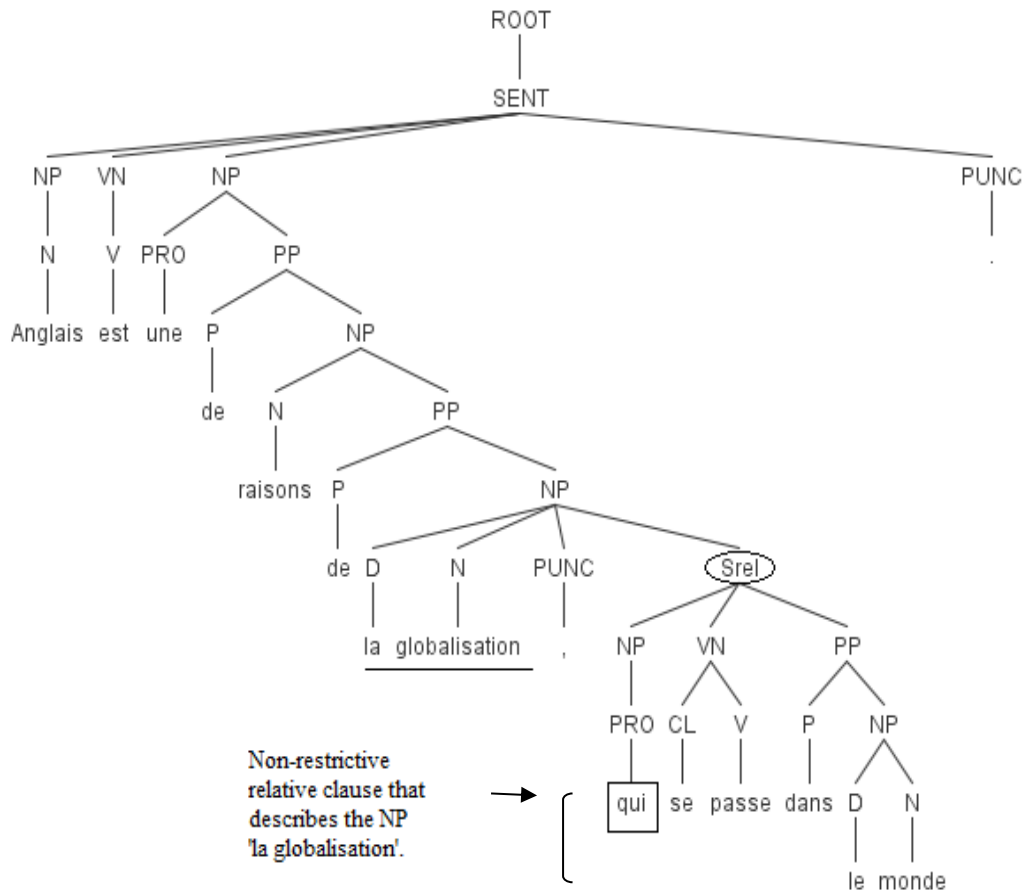
Figure 72 **Non-restrictive relative clause defining a non-personal antecedent in French**

In French (*Figure 73*) the sentence containing the non-restrictive relative clause [*qui se passe dans le monde*] subordinated to the main clause [Anglais est une de raisons de la globalisation] is similar to the one in English (*Figure 72*). The only difference is that in French the subject pronoun *qui* is utilized to describe both personal and non-personal nouns (see also level C), while in English *who* describes a personal noun and *that* – a non-personal noun.

Level C
>They also dramaticaly raised up <u>many taxes</u>, [*which* wouldn't let many companies work normaly]. (non-personal)
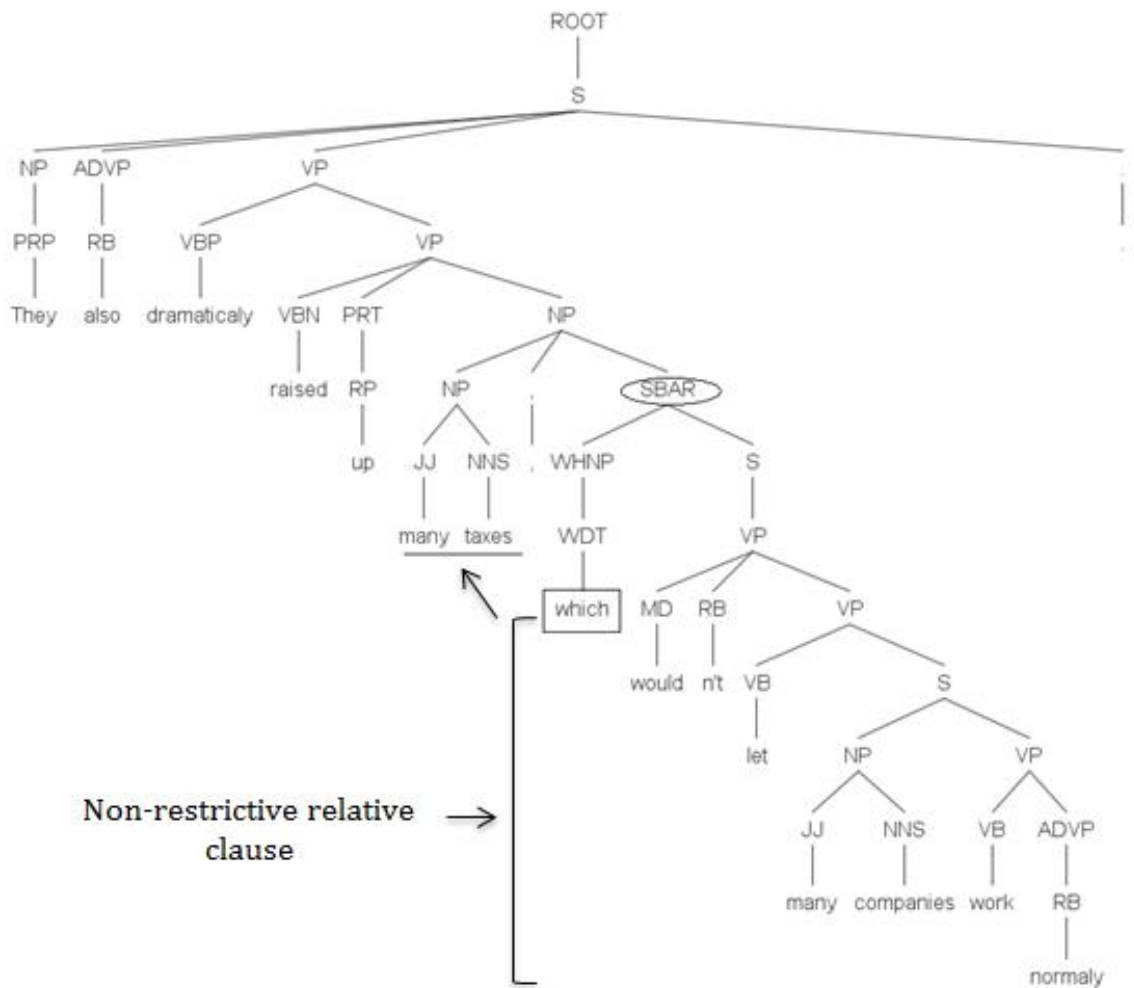
Figure 73 **Non-restrictive relative clause defining a non-personal antecedent in English**

Subject relative pronoun *which* introduces the non-restrictive relative clause [*which wouldn't let many companies work normaly*] that is subordinated to the main clause [They also dramatically raised up many taxes] and describes the noun [*taxes*], which is non-personal.

Par example mon amis Jānis _ [*qui à passé l'examen en englais « Gelts »* _ va etudiera a Denmark]. (personal)
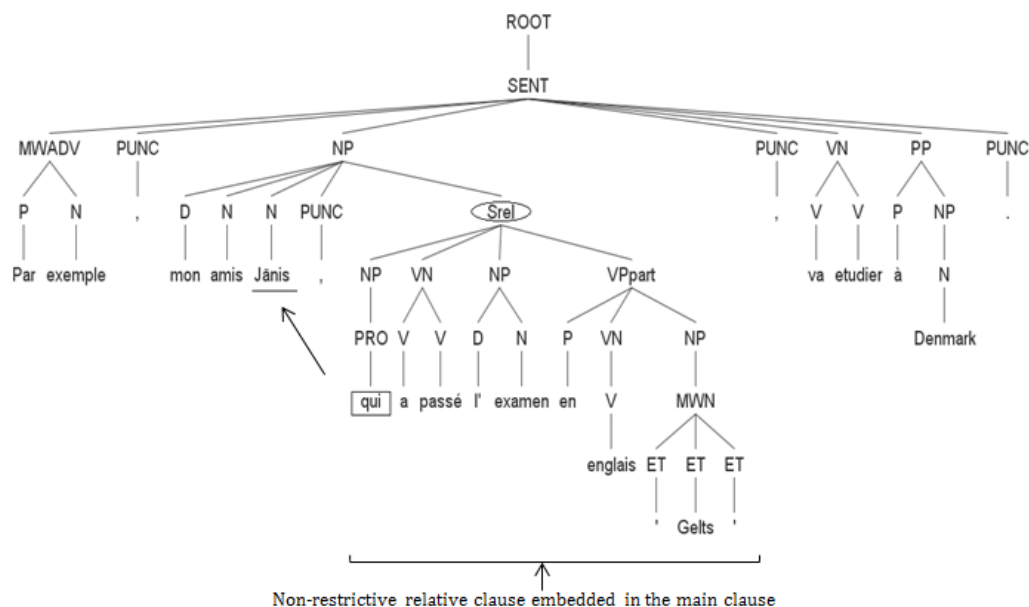
Figure 74 **Structure of the non-restrictive clause defining a personal antecedent in French**

The non-restrictive relative clause [*qui à passé l'examen en englais 'Gelts'*] is introduced by the subject relative pronoun *qui* and is embedded in the main clause [*Par example mon amis Jānis va etudiera a Denmark.*].

Level D

They start new life, [**which** is better then prewious]. (non-personal)

The examples retrieved from the compiled corpora allow the author to conclude that in English the students are aware of the differences between restrictive and non-restrictive relative clauses, and the relative pronouns to be used for a personal and non-personal noun which are described or defined by the subject relative pronouns *who, that* and *which.*

In French the students focus mainly on restrictive relative clauses. As for the non-restrictive relative clauses, they use them only to describe a personal noun. Moreover, they put neither comma before the relative pronoun *qui* when the non-restrictive relative clause is subordinated to the main clause, nor separate the embedded non-restrictive clause by putting it into commas (level C).

In both languages the main syntactic structures are: 1) main clause + subordinate relative clause; 2) subordinate relative clause embedded in the main clause.

Object relative pronoun in restrictive relative clause

Relative clauses in which the relative pronoun performs the function of *object* have been used less than the pronoun functioning as subject (in English 17 cases have been examined

160

and in French – 15). At all language proficiency levels the students of English and French have used relative clauses to define a *non-personal* noun, i.e. they have applied the relative pronouns *which, that* and *zero* in English, and *que* in French. Only one sentence in French has been marked in which the restrictive relative clause defines *a personal* noun at level B:

Level A

They get work experience [*that they would never get in Latvia*]. (non-personal)

Dans une langue on peut dire les choses [*qu'une autre langue ne permet pas*]. (non-personal)

Level B

Mostly Latvians and other immigrants can find only those jobs [*which loads doesn't want to do*]. (non-personal)

On peut peut-être communiquer en langue [*que les autres ne comprennent*]. (non-personal)

Pour devenir la personne [*que je veux devenir*], je dois savoir l'anglais, le français et les autres langues. (personal)

In the above sentence the restrictive relative clause [*que je veux devenir*] is embedded in the main clause [Pour devenir la personne je dois savoir l'anglais, le français et les autres langues], therefore no comma is needed after the relative clause.

Level C

Some part of society can't buy anymore things, [*that they have bought a couple of years ago*]. (non-personal)

The relative clause [*that they have bought a couple of years ago*] specifies the antecedent *things* in the main clause, therefore no comma is needed before the relative pronoun *that*.

Tu comprendrais tout [*que on dit*]. (non-personal)

Level D

And the money [*what we have* isn't so much for living a normaly life]. (non-personal)

In the sentence *what* is used incorrectly. It has to be replaced by the relative pronoun *that*. This example provides evidence how the native language influences the incorrect use of the word in a foreign language, especially at the lower levels of language proficiency.

Level E

One think in Latvia , [*I don't like wery much* is politic]. (non-personal)

The examples show that students tend to use two main sentence structures. The first one - the main clause is followed by the subordinate clause:
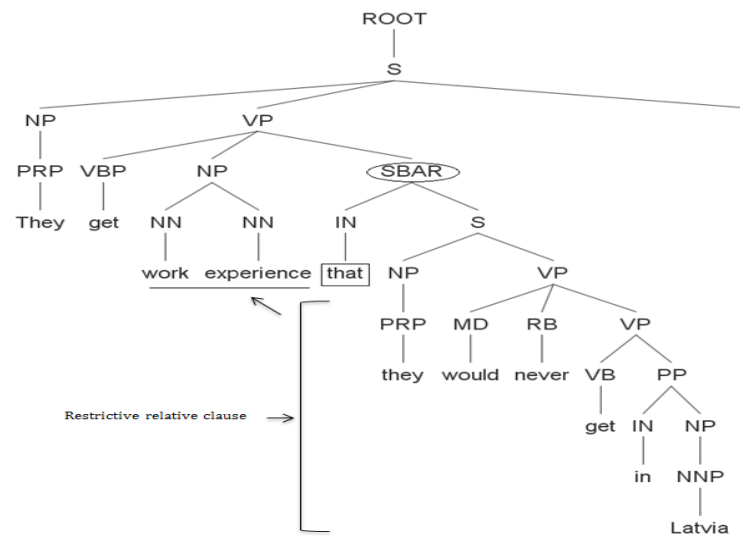
Figure 75 **Restrictive relative clause functioning as object in English**

Thus, the main clause [*They get work experience*] is directly followed by the restrictive relative clause [*that they would never get in Latvia*].

The second structure that students use is - the restrictive relative clause embedded in the main clause:
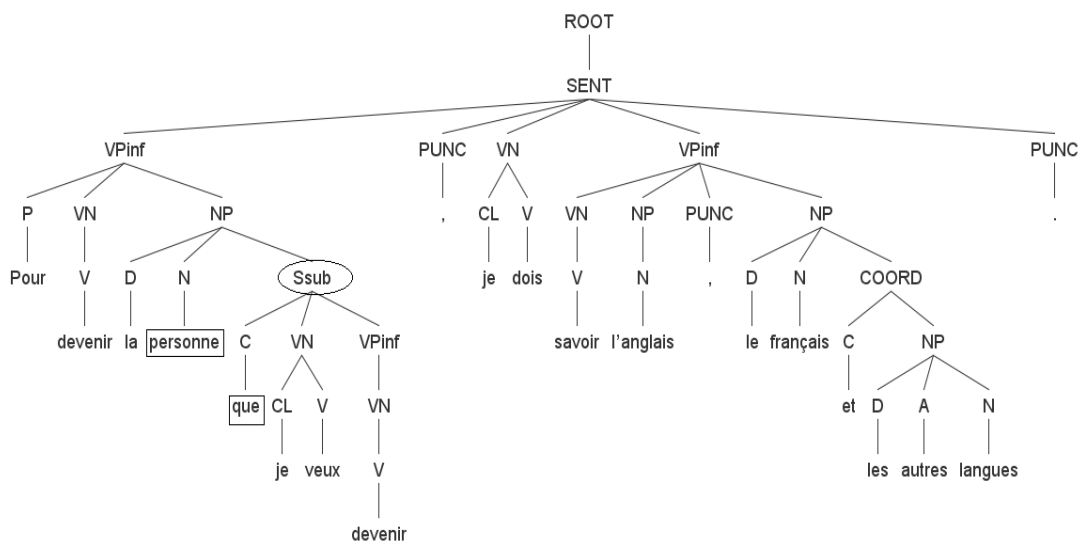
Figure 76 **Restrictive relative clause functioing as object in French**

In the above sentence the restrictive relative clause [*que je veux devenir*] is embedded in the main clause [*Pour devenir la personne je dois savoir l'anglais, le français et les autres langues*], and describes the personal antecedent [*la personne*].

Object relative pronoun in non-restrictive relative clause

The number of non-restrictive relative clauses in which the relative pronoun functions as *object* is very low – 1 sentence has been marked only in English at level D.

Level D

One of reason, [**which** I know] – it is education.
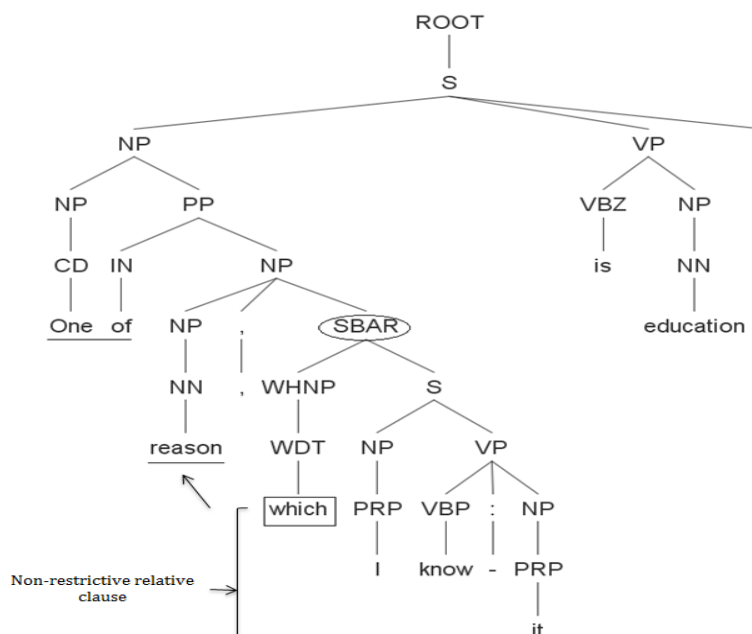
Figure 77 **Non-restrictive relative clause functioning as object in English**

In the example sentence the non-restrictive relative clause [*which I know*] emebedded in the main clause [*One of reason it is education*] is introduced by the relative pronoun *which* that describes the non-personal antecedent [*one of reason*]. However, it is obvious that the student who has produced this sentence does not know how to use the syntactic structures appropriately as the sentence has been built incorrectly.

The number of the examined sentences in which relative pronoun functions as object makes us conclude that the students of both English and French are not capable of using these syntactic structures appropriately. Moreover, almost all sentences contain the relative clause which defines or describes a non-personal antecedent. As for the position of the relative clause in a sentence, the students of English mainly produce the syntactic structures in which the relative clause is in subordination to the main clause (except at level D), while in French the students prefer to have the relative clause embedded in the main clause.

Relative pronoun as prepositional complement in restrictive relative clause

Only six sentences where the relative pronoun functions as prepositional complement have been marked in both languages:

Level A

[1]　　Lately Latvia suffers of criticism about the Government and <u>everything</u> [( ) *they are responsible for*]. (non-personal)
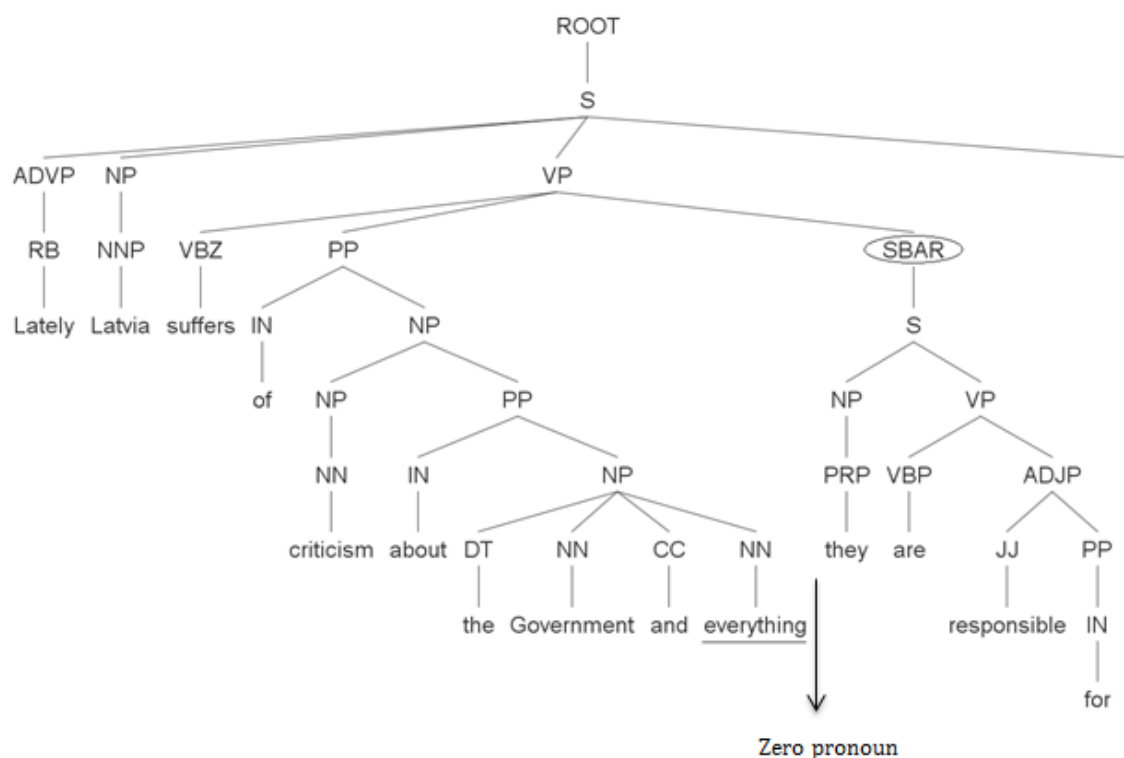
Figure 78 **Restrictive relative clause functioning as prepositional complement in English**

In the example [1] the *zero* pronoun defines the non-personal noun [*everything*], and the preposition *for* is deferred, i.e. it appears after the VP [*are responsible*]. This sentence does not clearly exemplify the use of the deferred preposition since it is more common to have this construction with prepositional verbs, e.g. That's the book ( ) he's been <u>looking</u> **for** (Quirk et al., 1987: 1253).

[2]

　　　Comme <u>la troisieme raison</u> [*pour laquelle il faut étudier des langues étrangères*] je peux mentionner les voyages. (non-personal)

Level C

[3]　　C'est la <u>langue</u> [*avec laquelle on peux gagne plus d'information*]. (non-personal)
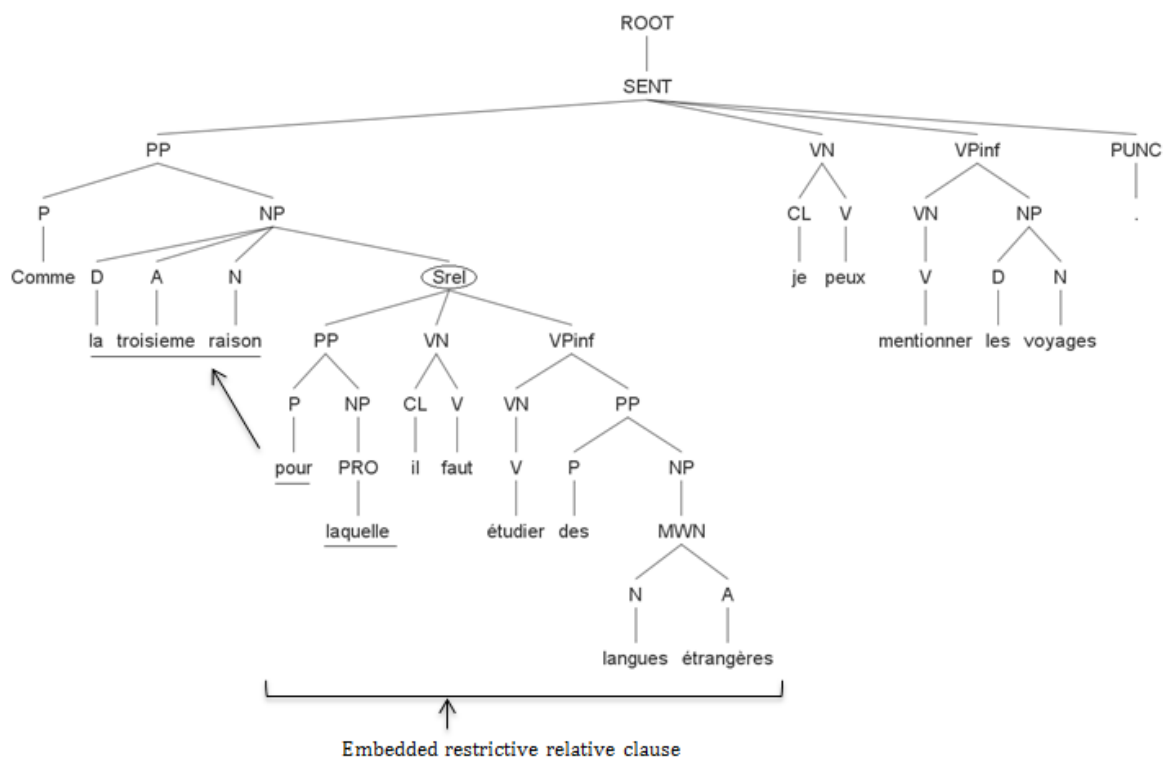
Figure 79 **Restrictive relative clause functioning as prepositional complement in French**

In sentence [2] the relative clause [*pour laquelle il faut étudier des langues étrangères*] introduced by the relative pronoun [laquelle], which agrees in number and gender with the non-personal antecedent [*raison*], is embedded in the main clause [Comme la troisieme raison je peux mentionner les voyages.]. Thus, in French, when the relative clause is introduced by pronouns *pour* and *avec,* a more specific relative pronoun *lequel* has to be used (sentences [2] and [3]).

If we compare the syntactic structures in both languages, then we see that in English the restrictive relative clause is subordinated to the main clause, whereas in French – embedded in the main clause. Moreover, in English the preposition might be used either together with the relative pronoun or is deterred, i.e. it appears at the end of the clause. In French the preposition always occupies the position before the relative pronoun.

Relative pronoun as adverbial in restrictive relative clause

The last function performed by relative pronouns, which has been marked in the sentences retrieved from the student essays, is the function of *adverbial*. It is the second most often used relative clause pattern; their number comprises 29 mono subordinate relative clauses

in English and 9 – in French. However, it should be stated that in English in many cases the students have constructed the relative clauses of *cause*, which have not been counted and further analysed as they are repetitive and have been copied from the task:

Usually on TV, radio and in newspapers we can hear reasons [*why people have left Latvia*].

But money is not the only reason [*why people leave*].

In the above examples the relative pronoun *why* specifies *the reason* for leaving Latvia, and the wording used in both sentences is rather similar. According to Quirk et al. (1985: 1256), the construction 'reason why' is tautologous; it is preferable to use *zero* pronoun or a nominal relative clause instead.

In French, only one sentence has been marked in which the student has used the relative clause of *cause*:

Level B

Une autre raison [*pourquoi l'anglais est devenu si important et populaire*] est grâce à sa simplicité.
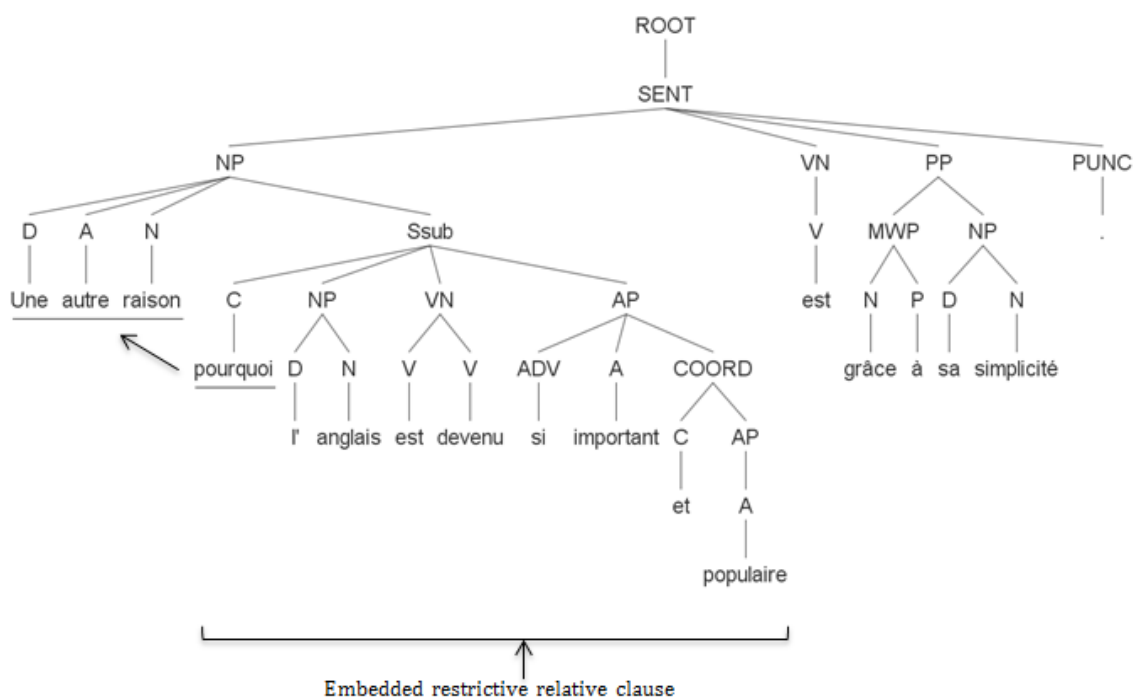
Figure 80 **Restrictive relative clause functioning as adverbial in French**

In the example sentence the relative clause [*pourquoi l'anglais est devenu si important et populaire*] specifies the antecedent [*une autre raison*] and is embedded in the main clause [*Une autre raison est grâce à sa simplicité.*].

166

As for other adverbials, several patterns have been marked in English in which students instead of using the construction 'preposition + pronoun' (e.g. *in which*) have replaced it by special adverbials *when* and *where*:

Level A

Therefore people decide to leave Latvia to find a place [*where the climate is warm and sunny during all the year*].

Level B

There was a time, [*when all people were happy, living their life, without problems*]. Latvia then is just place [*where they were born for them*].

Level C

In another countries Latvian people can find very good and respect work [*where worker can chaste normal money*].

Level D

We had many monuments, theatres, museums, cinemas, places [*where people can relax*].

Level E

Popular countries [*where people going*] are England and Ireland.

Latvian people make a days, [*when they cleaned nature*].

Level F

Five years ago Latvija do not be at the time, [*when we are live now*].

In several sentences (at levels B, E, F) students have used commas before the adverbial *when*, which is inappropriate. This might be signalling the influence of their native language (e.g. Latvian) as in Latvian a comma is always needed before the conjunction in complex sentences.

In French only 9 sentences have been marked which contain the adverbials of *time* and *place:*

Level A

(i)   Je ne peux pas imaginer une situation [*quand savoir une langue est inutile*].

(ii)   Toutes les langues étrangères laissent mieux connaître la monde [*où on vive*].

Level C

Il y a beaucoup des travailles [*où il y a necessaire la langues étrangères par example dans le bussines d'hotele*].

In sentence (i) the adverb *quand,* which denotes time, has been used incorrectly as it does not have a function of a relative adverb. Instead, one should use the relative adverb *où,* which denotes both time and place when used in relative clauses. Such occurrences could be explained by the influence of the native language as well in which the adverb of time could be used as a relative adverb.

In all the examples the same sentence structure has been used both in English and French, i.e. the relative clause is directly subordinated to the main clause.

Relative pronoun as adverbial in non-restrictive relative clause

Only one non-restrictive relative clause in which relative pronoun functions as adverbial has been marked at level A in English:

Movies tell us about a care-free life in America, [*where people have cars for almost everyone in the family, a big house with a white fence – an American dream*].
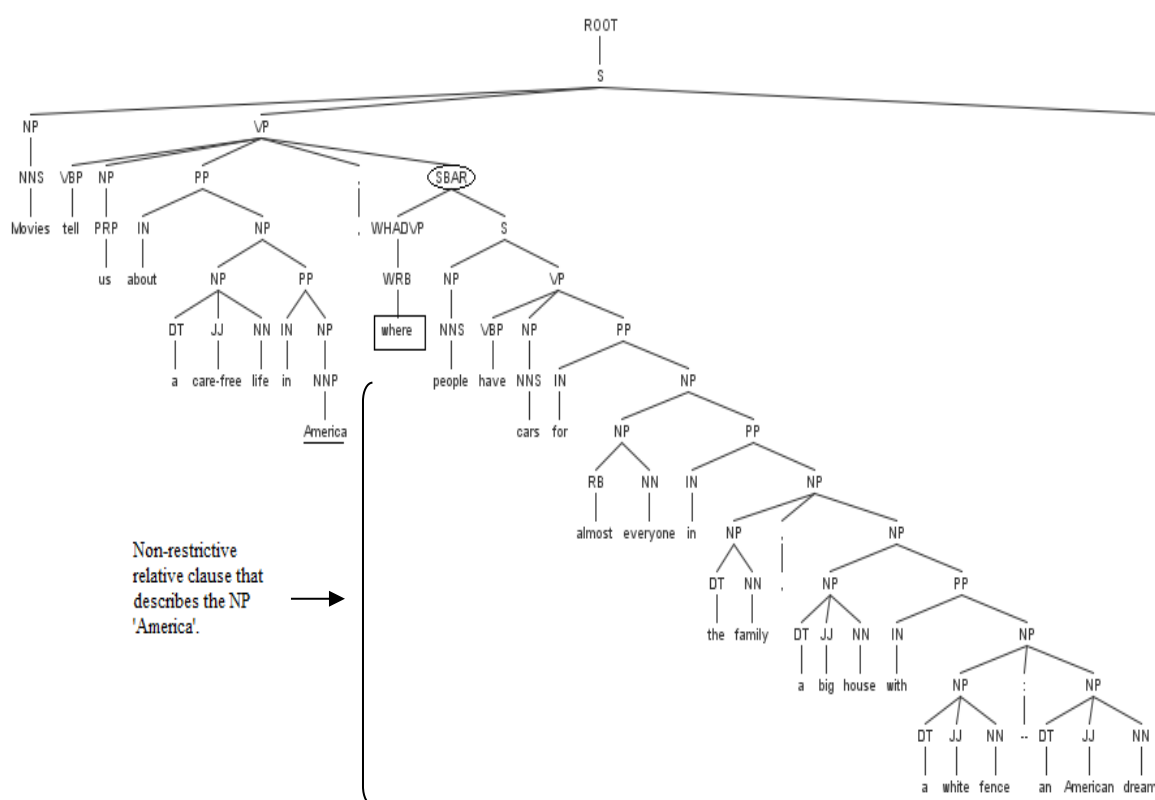
Figure 81 **Non-restrictive relative clause functioning as adverbial in English**

In the above sentence the non-restrictive clause [*where people have cars for almost everyone in the family, a big house with a white fence – an American dream*], introduced

by the adverbial *where,* is subordinated to the main clause [*Movies tell us about a care-free life in America*] and describes the non-personal antecedent [*America*].

In brief, the relative clauses in which the pronoun functions as adverbial are the second largest group marked in the student essays. However, the majority of test-takers prefer to use adverbials instead of relative pronouns. What is more, quite often there are mistakes in putting commas to decipher between restrictive and non-restrictive relative clauses, which might signal the lack of knowledge in using relative clauses appropriately.

As for the structures, [the main clause + subordinate clause] as well as [the relative clause is embedded in the main clause] have been utilised.

Having analysed different types of relative clauses containing a mono subordinate clause, the author of the present research concludes that:

–   restrictive relative clauses have been utilised more frequently than non-restrictive relative clauses, which complies with the research on native speaker language (e.g. Biber et al., 1999);

–   the learners of English and French are familiar with different types of relative clauses, though they are inattentive when deciphering between restrictive and non-restrictive relative clauses as there are rather many cases where commas are used inappropriately;

–   the majority of sentences marked in learner corpora contain four types of relative clauses in which the relative pronoun performs the function of a subject, object, prepositional complement and adverbial. No clauses have been marked with the function of complement and possessive determiner;

–   the main syntactic patterns utilised in relative clauses are:

Table 26 **Syntactic patterns of relative clauses**

in English	in French
who + VP that + VP	qui + VP
that + NP + VP which + NP + VP	que + NP + VP
( ) + NP + VP + preposition why + NP + VP where + NP + VP when + NP + VP	préposition + laquelle + NP + VP préposition + quoi + NP + VP où + NP + VP

–   the main sentence structures are: 1) the relative clause is embedded in the main clause; 2) the relative clause is subordinated to the main clause.

Having examined various types of mono subordinate clauses (noun, adverbial and adjectival) it could be concluded that the students in both languages – English and French – use the same subordinate clause types that perform the same functions in relation to the main clause, which complies also with the findings of Hawkins and Buttery (2010) on different types of relative clauses in English.

However, in order to provide evidence to the assumption that the syntactic complexity increases along with the foreign language proficiency development, a more profound analysis of the aforementioned syntactic patterns has been carried out by examining the complex sentences containing two subordinate clauses.

3.3.2 Complex Sentences Containing More Than One Subordinate Clause

In the following stage of the present research different multi subordinate clause types were manually marked in the complex sentences and consequently grouped according to their clause type. On the whole 46 variations were distinguished (see Appendix 10), ranging from two subordinations to six subordinations per sentence. However, in majority of cases two subordinations predominate both in English and French containing noun, adverbial and adjectival clauses in different combinations (see *Table 27*). It is evident that <N,ADV> clause combination occupies the first place both in English and French (95 cases were marked), the second place has been taken by <N,ADJ> clause combination marked in 66 cases, while the third most frequent clause combination is <ADV,ADJ>, which has been used 40 times. Withal, there is nothing unexpected since the research data on mono sub-clauses have already showed that the complex sentences containing noun and adverbial clauses predominate in the test-taker written essays at all levels of language proficiency both in English and French.

Table 27 **The most frequent multi sub-clause combinations across levels**

Types of multi sub-clauses	*English*						*French*				*Total*
	A	**B**	**C**	**D**	**E**	**F**	**A**	**B**	**C**	**D**	
<N,N>	2	3	3	4			1	5			18
<N,ADV>	8	11	15	16	8	4	6	13	14		95
<ADV,N>		7	1	2		1		5	3		19
<ADV,ADV>		4	3	3			1	7	3		21
<N,ADJ>	2	9	18	13	6	2	1	13	2		66
<ADJ,N>	3	5	3	4	3	1	3	2			23
<ADJ,ADJ>	2	7	3	2			3	6	3		26
<ADJ,ADV>	4	5	2	6	1	1		3	2		24
<ADV,ADJ>	3	7	3	7	1	1	2	12	4		40

Thus, for further analysis only nine (the most frequent variations) have been chosen.

Complex sentences containing <N, ADV> sub-clauses

The clause type <N, ADV> occupies the first place among all multi subordinate clause types marked in the student essays both in English and French. Their number reaches 95 sentences (see *Table 27*) and their distribution across levels is rather even. In English they range from 8 sentences at level A to 11 - at level B; 15 - at level C; 16 - at level D and then drop to 8 sentences at level E, and 4 – at level F. In French the highest number has been observed at level C – 14 sentences; at level B – 13 sentences, which falls at level A to 6 sentences. Thus, it could be maintained that this clause type has been used most effectively by the students at all language proficiency levels.

Level A

[1] I think $_1$[*people leave Latvia*] $_2$[***because** they have no future in this country*].

[2] En plus, je pense $_1$[***que** l'apprentissage des langues étrangères est très interessant et amusant*] $_2$[***parce que** ça nous donne une chance à se rendre avec les gens et faire connaissance*].

Level B

[3] There is a stereotype among people [**that** people have left Latvia during the last few years], [***just because** they are better paid in other countries*].

[4] They believe $_1$[***that** $_2$[***if** they know better other cultures*], *they will be more interesting for others, for themselfs*].

[5] Je crois $_1$[***que** personne a une bonne éducation*] $_1$[***s'il** parle quatre ou cinq langues*].

Level C

[6] I've heard $_1$[***that** many latvians have been returned from England*], $_2$[***because** even there they could not find job*].

[7] I think, $_1$[$_2$[***if** people want something very strong*], *they always get this in the end*].

[8] Je pense $_1$[***que** nous devons apprendre beaucoup de langues*] $_2$[***parce que** nous pourrions parler avec des étrangère*].

[9] Je pense $_1$[***que** $_2$[***si** je parle plus langues*] *je peux travailer aux pays different*].

Level D

[10] I suppose ₁[***that** many people leave Latvia*] ₂[***because** they don't know Latvian language*].

[11] I don't know ₁[***what** I will do*] ₂[***when** I graduate this school*].

Level E

[12] I think ₁[*in novadays many young people have left Latvija*], ₂[***because** their parents live in other contries*].

Level F

[12] I think ₁[*the people is going to Latvia*], ₂[***because** there not job and mony...*]

The examples demonstrate that in majority of cases the students produce the noun clauses having the function of direct object. As for adverbial clauses, the test-takers mostly use the adverbial clause of reason, which was noticed also when analysing mono sub-clauses. The second most frequently used adverbial clause is the adverbial clause of condition. However, there are some more variations, e.g. at level D, sentence [11] contains the interrogative noun clause [*what I will do*] functioning as direct object of the main clause, and the adverbial clause of time [*when I graduate this school*].

At level A (sentence [1]) the test-takers of English utilise zero pronoun at the beginning of the noun clause, which indicates that they are familiar with different syntactic structures. Though at lower levels the omission of the pronoun at the beginning of the noun clause may signal the restricted knowledge of linguistic structures, therefore the students make an extensive use of phrases that predominate in conversational language.

All in all, the above mentioned examples enable us to distinguish two main syntactic patterns that prevail in the student written essays:

1) In *Figure 83* the noun clause [*people leave Latvia*] introduced by zero pronoun and having the function of direct object is subordinated to the main clause [*I think*], while the adverbial clause of reason [*because they have no future in this country*] is syntactically dependent on the noun clause.
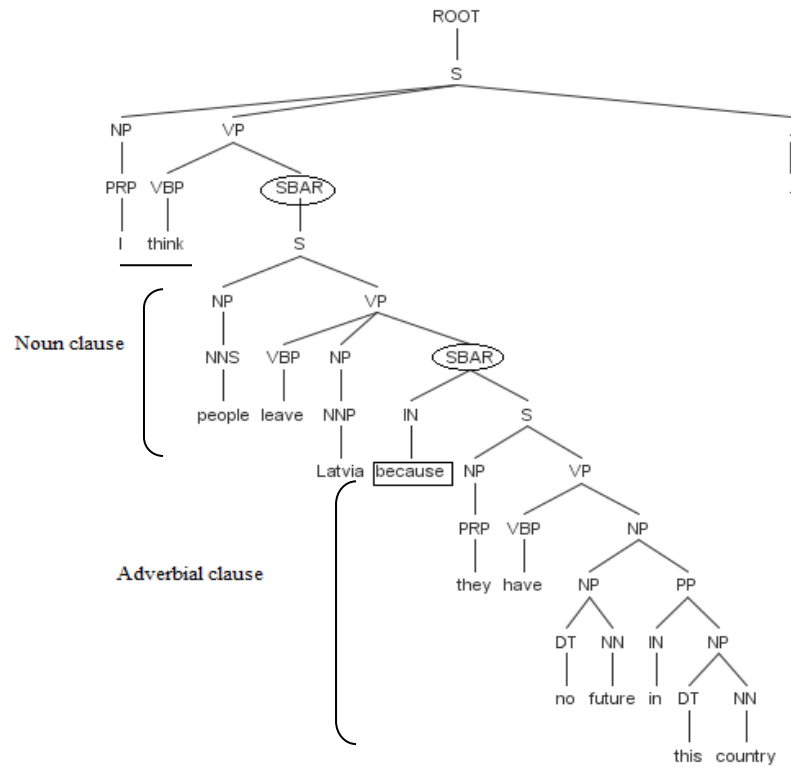
Figure 82 **<N, ADV> clause type in English**

2) In *Figure 84* the adverbial clause of condition [*si je parle plus langues*] is embedded in the noun clause [*que je peux travailler aux pays different*], which is introduced by the pronoun *que* and has the function of direct object. Moreover, it is subordinated to the main clause [Je *pense*].
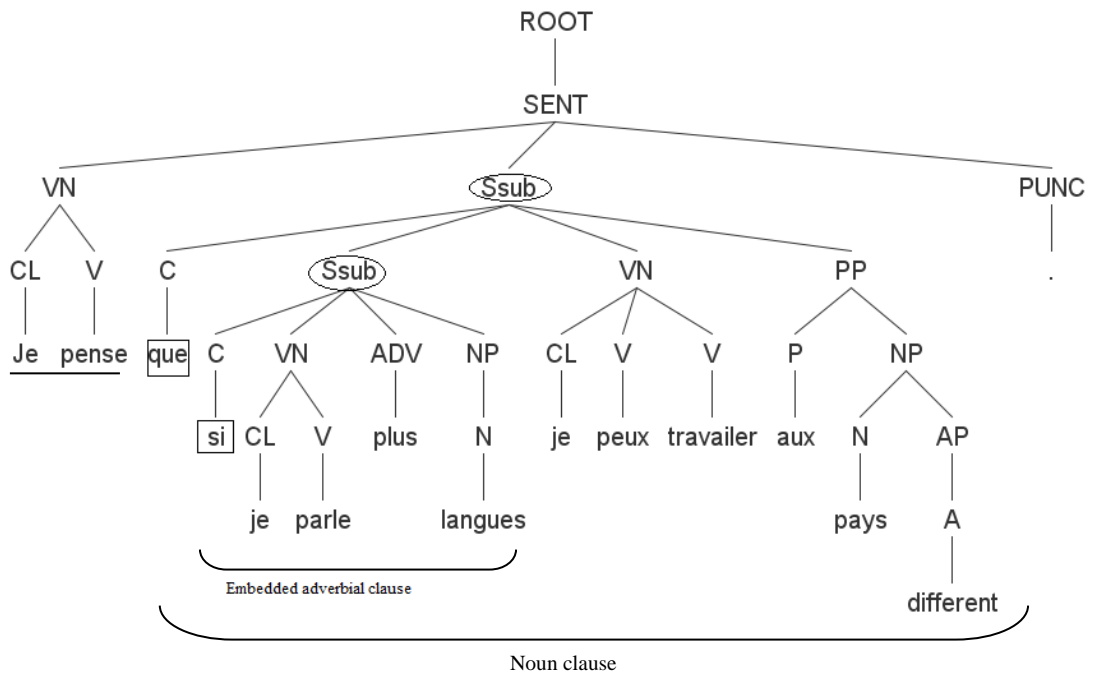
Figure 83 **<N, ADV> clause type in French**

Thus, the two main syntactic structures are: 1) [main clause + subordinate noun clause + subordinate adverbial clause]; 2) [main clause + subordinate noun clause with the embedded adverbial clause].

Complex sentences containing <N, ADJ> sub-clauses

The second most constantly marked complex sentence is the one containing <N, ADJ> clauses. The total number comprises 66 complex sentences (see *Table 27*). This clause type has been used most extensively at levels B – D. In English 9 sentences have been marked at level B; 18 – at level C; 13 – at level D. In French the students at level B have produced this clause type 13 times, while at level C – only twice.

Level A

[1] In conclusion I would like to say ₁[*that all my previous statements can be solved by simply paying more money to people*] ₂[*who work in Latvia*].

[2] Mais la popularité de l'anglais ne signifie pas ₁[**qu**'il *faut ignorer les autres langues etrangéres*] ₂[**qui** *sont moins connues*].

Level B

[3] I belive ₁[***that** somedays all these people* ₂[**who** *leaved Latvia*] *will come back for better life here*].

[4] I think, ₁[***that** is only one reason*], ₂[**why** *Latvians need to stay in Latvia*].

[5] D'abord il est très bien ₁[***que** tout le monde sait cette unique langue*] ₂[**qui** *est apprecié par tous les gens*].

[6] D'abord, je trouve ₁[***qu**'il y a beaucoup de pays*] ₂[**où** *on ne parle pas l'anglais*].

[7] Je crois ₁[***que** c'est une de meilleures façons*] ₂[**comment** *on peut passer le temps libre*].

[8] Je peux aussi dire ₁[***que** dans beaucoup de pays* ₂[**que** *je visitais*] *on ne sait pas l'anglais*].

Level C

[9] I think ₁[*we need to make a new industries*] ₂[**where** *people can work and make engeneering factories in Latvia*].

[10] First of all I think ₁[*we need new goverment*] ₂[**who** *find and distrayed the problems*].

[11] Je pense ₁[***que** il y a beaucoup des mots internationalles*] ₂[**qui** *aide apprendre cettes langues*].

Level D

[12] But I think ₁[***that** anywhay people* ₂[**who** *left Latvia*], *will come back*].

Level E

[13] I think ₁[*first reasons it is give work a young people*] ₂[**who** *finished the Professional school*].

Level F

[14] Im not teeling ₁[**ho** *this peopel* ₂[**ho** *left of Latvia*]].

      At the highest level of language proficiency (A) only one syntactic pattern has been marked (sentences [1], [2]) in both languages, i.e. the noun clause functioning as direct object is subordinated to the main clause and the restrictive relative clause, in which the relative pronoun functions as subject, is subordinated to the noun clause defining either personal or non-personal antecedent.

      At level B the test-takers in both languages have produced the adjectival clauses introduced by different relative pronouns, thus having various functions in relation to their antecedent. For example, sentence [4] contains the restrictive relative clause of reason [*why Latvians need to stay in Latvia*]; sentence [6] is an example of the restrictive relative clause of place [*où on ne parle pas l'anglais*]; in sentence [7] the restrictive relative clause of manner [*comment on peut passer le temps libre*] has been used. Nonetheless, starting from level D to level F only one type of restrictive relative clause, i.e. the one in which the relative pronoun functions as subject, has been produced. Moreover, in English at the lower levels of language proficiency E and F the sentences are rather intelligible.

      As for syntactic structures – in majority of cases multiple subordination has been utilized (see *Figure 85*), i.e. the main clause [*D'abord ile est très bien*] is directly subordinated by the noun clause [*que tout le monde sait cette unique langue*], which is further subordinated by the restrictive relative clause [*qui est apprecié pour tous les gens*].
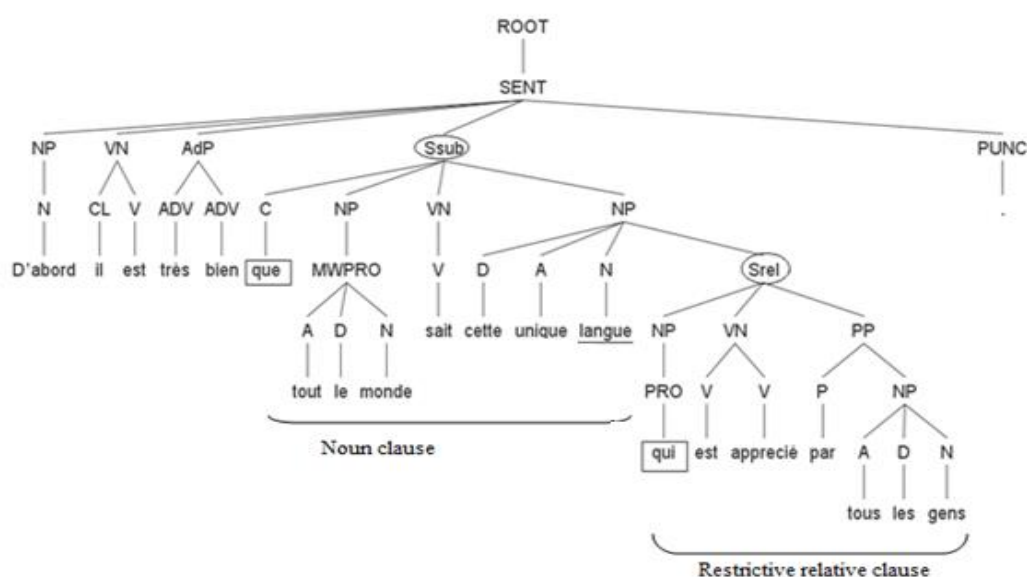
Figure 84 **<N, ADJ> clause type in French**

However, there are complex sentences containing the subordinate clause embedded in another subordinate clause. For example, in *Figure 85* the main clause [*I believe*] is subordinated by the noun clause [*that some day all these people will come back for better life here*] functioning as direct object and containing the embedded restrictive relative clause [*who left Latvia*], which is introduced by the subject relative pronoun *who* and defines the personal antecedent [*people*].
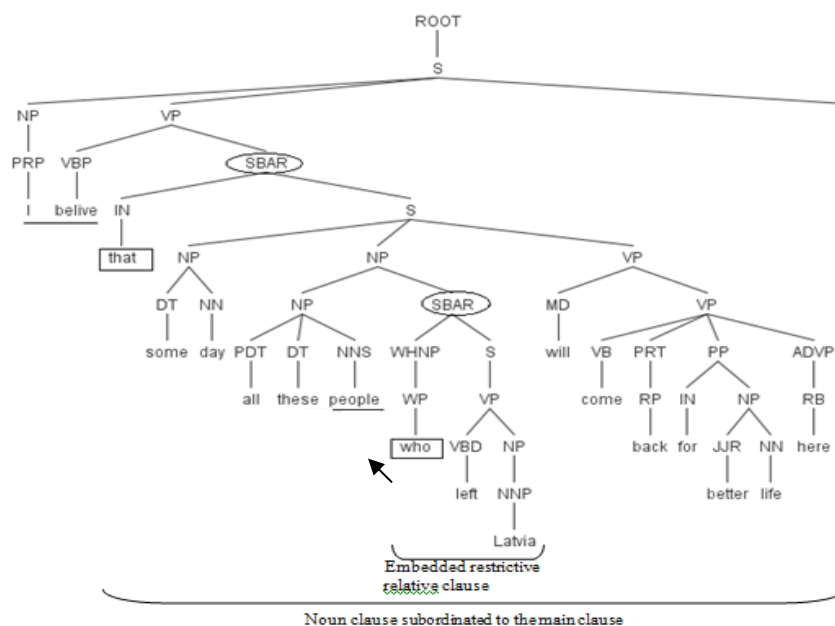
Figure 85 **<N, ADJ> clause type in English**

Consequently, we might distinguish two main syntactic patterns: 1) [main clause + subordinate noun clause + subordinate relative clause]; 2) [main clause + subordinate relative clause embedded in the noun clause].

Complex sentences containing <ADV, ADJ> sub-clauses

The third most often produced variant of clause combinations is a complex sentence containing adverbial and adjectival clauses (see *Table 27*). In English the number of such sentences varies from 3 sentences at levels A and C to 7 sentences at level B; from 7 sentences at level D to only one sentence at levels E and F. The sentence produced at level F is not worth analysing as it has been rewritten from some source. In French the biggest number of clause type <ADV, ADJ> has been marked at level B (12 sentences); only 2 sentences have been produced by the students at level A and 4 sentences – at level C.

Level A

[1] But this method is risky, ₁[***couse*** *you can end up being tricked into paying a* <u>*company*</u>], ₂[***that*** *will eventualy do nothing and leave you in another country without money*].

[2] ₁[***Si** quelqu'un est séduit par un pays exotique et veut y vivre*], ce sera plus utile d'apprendre la langue ₂[***qu**'on parle là-bas*].

[3] On peux voyager plus simplement ₁[***si** on sait la langue parlée du pays*] ₂[***qu**'on visite*].

Level B

[4] And ₁[***if** we will try to make each other happy*] there will be less people ₂[***who** would like to go abroad*].

[5] ₁[***Although** I do not understand those people*], I could figure out few reasons ₂[***why** have they done that kind of action*].

[6] C'est les sentiments inoubliable ₁[***quand** les gens de pays exotics te comprennent et essayent de t'aider avec la langue*] ₂[***que** tu ne connais pas très bien*].

[7] On ne va pas aller au magasin et parler avec la vendeuse en anglais ₁[***si** on est dans la ville native*] ₂[***où** la langue nationale n'est pas anglais*].

[8] ₁[***Si** on voyage*], c'est très bonne possibilité ₂[***comment** fera des connaisances sur la culture de la pays visitée*].

Level C

[9] ₁[***If** thay come back*], that would be very good for our situation ₂[***wich** is very bad right now*].

[10] Je parle l'anglais ₁[***parce que** j'ai un ami*] ₂[***qui** habite à Londres*].

[11] ₁[***Quand** vous avez le talent pour apprendre des langues*] il y a plus de possibiltés ₂[***qui** faire dans votre vie*].

Level D

[12] I don't want to leave Latvia ₁[***because** it is place*] ₂[***where** I was born*].

[13] And they go work abroad ₁[***because** they need money for living, for kids*] ₂[***who** study at school, etc*].

Level E

[14] ₁[***If** Latvia have money*] then he can do the factory for people ₂[***who** don't have work*].

The clause type combination <ADV, ADJ> is more often used by the students at higher levels of language proficiency in both languages. From the above examples it is obvious that the students give preference to combining the adverbial clauses of reason and condition with the restrictive relative clauses containing both relative pronouns and adverbials that specify the antecedent either in the main clause or in the adverbial clause.

In *Figure 87* we can see the consecutive subordination, i.e. the main clause [*Je parle l'anglais*] is subordinated by the adverbial clause of reason [*parce que j'ai un ami*], which

is further subordinated by the restrictive relative clause [*qui habite à Londres*] introduced by the subject relative pronoun *qui* and which defines the personal antecedent [*ami*] of the adverbial clause.
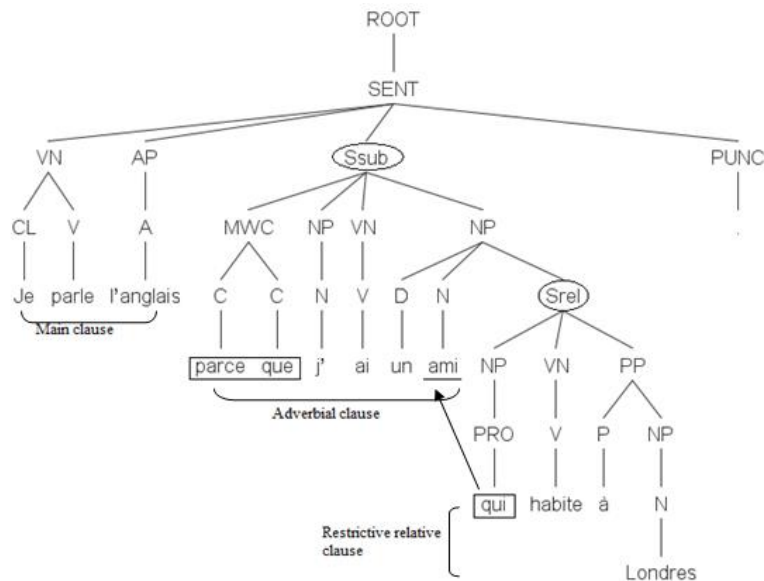
Figure 86 **<ADV, ADJ> clause type in French**

Figure 88 demonstrates another arrangement of subordinate clauses, i.e. the sentence begins with the adverbial clause of concession [*Although I do not understand those people*] followed by the main clause [*I could figure out few reasons*] subordinated by the restrictive relative clause [*why have they done that kind of action*], which defines the non-personal antecedent [*reasons*].
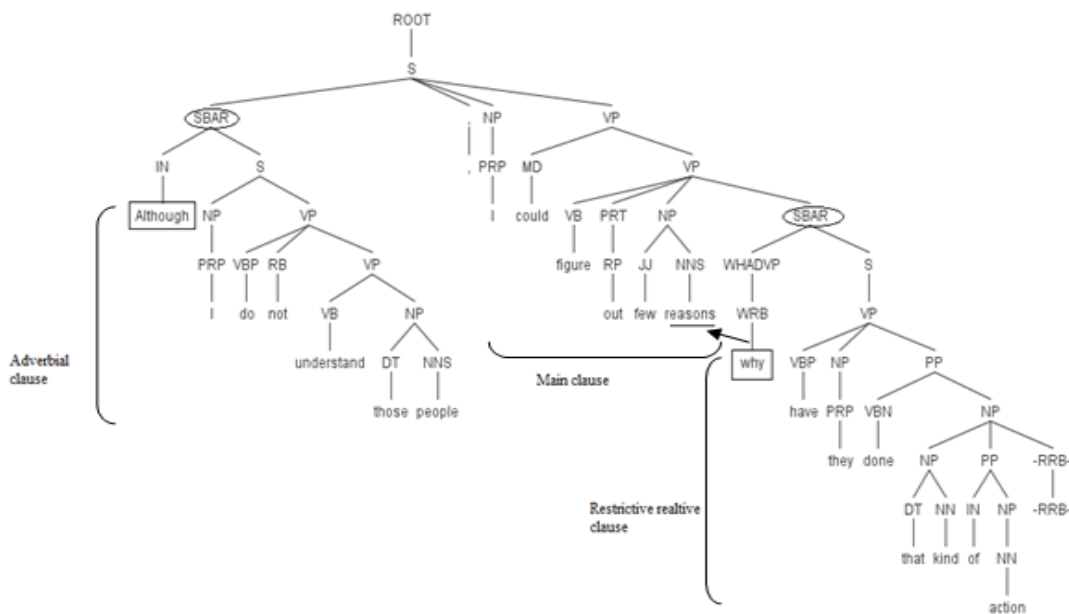
Figure 87 **Structure (b) of <ADV, ADJ> clause type**

In brief, two main syntactic patterns predominate: 1) [main clause + subordinate adverbial clause + subordinate relative clause]; 2) [subordinate adverbial clause + main clause + subordinate relative clause];

Along with already analysed most frequent sub-clauses, there are several more that have been marked in the student written essays, although their number is not that high. Nevertheless, these complex sentences are also noteworthy as they appear almost twenty times at different levels of language proficiency in both languages (see *Table 27*).

Complex sentences containing <N, N> sub-clauses

The number of sentences containing two noun subordinate clauses is considerably lower than the sentences in which the main clause is subordinated by only one noun clause (*Figure 35*). The data in *Table 27* reveal that in English only 12 sentences comprising the main clause and two noun subordinate clauses were produced at levels A – D, whereas in French 6 cases were marked, mainly at level B.

Level A

[1] Summing up all the facts, I think ₁[*it is peoples' free choice to go and do*] ₂[*what ever they want*].

      In sentence [1] the first noun clause [*it is peoples' free choice to go and do*] functions as direct object and is directly subordinated to the main clause [*I think*], while the second noun clause (nominal relative clause) [*what ever they want*] syntactically depends on the first subordinate clause [*it is peoples' free choice to go and do*].

[2] Car je pense aussi ₁[*que ce n'est pas possible*] ₂[*qu'anglais devient la langue de communication mondiale*].

      In sentence [2] the first noun clause [*que ce n'est pas possible*] is the direct object of the main clause [*Car je pense aussi*]; the second noun clause [*qu'anglais devient la langue de communication mondiale*] is subordinated to the first noun clause [*que ce n'est pas possible*] performing the function of the subject.

Level B

[3] The problem is ₁[*that we are said*] ₂[*that it will become better in few years*].
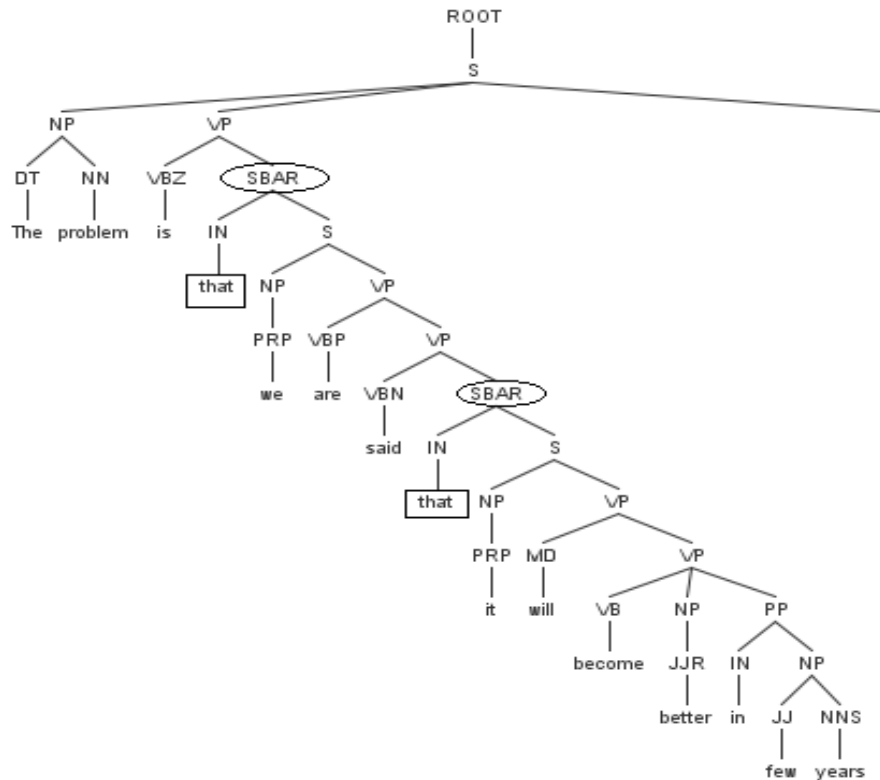
Figure 88 **<N, N> clause type in English**

In sentence [3] the main clause [*The problem is*] is subordinated by the noun clause [*that we are said*] having the function of the subject complement, whereas the second noun clause [*that it will become better in few years*] lies in subordination to the first noun clause performing the function of the object.

[4] Je voudrais dire ₁[*que ce n'est pas vrait*] ₂[*que tout le monde parle anglais*].
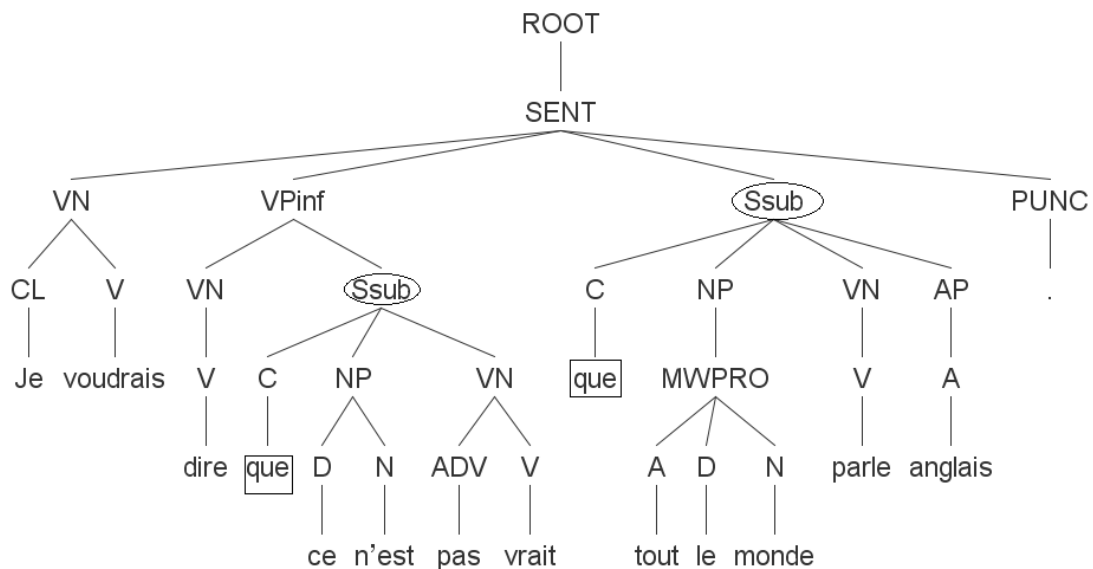
Figure 89 **<N, N> clause type in French**

In French (sentence [4]) the first noun clause [*que ce n'est pas vrait*] functions as direct object and is subordinated to the main clause [*Je voudrais dire*]; the second noun clause [*que tout le monde parle anglais*] is the subject noun clause subordinated to the first noun clause. Thus, the first noun clause is subordinated to the main clause and superordinated to the second noun clause.

Level C
[5] I think ₁[*that it is popular*] ₂[*why people are going to another countries*].

In the example sentence [5] the first subordinate clause [*that it is popular*] is the noun clause functioning as direct object; the second noun clause [*why people are going to another countries*] functions as adjectival complementation and is subordinated to the first noun clause.

Level D
[6] I think ₁[*the bigest problem is in our days*], ₂[*that many people don't have any job her*]!

In the above sentence the main clause [*I think*] is subordinated by the noun clause [*the bigest problem is in our days*] functioning as direct object, while it is rather difficult to decipher whether the second subordinate clause [*that many people don't have any job her*], which is syntactically dependent on the first subordinate clause, is noun clause or relative clause. This is due to the fact that the student has produced a grammatically incorrect sentence, which encumbers the understanding of whether the subordinate clause [that many people don't have any job her] performs the function of the subject complement or specifies the ADVP [our days].

Nonetheless, the analysed example sentences show that the most frequently used noun clauses function as subject, direct object, subject complement, adjectival compliment and nominal relative clause, which was already highlighted previously when examining the complex sentences containing one noun clause. As for the structure, in all the above sentences the first noun clause is subordinated to the main clause, whereas the second noun clause is dependent on the first noun clause (*Figures 89* and *90*).

Complex sentences containing <ADV, ADV> sub-clauses

The number of multi subordinate adverbial clauses, if compared with the frequency of use of mono subordinate adverbial clauses, is remarkably low, i.e. only 10 sentences were

marked in the essays produced by the test-takers in English and 11 sentences – in French (*see Table 27*). Moreover, this sub-clause pattern does not appear at the highest level of language proficiency (only one sentence has been marked at level A in French), though their frequency of use increases at level B comprising 13 sentences, which validates the assertion that adverbial clauses are the syntactic feature characteristic of levels B – D.

The following qualitative analysis of multi adverbial clauses reveals the different functions of subordinate clauses in relation to the main clause as well as their placement within a sentence in respect to the main clause and to each other.

Level A

[1] 1[*Si on veut ètudier la culture*] on doit étudier la langue et autrement 2[*parce que elles sont liés toujours*].
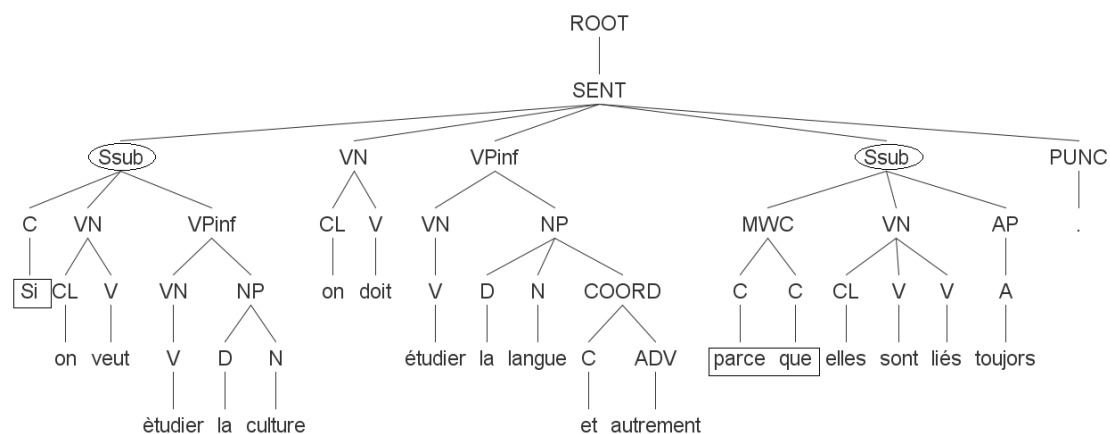
Figure 90 **<ADV, ADV> clause type in French**

The main clause [on doit étudier la langue et autrement] occupies the central position in the sentence as: 1) the direct subordinate conditional clause [*Si on veut ètudier la culture*] introduced by the subordinator *if* stands at the initial position in relation to the main clause performing the function of an adjunct; 2) the subordinate adverbial clause of reason [*parce que elles sont liés toujours*] comes after the main clause explaining why it is important to learn languages. Thus, the sentence contains two adverbial subordinate clauses – the conditional clause and the clause of reason.

Level B

[2] 1[*If I could make some diference*], I would change our press atitude, 2[*because it is too negative*].
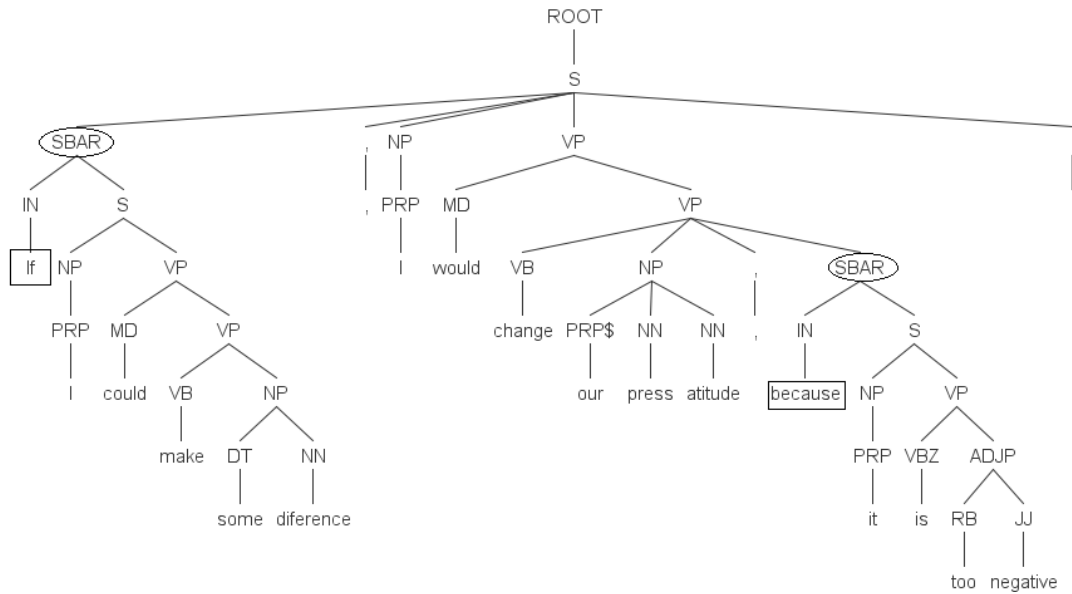
Figure 91 **<ADV, ADV> clause type in English**

In English the structure of the sentence [2] containing the conditional clause and the clause of reason is similar to sentence [1] in French, i.e. the conditional clause [*If I could make some diference*] is placed at the initial position of the main clause [I would change our press atitude], which is followed by the clause of reason [*because it is too negative*].

[3] Cela nous aide ₁[*quand nous allions quelque part en etranger*], ₂[*parce que l'anglais est actuellement la langue de communication mondiale*].
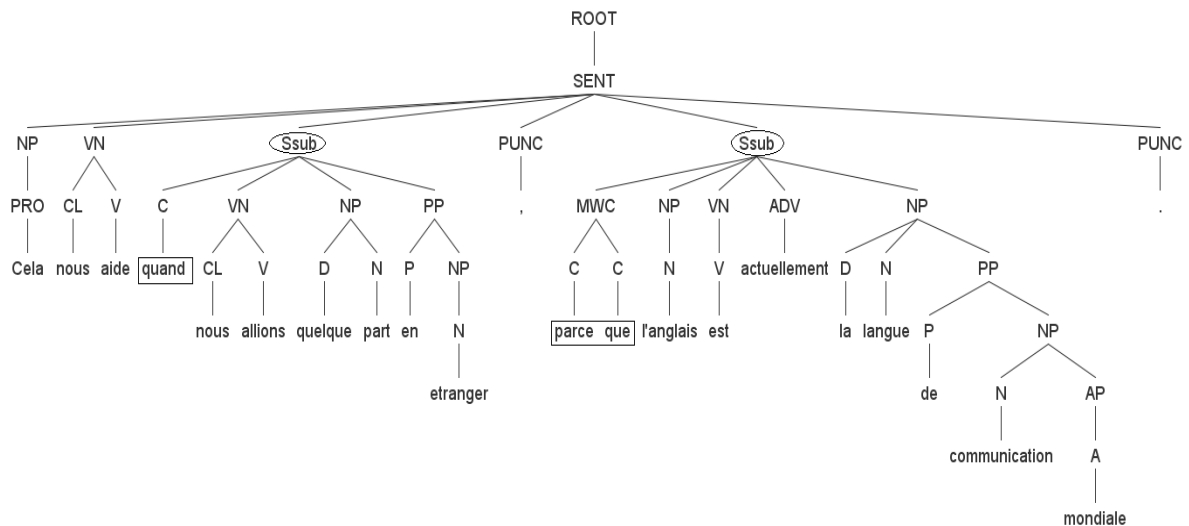
Figure 92 **<ADV, ADV> clause type (a) in French**

Sentence [3] begins with the main clause [*Cela nous aide*] followed the subordinate clause of time [*nous allions quelque part en etranger*], which in its turn is subordinated by the

clause of reason [*parce que l'anglais est actuellement la langue de communication mondiale*].

Level C
[4] Parce-que tu comprend plus de choses ₁[*que les autres*] ₂[*si tu vien aux autres pays*].
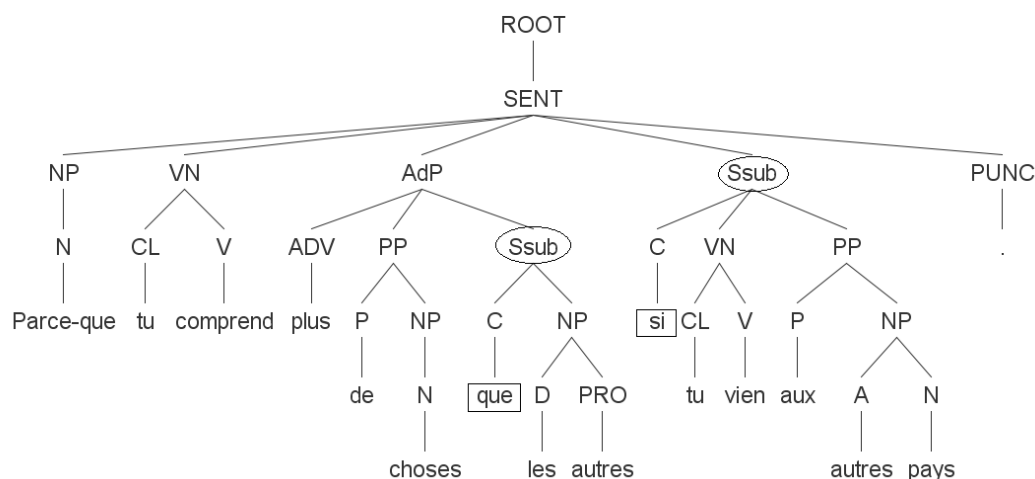
Figure 93 **<ADV, ADV> clause type (b) in French**

Sentence [4] begins with the main clause [*Parce-que tu comprend plus de choses*] that contains the embedded clause of comparison [*que les autres*] and is subordinated by the adverbial clause of condition [*si tu vien aux autres pays*]. Stylistically the sentence is incorrect as the main clause starts with the subordinator *parce que*.

Level D
[4] ₁[*If man really want something*], than he do it, ₂[*because he find a possibility*].

Sentence [4] has the same structure as sentences [1] and [2], i.e. the clause of condition [*If man really want something*] is positioned at the beginning of the main clause [*than he do it*], while the clause of reason [*because he find a possibility*] follows the main clause.

The analysed sentences show that the students prefer to use mainly the clauses of condition, time and reason, which was already found previously when examining mono sub-clauses. As for the sentence structures, in majority of cases: 1) the subordinate clause occupies the initial and final position in a sentence in relation to the main clause; 2) the main clause contains two subordinate clauses, where the second sub-clause is immediately superordinate to the first sub-clause which is subordinate to the main clause.

Complex sentences containing <ADJ, ADJ > sub-clauses

There were 26 sentences containing two adjectival subordinate clauses marked in the student written corpora. This clause type was most frequently used by the students at level B in both languages: 7 sentences in English and 6 – in French. At all other levels of language proficiency their numbers range from 1 to 3. Moreover, at lower levels E and F no sentences consisting of two adjectival subordinate clauses were marked, which was also the case when analysing the complex sentences containing only one adjectival clause.

Level A
[1] I myself want to study at a highschool ₁[***that** gives opportunity for an exchange*
   *program, preferably one] ₂[**that** involves Germany].*

[2] There are a lot of people ₁[**who** *specializes in spheres] ₂[**that** are just not needen in*
   Latvia].

   In sentence [1] the main clause [*I myself want to study at a highschool*] is subordinated by the restrictive relative clause [*that gives opportunity for an exchange program, preferably one*] that is in its turn subordinated by another restrictive relative clause [*that involves Germany*]. The subject relative pronoun *that* defines the non-personal antecedent *highschool* in the main clause and *one* in the subordinate clause.

   In sentence [2] the main clause [*There are a lot of people*] is defined by the restricitve relative clause [*who specializes in spheres*] that is further subordinated by the restrictive relative clause [*that are just not needen in Latvia*]. The subject relative pronoun *who* defines the noun head [*people*] in the main clause, while another subject relative pronoun *that* defines the non-personal antecedent [*spheres*].

[3] Il y a plein de pays ₁[***où** l'anglais n'est pas la langue unique] ₂[**dont** on parle].*

   In the complex sentence [3] the main clause [*Il y a plein de pays*], which does not contain the subject, is subordinated by the restrictive relative clause [*où l'anglais n'est pas la langue unique*] whose adverbial *où* specifies the non-personal antecedent [*pays*] of the main clause, whereas another subordinate restrictive relative clause [*dont on parle*] with its relative adverb *dont* defines the non-personal antecedent [*langue*] in the preceeding subordinate clause.

[4] À cause de mondialisation on commence à avoir le concepte ₁[***que** l'anglais est la seule*
   *langue necessaire] et ₂[**que** les autres sont inutiles].*

Sentence [4] is an example of parallel subordinantion, where the main clause [*À cause de mondialisation on commence à avoir le concepte*] is followed by two subordinate clauses [*que l'anglais est la seule langue necessaire*] and [*que les autres sont inutiles*] joined by the coordinating conjunction *et.* Thus, both restrictive relative clauses define the non-personal antecedent [*le concept*] in the main clause, and the relative pronoun *que* has the function of object.

Level B

[5] Why stay in a <u>coutry</u> ₁[*were people* ₂[*who are in charge*] *lies and cheats, and takes all money for themselves*]?
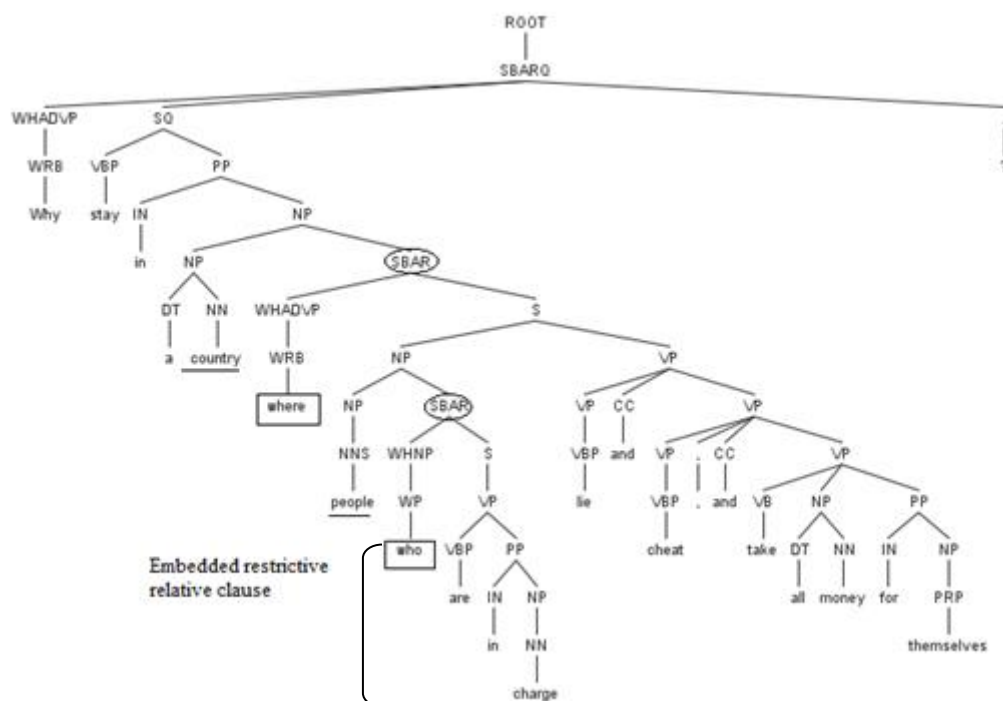
Figure 94 **<ADJ, ADJ> clause type in English**

In the above sentence [5] the main clause [*Why stay in a coutry*] is subordinated by the restrictive relative clause [*were people lies and cheats, and takes all money for themselves*] that contains the embedded restrictive relative clause [*who are in charge*]. The adverbial *where* specifies the non-personal antecedent [*country*] of the main clause, while the subject relative pronoun *who* defines the personal antecedent [*people*] in the subordinate clause [*were people lies and cheats, and takes all money for themselves*].

[6] A mon avis la mondialité de la lange anglaise ne change rien pour <u>ce</u> ₁[*qui aiment les languages*] et ₂[*qui veulent les apprendre*].

Sentence [6] is another example where the main clause [*A mon avis la mondialité de la lange anglaise ne change rien pour ce*] is subordinated by two restrictive relative clauses [*qui aiment les languages*] and [*qui veulent les apprendre*] linked by the coordinating conjunction *et* (see also sentence [4]). However, in this case both relative clauses are introduced by the subject relative pronoun *qui* and define the personal antecedent *ce* of the main clause.

Level C

[7] The only opinion ₁[***wich** is left*] is to look for life in other countries ₂[***where** the situation is better*].

The complex sentence [7] contains one embedded restrictive relative clause [*wich is left*], introduced by the subject relative pronoun *which* specifying the non-personal antecedent [*opinion*], and one subordinate restrictive relative clause [*where the situation is better*] introduced by the adverbial *where* which defines the noun head [*countries*] of the main clause.

[8] Et vous pouvez faire les ami dans tout les pays, ₁[***où** vous allez*], ₂[***qui** vous aideraient*].
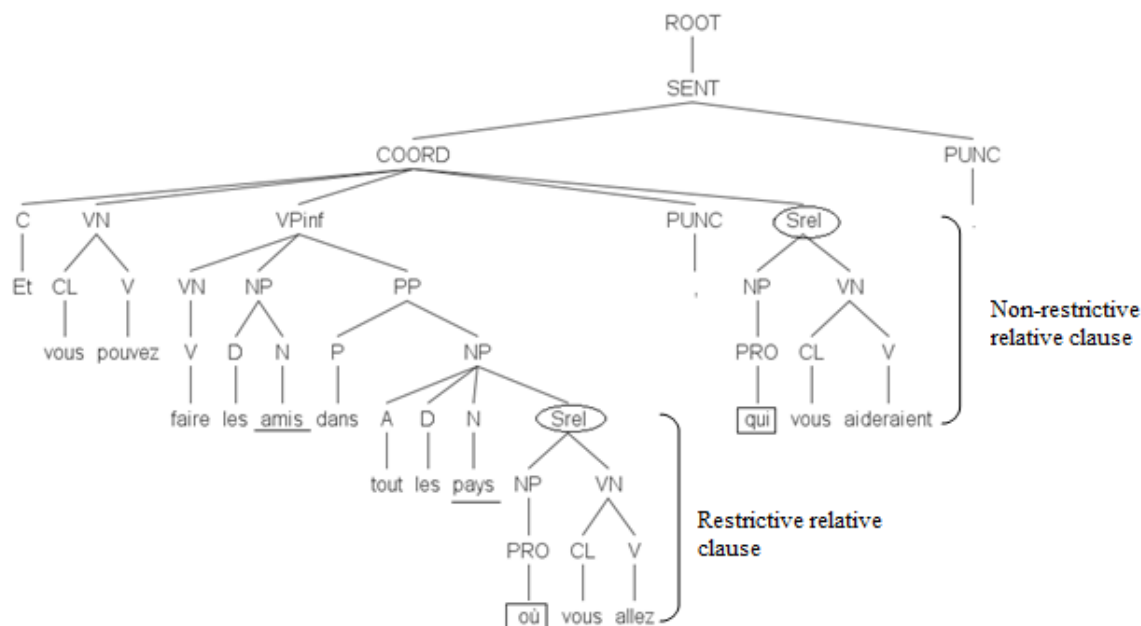
Figure 95 **<ADJ, ADJ> clause type in French**

The above sentence structure is rather unique as the clause linking patterns are unusual. Moreover, it contains both restrictive and non-restrictive relative clause. Thus, the main clause [*Et vous pouvez faire les ami dans tout les pays*] is subordinated by the restrictive relative clause [*où vous allez*] and as if by the non-restrictive relative clause [*qui vous*

187

aideraient]. However, grammatically it would be more appropriate to place the non-restrictive relative clause [*qui vous aideraient*] right after the NP [*les amis*], which it describes, and have an embedded construction within the main clause.

Level D

[9] These people ₁[***who** don't have job*] and ₂[***who** can't find job*] just leave Latvia and go to other cantries.

The complex sentence [9] consists of the main clause [*These people just leave Latvia and go to other countries*] and two embedded restrictive relative clauses [***who** don't have job*], and [***who** can't find job*] linked in coordination, thus making a parallel subordination. Both relative clauses are introduced by the subject relative pronoun *who,* which specifies the personal antecedent [*people*].

It is obvious from the examined multi adjectival sub-clauses that the students in both languages are familiar with different sentence structures (embedding and subordination). However, in English the students prefer using only restrictive relative clauses, which was also observed when analysing the sentences containing one subordinate clause. In French there are at least some cases when the students have used non-restrictive clauses as well. Moreover, at level A (sentence [3]) the restrictive relative clause introduced by the adverbial *dont*, which is considered to be one of the most complicated syntactic patterns, was marked. Thus, it might be claimed that this structure is criterial for higher levels of language proficiency.

Complex sentences containing <ADV, N > sub-clauses

Complex sentences composed of the main clause subordinated by the adverbial and noun clauses have been used 19 times, particularly at level B both in French (5 sentences) and English (7 sentences).

Level B

[1] ₁[***When** they finish their trip*] they realise ₂[***that** they haven't been to Latvia for a long time*].

[2] More and more people are leaving Latvia ₁[***because** they think*] ₂[***that** in other countries they will have better opportunities and life*].

[3] In my opinion, it is good to travel, ₁[***if** I know*] ₂[***when** I will be back*].

[4] Je voudrais apprendre encore quelques langues, ₁[***parce que** je crois*] ₂[***que** ça me fera plus intélligente et capable à communiquer au presque tout le monde*].

[5] ₁[***Quand** vous allez en voyage, par exemple, en Afrique*] vous voulez voir tout ce ₂[***qui** est intéressant*].

[6] 1[***Même si*** *l'anglais est la deuxiéme langue le plus parlé de monde*], je crois 2[***que la langue française, russe, allemand et espagnol sont vraiment importantes***].

Level C
[7] People after leave their countries 1[***because*** *they actually don't know*] 2[***what*** *they want in their lives*].

[8] C'est utile pour chaque personne, 1[***parce que*** *je pense*] 2[***que les langues sont comme les mains***].

[9] 1[***Si nous savons quelque langue***], nous pouvons voyage, lit, et regarde 2[***coment*** *vivre otres persones en diferantes cultures*].

Level D
[10] 1[***When*** *you leave at school*] you understand 2[***that*** *the time of your independence life and the beginning of a for more serious examination of you abilities and character has come*].

In majority of sentences the students have used the adverbial clauses of condition ([1],[3],[5],[6],[9]), reason (sentences [2],[4],[7],[8]) and time (sentences [1],[5],[10]) introduced by the subordinators *if, even if, when, because* in English and *même si, si, quand, parce que* in French. In the provided examples the adverbial clauses appear either at the beginning of the main clause or are directly subordinated to the main clause. As for the noun clauses, the ones having the function of direct object predominate. In addition, they are subordinated either to the main clause or the adverbial clause. For example, in *Figure 96* the adverbial clause of time [*When they finish their trip*] occupies the initial position in the sentence followed by the main clause [*they realize*] which is subordinated by the noun clause [*that they haven't been to Latvia for a long time*] functioning as direct object of the main clause. Thus, the main clause has been placed in the medial position between two subordinate clauses.
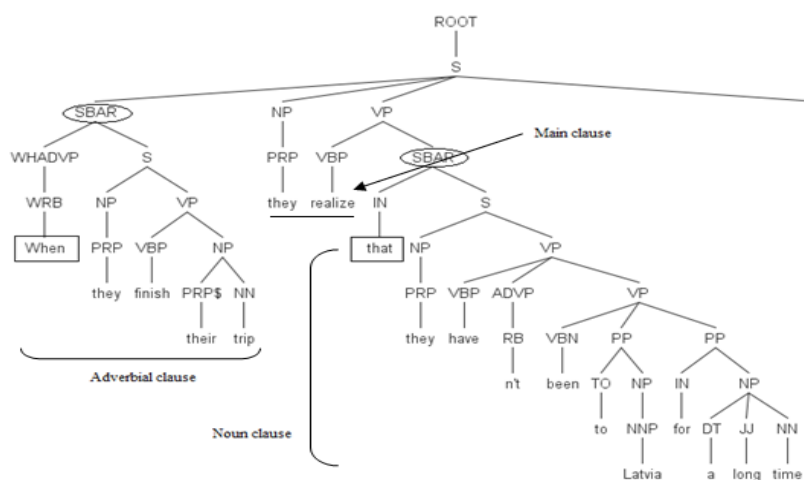
Figure 96 **<ADV, N> clause type in English**

189

In *Figure 97* the traditional arrangement of the two subordinate clauses in relation to the main clause can be observed, i.e. the sentences is introduced by the main clause [*C'est utile pour chaque personne*] followed by the adverbial clause of reason [*parce que je pense*] subordinate to the noun clause [*que les langues sont comme les mains*] having the function of direct object.
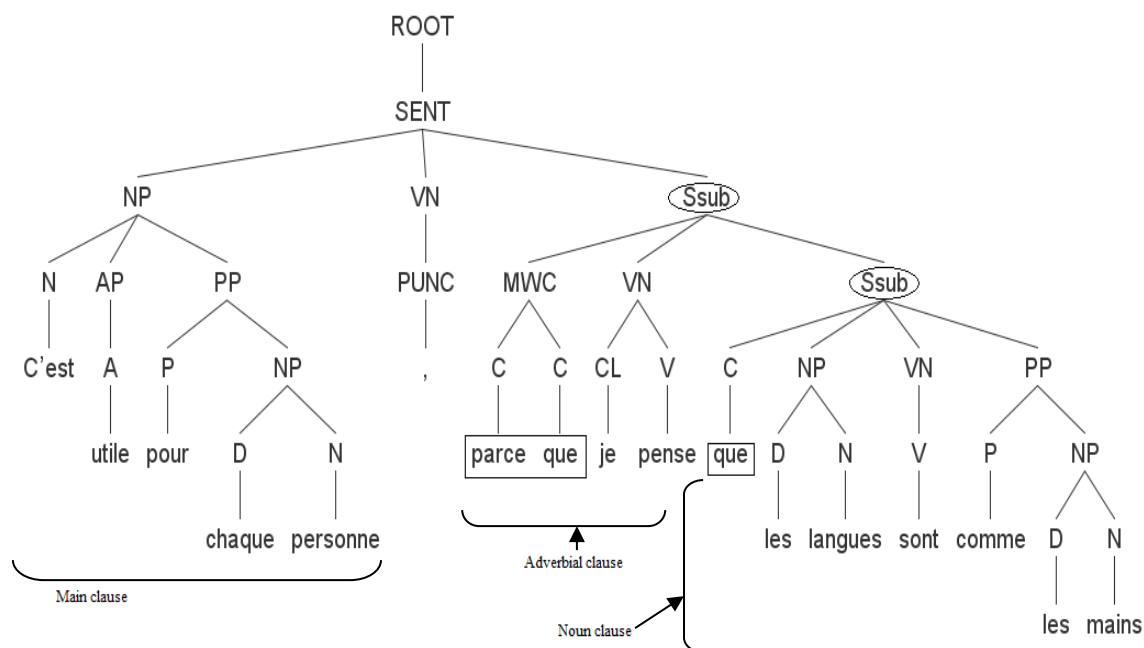
Figure 97 **<ADV, N> clause type in French**

To conclude, two main syntactic patterns are typical of the aforementioned complex sentences containing adverbial and noun clauses: 1) [subordinate adverbial clause + main clause + subordinate noun clause]; 2) [main clause + subordinate adverbial clause + subordinate noun clause].

Complex sentences containing <ADJ, N > sub-clauses

Complex sentences in which the adjectival clause is subordinated by the noun clause has been marked almost 3 times less (23 times) than those complex sentences in which the noun clause preceeds the adverbial clause. Apart from that, in English the students use this clause type at all levels of language proficiency, while in French it has been marked only at levels A and B.

Level A

[1] The only thing 1[( ) *I would like*] is 2[***that people would come back after some years***].

[2] Une autre raison ₁[***pour laquelle** je considère*] ₂[***qu****'il faut apprendre pas seulement l'anglais*], c'est à cause de l'éducation.

[3] Les personnes ₁[***qui** ont choisi l'espagnol comme leur troisiéme langue étrangère*] dit ₂[***que** c'est similair au français ou italien*].

Level B

[4] People, ₁[***who** have left Latvia*] say ₂[***that** they are better paid in other countries*].

[5] One of the main resons ₁[***why** people have left Latvia during the last few years*] is ₂[***that** they are better paid in other countries*].

[6] Les hommes ₁[***qui** vivent dans l'étrangeres*] aime ₂[***quand** le touristes parlent dans leur lange*].

[7] Grâce aux pays puissants et le developpement des médias ₁[***qui** assurent la circulation rapide de l'information*], on peut considéres ₂[***que** l'anglais est la langue de communication mondiale*].

Level C
[8] Other reasons ₁[***why** citizens leave country*] is ₂[***that** in others countrys you have more helps for your kids from goverment*].

[9] So, the people, ₁[***who** are leaving our country*], think, ₂[***thet** there they can live all their life, growing their children there and be more garnteed in other countries*].

Level D
[10] Other reason ₁[***why** we leave our country*] is ₂[***that** we can't feel happy here*].

Level E
[11] And people ₁[***who** don't have so mutch many to by someting*] you been pore, unlacky and you never get's ₂[***what** you want*].

Level F
[12] I reoson ₁[***why** people hove left Latvia*] – ₂[*people don't love country*].

In English the students at all levels tend to reproduce the same syntactic structures given in the task rubric (sentence [5]), which contains the main clause [*One of the main reasons is*] with the embedded adjectival clause [*why people have left Latvia during the last few years*] followed by the subordinate noun clause [*that they are better paid in other countries*]. Parallels can be drawn with the findings when different types of adjectival clauses were analysed in the complex sentences with mono subordination, since the students had reproduced the same syntactic patterns. As a result, the clauses specifying the antecedent *reason* were not counted and further examined. Though in majority of cases the adverbial *why* introduces the restrictive relative clause defining the non-personal antecedent *reason* in the complex sentences containing multi sub-clauses, the students have

utilised also other restrictive relative clause types introduced by the relative pronouns *who, that, zero pronoun* in English and *qui, que* in French, thus performing the function of subject [3], [4], [6], [7], [9], [11] and object [1].

As for the noun clauses, they are mainly subordinate declarative clauses introduced by the pronoun *that* in English and *que* in French (in sentence [6] the conjunction *quand* has been used), and functioning as direct object or subject complement.

There is one main syntactic structure prevailing in the given examples, which is reflected in *Figure 98*. The sentence consists of the main clause [*Les homme aime*] which contains the embedded restrictive relative clause [*qui vivent dans l'étrangeres*] introduced by the subject relative pronoun *qui*. The main clause is further subordinated by the noun clause [*quand le touristes parlent dans leur lange*].
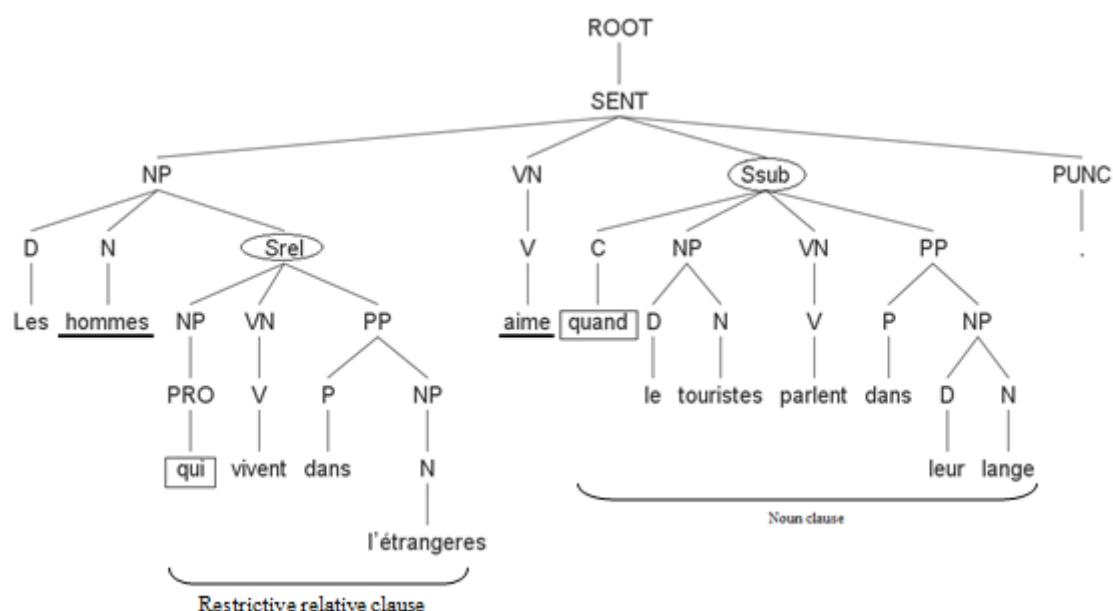
Figure 98 **<ADJ, N> clause type in French**

Thus, we can distinguish the syntactic pattern [main clause containing an embedded restrictive relative clause + subordinate noun clause].

Complex sentences containing <ADJ, ADV > sub-clauses

The <ADJ, ADV> clause combination has been produced by the students 25 times: 19 sentences have been marked in English and 5 sentences – in French written performance at levels B and C (see *Table 27*).

Level A
[1] On the one hand, I, of course, understand people ₁[***who** are now leaving*], ₂[***because** the*

current situation is really bad].

[2] Another <u>reason</u> ₁[(   ) *people are leaving*] is ₂[***because** they can't even find a job here*].

Level B

[3] Personally I know a lot of <u>students</u> ₁[***who** will leave Latvia after school graduation*], ₂[***because** they will study abroad*].

[4] In <u>Latvia</u> ₁[***which** is the country with one of the lowest average salary in Europe*], higher education is expensive, specially since year 2009 ₂[***when** the price for education almost doubled*].

[5] Moi, je suis une <u>étudiante</u> ₁[***qui** aime passionnement les langues étrangères*] ₂[***autant que** tous le pays autour des monde*].

Level C

[6] Now, <u>childrens</u> ₁[***who** grawing up*], will not Latvias patriots, ₂[***because** goverment didn't do enything for their*].

[7] L'utilisation d'une langue dans tout la monde est une <u>signe</u> de „globalisation" ₁[***qui** n'est pas un bon process*] ₂[***parce que** les economies locales ne functionne jamais*].

Level D

[8] I agree with <u>people</u>, ₁[***who** go abroad*], ₂[***because** Latvia is too little salary*].

Level E

[9] Another <u>reason</u> ₁[***why** people leave Latvia*] is ₂[that, ***because** they want to travel and look out for the other city, country*].

Level F

[10] Very much <u>people</u> ₁[***who** live in Latvia*] now ride in other countries ₂[***becose** in other countries life they children, parents*].

The adjectival clauses produced by the students of English and French are almost all restrictive relative clauses introduced by the relative pronouns *who* and *qui* (with the exception when the relative clause begins with the adverbial *why*) having the function of subject, thus defining the personal and non-personal antecedent in the main clause. In sentence [2] the adverbial *why* has been omitted to avoid tautology, which is a noteworthy linguistic competence. Sentences [4] and [7] contain non-restrictive relative clauses. In sentence [4] the relative clause describes the antecedent *Latvia* while in sentence [7] the non-restrictive relative clause [*qui n'est pas un bon process*] is more like a comment. However, in sentence [4] no comma has been put before the relative pronoun *which*, and in sentence [7] the restrictive relative clause has not been separated by commas, which means that students do not pay much attention to deciphering between restrictive and non-restrictive relative clauses.

The main type of adverbial clauses used in the above sentences is the adverbial clause of reason, while sentence [4] contains the adverbial clause of time and sentence [5] – the adverbial clause of comparison, which has been a rarity in the marked sentences.

As for the position of the subordinate clauses in relation to the main clause, we may discern two cases. In the first one (*Figure 99*), the restrictive relative clause [*who are now leaving*] is subordinated to the main clause [*On the one hand, I, of course, understand people*] and superordinate to the adverbial clause of reason [*because the current situation is really bad*].
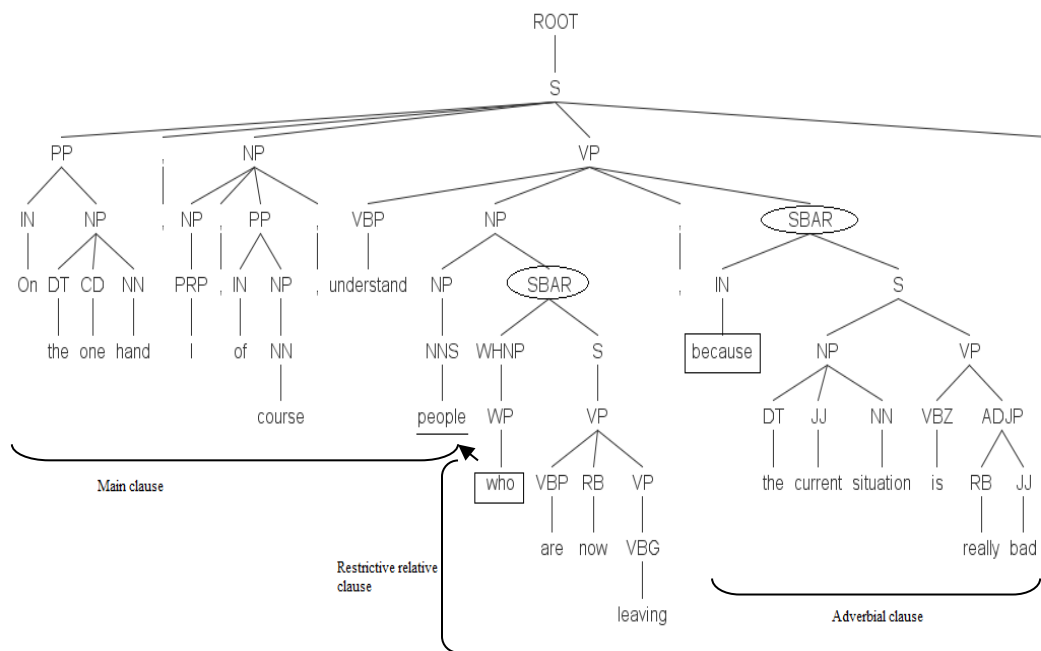
Figure 99 <ADJ, ADV> clause type (a) in English

In the second case (*Figure 100*), the embedded restrictive relative clause [*who grawing up*] is introduced by the subject relative pronoun *who* and defines the personal antecedent [childrens] of the main clause [*Now childrens will not Latvias patriots*], which is superordinate to the adverbial clause of reason [*because government didn't do anything for their*].

Figure 100 **<ADJ, ADV> clause type (b) in English**

It is evident that the test-takers are familiar with the following syntactic patterns: 1) [main clause + subordinate adjectival clause + subordinate adverbial clause]; 2) [main clause containing an embedded relative clause + subordinate adverbial clause].

Having examined the main types of multi subordinate clause combinations within complex sentences, the author of the present research concludes that the students in both languages produce the same syntactic patterns marked in the complex sentences with one subordination, except for some cases in French, when the students have utilised a more complex patterns, such as the adverbial clause introduced by the adverbial *même si* at level B, and the adjectival clauses introduced by the relative pronoun *pour laquelle* and the adverbial *dont* at level A. This allows us to assert that the rare occurrence of more complex structures might be used as distinguishing features for attributing a higher level of language proficiency at the examination, which complies with the claim that 'the more complex a structure, the less frequently it is used' (Hawkins and Filipovič, 2012: 90).

As to the placement of subordinate clauses in a sentence, several patterns have been distinguished:

→ [main clause + subordinate noun clause + subordinate adverbial clause];

→ [main clause + subordinate noun clause with embedded adverbial clause];

→ [main clause + subordinate noun clause + subordinate relative clause];

195

→ [main clause + subordinate noun clause with embedded relative clause];

→ [main clause + subordinate adverbial clause + subordinate relative clause];

→ [main clause + subordinate adverbial clause + subordinate adverbial clause];

→ [superordinate adverbial clause + main clause + subordinate relative clause];

→ [superordinate adverbial clause + main clause + subordinate noun clause];

→ [main clause + subordinate adverbial clause + subordinate noun clause];

→ [main clause containing an embedded relative clause + subordinate noun clause];

→ [main clause + subordinate relative clause + subordinate adverbial clause];

→ [main clause containing an embedded relative clause + subordinate adverbial clause].

It is evident that the students produce the same syntactic structures irrespective of their foreign language (English or French), which had to be proved by carrying out the contrastive studies of different clause types marked in the EFLWT corpus. However, the complexity and accuracy of the structures differs according to the level of language proficiency. This phenomenon might be explained by applying Pienemann's Processability Theory, according to which the most important aspect in acquiring a language is not the proximity of L1 and L2, but the individual's capacity of processing different syntactic patterns; though, we cannot deny the fact that if the particular syntactic structure is not linguistically marked in the test-taker L1, the acquisition of this structure might be relatively more difficult.

Before determining the syntactic features criterial for each language proficiency level, the present research data are to be compared with the grammatical criterial features elaborated within the framework of the English Profile Programme (EPP). As no such research has been accomplished in French, the data of the present research will be compared with the syntactic criterial features of each language proficiency level in English, since the aforementioned findings substantiate the hypothesis drawn at the beginning of the present study that the same criterial features may be used when assessing student grammatical competence cross-linguistically. In order to have a clear understanding of how the language proficiency levels (A − F) used in assessing the test-taker linguistic performance in Latvia till 2012 comply with the ones defined by the CEFR (A1 − C1), the subsequent analysis will contain both of them, e.g. level D will be related to the CEFR level A2 based upon the test relation research effected by Kalnbērziņa (2007)

and Kunda (2011). Moreover, the EP syntactic criterial feature for each level will be compared with the present research findings in English and French.

Level A2/D

1) **[Verb + subordinate clause with or without *that*]**, which comprises one main clause containing a NP and a V and a subordinate complement clause with a finite Verb.

 This syntactic structure has been extensively used (most frequently if compared with other clause types) by the students at all levels of language proficiency (A/C1 – F/A1). However, at lower levels E/A1 and F/A1 the syntactic structures produced are rather short, primitive, repetitive and in many cases illigible. At level D/A2 the sentences become longer consisting of different clause elements, though a lot of the sentences are ungrammatical.

Level B1/C

2) **It + Verb + subordinate clause with or without *that***

 [It extraposition with finite clauses]

 This clause type has been marked at levels A/C1, B/B2, C/B1 both in English and French. It is obvious that foreign language learners are familiar with this structure. However, the frequency of its use is rather low, ranging from 1 to 6 cases per level.

3) **Verb + Prepositional Phrase + subordinate clause with or without *that***
 [Verbs with a PP plus finite complement clause, NP-V-PP-S]

 This syntactic structure is typical of the learners of English as it does not exist in French. However, none of the students has produced such a complex sentence.

4) **Relative clauses with *whose***

 [Relative clauses formed on a genetive position]

 None of the students in either language has used the relative clause beginning with *whose.*

5) **Wh-word + Noun Phrase + Verb clauses used as subject or object**
 [**Pseudoclefts type (i)** WH-NP-VP]

 This syntactic structure has not been produced by the students in either language.

6) **Indirect WH-questions**

 The syntactic structure where the interrogative subordinate clause functions as direct object has been used by the students at levels C/B1 and D/A2 in English only, e.g. After 12th form students don't know *[**where** they would like to study]* (level D/A2). In

the example sentence the subordinate clause begins with a *wh*-word (where) followed by the noun [*they*] and the finite Verb [*would like*].

Level B2/B

7) **Wh-word + Verb clauses used as subject or object**

   [**Pseudoclefts type (ii)** WH- VP]

   It appears rarely in student written performance. What is more, it has been marked once at level A/C1 and once at level D/A2 in English:

   > That is [*what* *drives young, powerful people*] *into the unknown.*] (level A/C1)

   > [*What* *compels the people of old faith, having deep roots in Latvia*], they do not know the languages of foreigners, to abandon their homes? (level D/A2)

   In French no such structure was marked in the FLWT sub-corpus.

8) **Verb + object + subordinate clause with or without that**

   [**Verbs with a NP plus finite complement clause, NP-V-NP-S**]

   This structure has been utilised very seldom, i.e. it has been marked only once at levels B/B2 and C/B1 in the essays produced by the learners of French. Moreover, in French, if the object is expressed by a pronoun, it appears:

   a) before the verb, e.g. Pour conclure, je voudrais <u>vous</u> **dire** *[que la langue c'est la communication entre plusieurs hommes]*. Thus, the syntactic structure in this sentence is more complex – NP-V-NP-V-S.

   b) after the verb if preceded by a preposition, e.g. Je peux **recommender** <u>à chaqun</u> *[qu'il doit apprendre des langues étrangères]*. In this case the syntactic structure is NP-V-V-PP-NP-S

   When analyzing the clauses, one should be aware of the structural differences in both languages, which is natural, as the main principles of syntactic structure generation are universal both in English and French, whereas certain parameters enforce irregularity, e.g. the placement of the adverb or clitic pronoun.

   The overall results make the author of the present research conclude that the students learning English and French as a foreign language tend to produce mainly uncomplicated syntactic structures, which are rather repetitive. What is more vital is the fact that the research results substantiate the assumption that the language competence (syntax) can be assessed by applying the same criteria in both languages as they almost comply with the criterial features enumerated in the English Profile:

1) At level A2/D learners are to produce more complex sentences, i.e. a sentence containing a noun clause functioning as direct object introduced or not by the subordinator *that/que*. Thus, this level is characterized by relatively simple syntactic structures, which has been proved by the research results showing that this clause type has been used by the students of English most efficiently at level D. In French this clause type has been utilized excessively also at levels B and C.

2) At level B1/C learners already make use of simple relative clauses introduced by relative pronouns *who* and *which (que)*. What they should start producing at this level is the relative clause introduced by *whose/dont*. However, the relative clauses marked in learner essays in both languages do not contain the structure [whose + NP + VP].

3) At level B2/B the number of complex sentences considerably increases, which has also been proved by the research data, namely, the number of complex sentences marked in student essays increases towards level B in English (reaching the highest number – 37%), whereas in French the highest number of complex sentences has been observed already at level C comprising 43% and then falling down to 39% at level B.

   As for the structures, this level of language proficiency is caharacterised by the use of a sentence beginning with *wh*-clause functioning as the subject of the main clause. However, the frequency of this structure is extremely low – it has been used only twice in English, which might be explained by the fact that this structure is not frequent in the native language as well. Another structure that characterizes this level is the main clause containing a direct object followed by a subordinate complement clause with *that/que* (or without that in English). This structure as well as the previous one appears in few cases, this time only in French.

4) At level C1/A there are no particular grammatical criteria as learners are supposed to use 'structures mastered at earlier levels but with a much wider range of vocabulary and in more accurate ways' (English Profile, 2011:15).

   If we compare the criterial features elaborated by the English Profile team members with the analytic marking scale provided by the NCE for assessing the student written performance based on the CEFR descriptors of grammatical accuracy, then we see (Appendix 8) that:

1) At level A2 students are expected to use only *simple structures* in which errors predominate. This does not conform with what has been stated in the English Profile according to which more *complex sentences* are to be produced.

2) At level B1 still simple structures with errors predominate in the NCE scale while in the CEFR (Council of Europe, 2001: 114) there is 'generally good control though with noticeable mother tongue influence'. In the EP the learners are supposed to make use of different relative clauses, which does not comply with the other two level descriptors.

3) At the beginning of level B2 the students in Latvia are to use both simple and complex sentence forms and only towards a higher level of proficiency within level B2 a relatively high degree of grammatical control is expected, which corresponds to the EP level B2 descriptors, i.e. the number of complex sentences considerably increases.

4) At level C1 both the NCE and the CEFR state that learners are to maintain constantly a high degree of grammatical control of complex sentences, which fully corresponds to the EP.

The contrastive analysis of the aforementioned documents reveals that the descriptors are more alike for the higher levels of language proficiency. What needs to be thought over are the lower level descriptors as they are rather indistinct. Moreover, the author of the present research finds the descriptors to be insufficient to assess the student written performance both in English and French as they do not clearly state (concretely specify) what syntactic structures are to be searched for when attributing a certain level for student written performance. Furthermore, it is noteworthy what Pienemann (1985: 24) states, i.e. 'an identification of simple and frequent grammatical and lexical points can serve as the basis for the teaching of more advanced linguistic features, thereby improving the quality of L2 instruction and learning.' Therefore, for the sake of learners, teachers and assessors the criterial features characteristic for each level (starting with the lower ones) were elaborated (*Tables 28* and *29*) taking into consideration the research results, the CEFR descriptors as well as the EPP grammatical criteria. This would guarantee more objective and more reliable results at the examination that could be related to the CEFR as well as guide towards attaining the higher levels of learner linguistic competence.

Table 28 **Syntactic criterial features in English**

F/A1	**Simple sentences**		
E/A1	**Simple sentences**		
D/A2	**Compound sentences consisting of two or more main clauses, linked by coordinating conjunction.** It's late and I must go home.		
	Complex sentences		
	NOUN CLAUSE	**ADVERBIAL CLAUSE**	**ADJECTIVAL CLAUSE**
	NP + VP + (that) I think *that this number will be smaller in the future.* (know, think, believe, hope, say, tell)	**where** (place) I found the bag *where I always put it.*	
	Indirect questions with *if* I don't know *if I can do that.* He asks if he can go now.	**when** (time) *When they go abroad,* they see a lot of place.	
	After **mind** Do you mind if we go by air? I don't mind if you smoke.	**because** (reason, cause) **+ NP + VP** Nowadays a lot of people go to other countries *because they have some reasons to do it.*	
	Following certain complementing adjectives: **sure, certain** *I'm sure (that) it's very important for our country.*	**If** (condition) **+ NP + VP** if +present / present Nowadays we must work hard *if want to live in a beautiful and rich country.*	
		If + present/future ***If** they continue studying abroad,* there will be opportunities to get a job.	
C/B1	**Complex sentences limited to one subordinate clause with simple structure.**		
	NOUN CLAUSE	**ADVERBIAL CLAUSE**	**ADJECTIVAL CLAUSE**
	It (extraposed) + VP+ AdjP *It* is understandable *that people want to learn as much as possible.*	**before** (anteriority) The rain started ***before** he managed to get home.*	**That/which + VP** ⎤ (subject relative **Who + VP** ⎦ pronoun) There are many programmes *that offer this exchange.* People *who live in Latvia* are very helpful and friendly.
	NP + VP + AdjP I'm sure *that it's very useful for our country.*	**after** (posteriority) **After** he finished reading the book, he went out.	**whom +NP + VP** We are happy about the person *whom we have*

	NOUN CLAUSE	ADVERBIAL CLAUSE	ADJECTIVAL CLAUSE
			chosen for the post.
	NP + VP + PP + (that) <u>She said to Ann</u> (***that***) *she would like to go home.*	**Since** (duration, posteriority) There is silence ***since*** *the baby fell asleep.*	**PP + whom + NP + VP** He is the man ***at whom*** *the little girl laughed.* This is the car ***at which*** *I threw stones.*
	WH + NP + VP used as subject/object <u>*What I hate*</u> is work in the garden.	**As soon as** (simultaneity) ***As soon as*** *you finish this task*, we will go out.	**(who/that) + NP + VP + PP** He is the man (***who***) *the little girl laughed* ***at.*** This is the car (***that***) *I threw stones* ***at.***
	Indirect WH – questions He doesn't know ***what*** *they have to do.*	**While** (simultaneity, duration) ***While*** *there is crisis in Latvia,* a lot of people lose their jobs.	**whose + NP + NP + VP** The lady ***whose*** *daughter you like* is the wife of the president.
		Since (reason) ***Since*** *you were absent*, I will not ask you.	
		As (cause) ***As*** *they were late*, they missed dinner.	
		So/ so....(that) (effect) It was dark, ***so*** *he couldn't see the road.* The concert was ***so*** *nice* ***that I wanted to stay.***	
		Than **As....as** (comparison) **Not so ....as** The house was bigger **than** I thought. She practices **as** often **as** she can. His sister is **not so** intelligent **as** she thinks she is.	
B/B2	**Complex sentences with more than one embedded clause.**		
	NOUN CLAUSE	**ADVERBIAL CLAUSE**	**ADJECTIVAL CLAUSE**
	WH + VP as subject of the main clause <u>*What fascinated me the most*</u> was her negative attitude.	**Until/till** (duration, future) People will keep leaving Latvia ***till*** *the situation gets fixed.*	**That/which (que)+ NP + VP** (object **Who (que)+ NP + VP** relative pronoun) They get the work experience ***that*** *they would never get in Latvia.* We are happy about the person ***who*** *we have chosen for the post.*
	NP + VP + PP + WH At the beginning <u>they are afraid of</u> ***what*** *the others will*	**If (condition)**	

say.	if + past modal /past It would be a catastrophe *if everyone started speaking English.*	
NP + VP + NP + (that) I told him *(that) he could come in.*	if +past perfect/past perfect modal ***If I had not participated in the party***, I wouldn't have met her.	
NP + VP + that My conclusion is ***that*** people don't leave Latvia without a good reason.	if + past perfect/past modal ***If I had not broken my leg***, I wouldn't be in hospital now.	
Nominal relative clause I will start with *the statement **that** other countries pay better.*		
A/C1	**The structures mastered at level B2 are used in a more accurate way.**	

Table 29 **Syntactic criterial features in French**

F/A1	**Phrase simple.**		
E/A1	**Phrase simple.**		
D/A2	**Les phrases composées constituées de deux ou plusieurs clauses principales, reliées par conjonction de coordination.** Il est tard et je dois rentrer à la maison.		
	Phrase complexe.		
	LA SUBORDONNÉE COMPLÉTIVE	**LA SUBORDONNÉE CIRCONSTANCIELLE**	**LA SUBORDONNÉE RELATIVE**
	NP + VP+ (que) Je pense *qu'on peut parler des langues différentes.*	**où** (lieu) J'ai trouvé le sac *où je le mets toujours.*	
	Interrogative indirecte avec *si* Je ne sais pas *si je peux le faire.*	**quand** (temps) ***Quand ils vont à l'étranger***, ils voient beaucoup de l'endroits.	
		parce que (raison, cause) C'est très important en Lettonie ***parce que*** notre pays n'est pas très grand.	

		si (condition)	
		si présent / présent	
		Si la personne parle plusieurs langues, elle est plus intelligente.	
		si + présent/futur	
		Si on connaît beaucoup de langues, on pourra trouver du travail.	
C/B1	**Les phrases complexes limitées à une proposition subordonnée avec une structure simple.**		
	LA SUBORDONNÉE COMPLÉTIVE	**LA SUBORDONNÉE CIRCONSTANCIELLE**	**LA SUBORDONNÉE RELATIVE**
	Il (extrapose) + VP+ AdjP <u>*Il* est vrai</u> *que maintenant on apprend l'anglais dans toutes les écoles.*	**avant que** (antériorité) La pluie a commencé ***avant qu****'il réussisse à rentrer à la maison.*	**qui + VP** (pronom relative du sujet) Il y a plusieurs facteurs ***qui*** approuvent ça. Les gens ***qui*** connaissent beaucoup de langues sont respectes par les autres.
	NP + VP + AdjP <u>Je suis sûr</u> ***que*** ce serait aussi important d'apprendre des langues étrangères à l'avenir.	**après que** (postériorité) **Après qu**'il a fini de lire le livre, il est sorti.	**que +NP + VP** *Nous sommes heureux pour la personne* ***que*** *nous avons choisie pour le poste.*
	NP + VP + PP + que <u>Elle a dit à Anne</u> ***qu'****elle aimerait rentrer chez elle.*	**depuis que** (durée, postériorité) Il y a le silence ***depuis que*** *le bébé s'est endormi.*	**PP + qui/lequel + NP + VP** C'est une fille ***à qui*** j'ai donné mon livre préféré. C'est un homme ***sur lequel*** *la petite fille conte toujours.* Voici une poupée ***avec laquelle*** *j'ai joué toute la journée.*
	WH + NP + VP utilise comme sujet/objet *<u>Ce **que** je déteste</u> est le travail dans le jardin.* J'ai regardé dans la salle et <u>***ce que** j'y ai remarqué*</u> m'a choqué	**dès que** (simultanéité) ***Dès que*** *vous finissez cette tâche*, nous sortirons.	**dont + NP + NP + VP** La dame ***dont*** *la fille vous aimez* est celle du président.
	WH – questions indirectes Il ne sait pas ***ce qu****'ils doivent faire.*	**depuis que** (simultanéité, durée) ***Depuis que*** *la crise continue les* gens perdent leur emploi.	
		puisque (raison) ***Puisque*** *vous étiez absent*, je ne vais pas vous interroger.	
		comme (cause)	

		Comme ils étaient en retard, ils ont raté le dîner.	
		donc/ si....que (effet) Il faisait sombre, ***donc** il ne pouvait pas voir la route.* Le concert était *si beau **que** je voulais rester.*	
	.	**plus…que** **aussi....que** (comparaison) **si ....que** La maison était **plus** grande **que** celle que j'avais *prévue.* Elle pratique **aussi** souvent **qu'**elle le peut. Sa sœur n'est pas **si** intelligente **qu'**elle le pense.	
B/B2	**Les phrases complexes avec plus d'une proposition enchâssée**		
	LA SUBORDONNÉE COMPLÉTIVE	**LA SUBORDONNÉE CIRCONSTANCIELLE**	**LA SUBORDONNÉE RELATIVE**
	WH + VP comme le sujet de la phrase principale *Ce qui* me fascinait le plus, c'était son attitude négative.	**jusqu'a ce que** (durée, futur) Les gens continueront de quitter la Lettonie *jusqu'à ce que la situation soit corrigée.*	**que + NP + VP** (pronom relatif d'objet) Dans une langue on peut dire des choses *qu'une autre langue n'accepte pas.* Nous sommes heureux pour la personne *que nous avons choisie pour le poste.*
	NP + VP + PP + WH Au début ils ont peur de *ce que les autres vont dire.*	**si (condition)** **si + imparfait /conditionnel présent** *Si tout le monde commençait à parler anglais,* cela serait une catastrophe. **si+ plus-que-parfait/ conditionnel passé** *Si je n'avais pas participé à la fête*, je ne l'aurais pas rencontré. **si + plus-que-parfait/ conditionnel présent** *Si je ne m'étais pas cassé ma jambe*, je ne serais pas à l'hôpital maintenant.	
	NP + VP + NP + que Je lui ai dit *qu'il pouvait entrer.*	**Il y a ... (durée)** **Il y a dix ans** *que le monde est entré dans un nouveau millénaire.*	
	NP + VP + que Mon avis est que la connaissance de langues fait		

	penser à l'intelligence.		
	Phrase nominale relative Je suis d'accord avec l'*opinion **que** l'anglais est actuellement la langue de communication mondiale.*		
A/C1	**Les structures maîtrisées au niveau B sont utilisées d'une manière plus précise.**		

3.4 Empirical Validation of Syntactic Criterial Features

In order to evaluate the utility and applicability of the developed syntactic criterial features cross-linguistically when assessing student written performance in English and French, *the empirical validation* of the criteria was effected in two stages.

First, four experienced English language teachers were asked to assess different language proficiency level written scripts of six test-takers and four experienced French language teachers, who had been test-developers as well, were asked to assess four written scripts of year 2009 examination according to the marking scales (Appendices 5 and 6) elaborated by the NCE test-developers. It should be specified that the marking scale performance level descriptors ('the key referents that participants rely on when making whatever judgments a particular method requires' (Council of Europe, 2009: 73)) differed across the chosen foreign language. Moreover, in English each language proficiency level was allocated an integer (from 0 to 5) along with levels A – F, whereas in French only integers from 0 to 8 were provided to assess the student written performance, and only at the end the students were awarded levels for their language proficiancy examination. Thus, for the needs of the present research the marking scale was adapted by allocating each integer a corresponding level A – F before initiating the assessment of the sudent scripts. Moreover, the experts were also provided with the elaborated syntactic criterial features specific for each level of language proficiency and asked to utilize them when assessing grammar. Furthermore, the obtained results were arranged in tables for each foreign language apart and compared with the level awarded to the test-takers by the author of the present research (assessor 5 in *Tables 30* and *31*) as well as the level awarded at the centralized examination to see the decision consistency, i.e. the extent to which the decisions made by independent assessors would comply with each other and that of assessor 5.

Table 30 **Empirical validation results in English**

Script	Assessor 1	Assessor 2	Assessor 3	Assessor 4	Assessor 5	Level awarded at the examination
004	C	C	B	D	C	B
023	D	B	A	C	B	A
095	D	D	C	D	D	C
356	F	F	F	F	F	F
012	E	E	D	D	E	E
058	E	C	D	D	D	D

The data in *Table 30* show that in English in majority of cases the four independent assessor judgment complies with that of assessor 5, except for script 023 whose results vary from level D to level A. If compared with the results obtained at the examination, then the data reveal the discrepancy between the two judgments at the higher levels of language proficiency A - C. The NCE raters have assessed the student performance higher than the five independent markers. It might be explained by the fact that the criterial features used by the raters when assessing the written scripts were not available in 2009, thus the markers could not distinguish accurately between the two neighbouring levels, which was mentioned as the most important aspect when deciding upon the level to be awarded for grammar by the teachers participating in the present research data approbation.

Table 31 **Empirical validation results in French**

Script	Assessor 1	Assessor 2	Assessor 3	Assessor 4	Assessor 5	Level awarded at the examination
001	D	C	C	D	C	C
043	C	B	B	B	B	A
016	E	D	E	E	D	D
018	B	C	B	C	B	B

In French (*Table 31*) the level of proficiency of script 016 does not match the expected level D as three out of four assessors have awarded level E for the written performance. In addition, none of the markers have considered the written scripts to be at the highest level A, which was awarded for script 043 at the centralised examination. The reason for that might be the same as for the case in English, i.e. there is no clear distinction between levels as the descriptors included in the marking scale (Appendix 8) are rather general, and the assessors might not have been familiar with the grammatical criterial features for each proficiency level.

Thus, to have more substantiate data the empirical validation of the elaborated syntactic criterial features was carried out for the second time during year 2013 standardisation and benchmarking session of the test-taker written performance of the national centralised examination in English in cooperation with the test administrator and test developer from the NCE. Unfortunately, the repeated approbarion of the criterial features could be effected only in English as in French there was only one assessor who marked the student essays; moreover, she had already participated in the first stage of the validation.

Consequently, 100 experienced assessors from different schools all over Latvia participated in the validation session of the syntactic criterial features in English, which conformed to the requirements of having judges 'external to the institution producing the test' (Council of Europe, 2009: 38). During this session the assessors underwent several stages: 1) familiarization with the marking scales of written performance (Year 2013) common to all foreign languages (see Appendix 8) developed by the NCE experts in compliance with the CEFR proficiency level descriptors; 2) benchmarking, during which the assessors were provided 6 sample scripts of Task 1 containing a letter to a friend and 5 sample scripts of Task 2, in which the test-takers had to produce an essay of 200 - 250 words on the theme 'Films versus books' by expressing their opinion and providing reasons to support their point of view (it should be stated that the selected scripts represented different levels of language proficiency). During this stage the markers were asked to consider the performance (according to the marking scale), which was compared with the expert judgment. Before assessing Script 3 of Task 2 (Appendix 11) the panellists were handed out a slip of paper, on which they were asked to assess the grammatical competence apart from all the other parameters attributing scores from 5 (the highest) to 0 (the lowest or not enough to evaluate). It should be specified that each level was also linked to an integer (see *Table 32*) with the exception of level A1 as the test-takers of level A1 were not expected to perform the set task. Moreover, since 2012 the SEC test developers use a slightly different marking scale for assessing the second written task – essay. Level B2 has been sub-divided into B2.1 and B2.2 to have a more precise distinction between the higher levels of language proficiency (see the table below).

Table 32 **The correspondence of the language proficiency levels to integers**

Level	C1	B2.2	B2.1	B1	A2	0
Corresponding integer	5	4	3	2	1	0

Afterwards the assessors were familiarized with the elaborated syntactic criterial features typical for each level of language proficiency and asked to assess grammar once again by utilizing the provided criteria.

Further on, the quantitative analysis of the data was made to see the consistency of judgment. The first assessment (*Figure 101*) varies from 5 to 2 points (see Appendix 12), although the majority of the assessors, i.e. 70% awarded 3 points and only 21% – 4 points; 7%

− 2 points and only 5% of the assessors judged the grammatical competence by attributing 5 points.
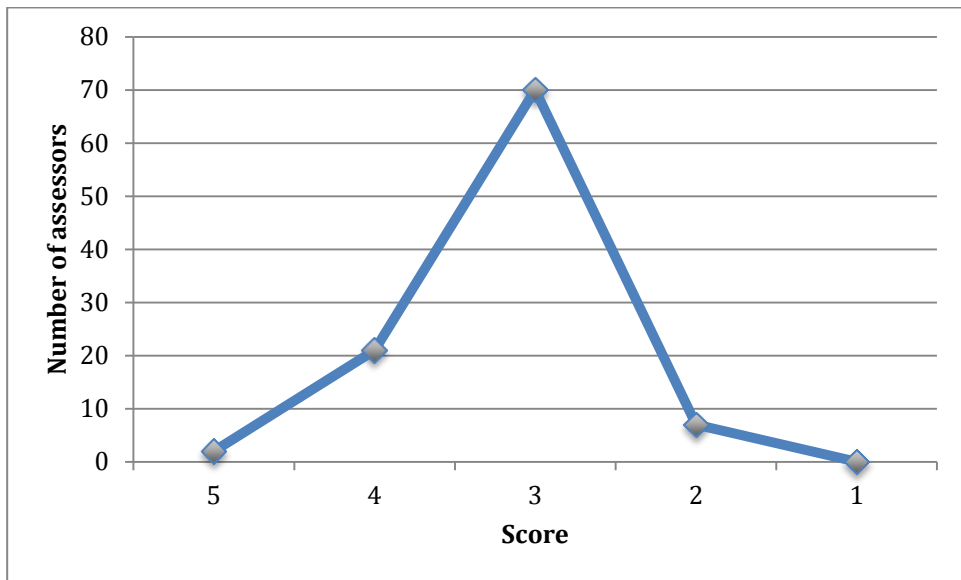
Figure 101 **Frequency of allocated scores for script 3 (assessment 1)**
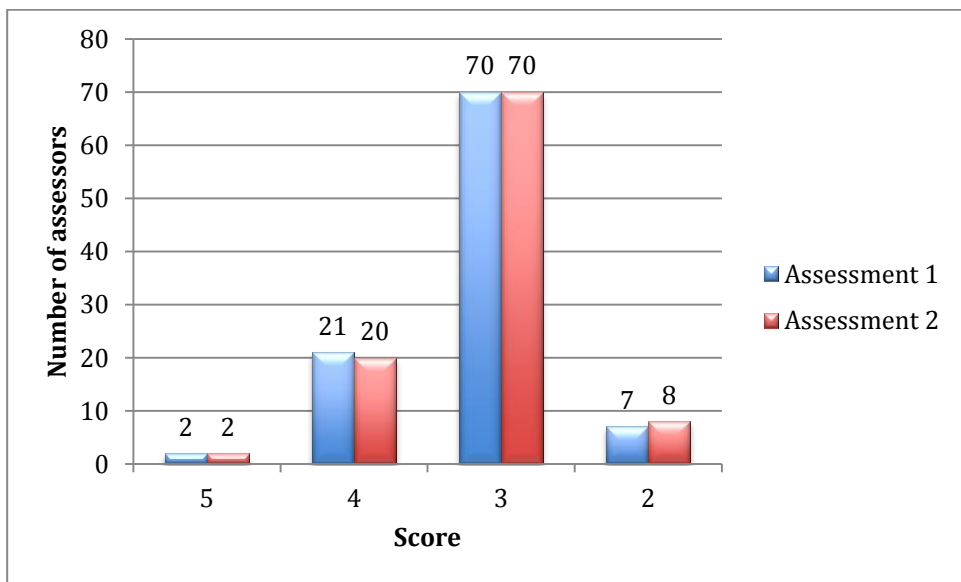
Figure 102 **Comparison of assessment 1 and 2 of script 3**

In the second assessment there is a slight difference as assessors Nr.47 and Nr.99 (Appendix 12) have judged the grammatical competence more severely when using the syntactic criteria (*Figure 102*). Overall, the obtained results allow us to see the distribution of scores, i.e. the *mean* (average score), which is 3.18 in assessment 1 and 3.16 in assessment 2 and the

correlation between the two judgments, which indicates that there is a strong relationship between assessment 1 and 2 as the correlation coefficient is very close to 1 (*Table 33*).

Table 33 **Statistic data of assessment 1 and 2 in English**

	Assessment 1	Assessment 2
Total score	318	316
Mean	3.18	3.16
Correlation coefficient	0,96162991	

These statistic data clearly indicate what the inter-rater reliability is (by inter-rater reliability we understand the 'tendency of different raters to give the same score to the same scripts' (Weigle, 2005: 135)), which determines the internal validity (accuracy and consistency) of the test results. In this particular case the expert point of view conforms to the majority (Appendix 11), who has awarded 3 points for grammar (level B2.1), which allows us to consider the assessment to be reliable. However, 30 markers have been either more severe or lenient in their judgment, which may signal the misunderstanding or misinterpretation of the marking scale descriptors. Two of the markers (Nr.37 and Nr.46) who awarded the highest level for the written performance should be removed from the process of assessing the written performance, as they are not familiar with the evaluation requirements.

The empirical validation results approve the necessity and validity of the syntactic criterial features in the process of assessing the test-taker written performance, in particular syntax, as they guarantee the reliability of judgment, which is a pre-requisite for the writing test construct validity. Moreover, the syntactic criteria allow us not only to make inferences about the test-taker writing abilities but they may also have a positive effect on instruction, as being disseminated among teachers they could be used in setting educational goals in the future, so that the students are familiar with the expected outcomes. However, we should understand that test validation in general is 'an on-going process and that the interpretations we make of test scores can never be considered absolutely valid' (Bachman and Palmer, 1996: 22). Therefore, the development of syntactic criteria, which was the aim of the present research, should be considered as an important stage in ensuring the validity of the written performance test.

CONCLUSIONS

The theoretical research on the linguistic development of language learners along with the compilation of the EFLWTC served as a basis for conducting a contrastive study on the syntactic structures produced by the test-takers of English and French, which resulted in elaborating the syntactic criterial features for different language proficiency levels. The effected research enables the author to draw the following conclusions.

- Various theoretical approaches in examining the human language development prove to be an integral part in comprehending the stages in language learning determined by various mental, cognitive, linguistic and psychological factors. The research on L1 and L2 discloses that there is a natural order in learning linguistic units similar to all learners irrespective of L1, L2 or foreign language. However, there may be variations in the order of accuracy. Moreover, the research on the first language is an essential pre-requisite in understanding the processes involved in learning a foreign language since it depends largely on interlanguage.

- The second language studies substantiate the notion of universality in generating syntactic structures as it reveals that learner grammars are restrained in the same way as natural language grammars, which supports the allegation of formal linguists that all languages contain principles that are language-invariant and parameters that allow variations, i.e. they are language-particular.

- The study of written performance assessment, which encompasses both competence and performance, accentuates that the language learner linguistic knowledge influences greatly the overall judgment. Moreover, as syntactic complexity is one of the most problematic areas to assess in student written performance, the markers should be familiarized with the stages of grammatical development in compliance with the CEFR levels to ensure validity and reliability of judgment at the high stakes examinations cross-linguistically.

- The contrastive analysis of different clause types based upon descriptive contemporary grammars serves as a theoretical grounding for the empirical research on syntactic structures since it unveils the similarities and differences in clause structure formation in English and French as well as demonstrates various semantic

functions of different clause types in relation to the main clause. Furthermore, it confirms the assumption that, in general, both languages are fixed order languages with some parametric variations.

- The corpus-based approach proves to be an indispensable tool in carrying out the quantitative, qualitative and contrastive analysis of different syntactic patterns, since it provides comparable data gathered according to the same sampling frame both in English and French. Moreover, being digitalized the compiled corpora may serve as a reusable source for further linguistic analysis.

- The quantitative and qualitative analysis of the complex sentences consisting of the main clause and one subordinate clause in the compiled EFLWT corpora reveals that noun clause is the most frequently used clause type at all levels of language proficiency followed by adverbial clause typical for lower levels of language proficiency, whereas the frequency of adjectival clauses increases towards the highest levels of language proficiency. The most frequently marked structures are: 1) *that*-clauses functioning as direct object; 2) adverbial clauses of time, reason and direct open condition, which perform the function of adjuncts; 3) restrictive relative clauses having the function of an object. The more complex patterns, which are rather rare, appear at higher levels of language proficiency. As for the placement of subordinate clauses in a sentence, noun clauses follow directly the main clause; adverbial clauses may be superordinated or subordinated to the main clause; adjectival clauses are either embedded in the main clause or subordinated to the main clause.

- The empirical research on the sentences containing multi subordinate clauses demonstrates that in majority of cases the marked sentences contain two subordinate clauses, yet their number may reach six clauses per sentence. Furthermore, the students tend to use almost the same syntactic structures marked in the sentences with mono subordinate clause, with the exception of relative clauses, as more complex structures, e.g. the restrictive relative clause of manner and place, have been utilised by the students at higher levels of language proficiency. As for the placement of subordinate clauses within a sentence, three main syntactic patterns predominate: 1) [main clause + subordinate clause + subordinate clause]; 2)

[subordinate clause + main clause + subordinate clause]; 3) [main clause containing an embedded clause + subordinate clause].

- The quantitative and qualitative research of different clause types in English and French enables the author:
  - To substantiate the assumption that overall, foreign language learners produce the same syntactic structures as the native language learners at the respective level of language proficiency. However, there are differences in the word order, which is determined by the fact that English and French belong to different Indo-European language groups (Germanic and Romance languages). This validates the findings of the formalists and functional-typologists that the main principles in structure generation are universal whereas certain parameters cause deviation from the norm.
  - To confirm Pienemann's (1998) allegation that there is a certain order in acquiring syntactic structures since adjectival clauses are a frequent feature in the student written performance of higher levels of language proficiency. Hence, adjectival clauses should be used as the discriminatory element when awarding qualification at the examination as their frequency of use increases towards the highest levels of language proficiency in both languages, which complies with the CEFR descriptors and English Profile grammatical features.
- The study of the EEP level descriptors and the ones provided by the NCE analytical marking scale highlighted that the lower level descriptors of the NCE scale are rather indistinct, i.e. they do not clearly state what syntactic structures are to be searched for when attributing a particular level for the language competence when assessing the test-taker written performance. This proved the necessity and utility for elaborating the syntactic criterial features for each language proficiency level based upon the data of the present research and the EEP.
- The empirical validation of syntactic criterial features affirms that they are of utmost significance when assessing the student grammatical competence since they facilitate deciphering between different language levels and ensure the objectivity and reliability of judgment cross-linguistically. Furthermore, the syntactic features may

be used in the learning process to set certain goals towards attaining the expected results.

The results attained during a thorough analysis of different syntactic structures extracted from the compiled EFLWT corpus allow the author to evaluate the present research as positive. Moreover, all the goals and objectives as well as the hypothesis advanced before commencing the doctoral thesis have been proved.

The developed syntactic criterial features will enable the educational administrators to ensure the uniform written performance assessment related to the CEFR levels, thus guaranteeing the test-takers receive reliable results at the examination. The compiled electronic bilingual (English and French) learner written text corpus may advance further studies on learner language. Since the present research was limited to only the finite clauses because of the labour intensive and time consuming corpus compilation and its grammatical annotation, the subsequent studies may examine the non-finite clauses in order to have a more profound analysis of different clause types. Moreover, the research on various syntactic structures as well as error analysis may be effected in longitudinal studies, which will provide more grounded data to be relied upon when setting educational goals. Finally, the cross-linguistic reliability of test results affected by the application of the elaborated syntactic criterial features may be investigated as well.

Bibliography

1.	Alderson, J. Ch. and Banerjee, J. (2002) Language Testing and Assessment. In *Language Teaching.* Volume 35/Issue 02, (pp.79–113) Cambridge: Cambridge University Press.
2.	Alderson, J. Ch. (2005) *Diagnosing Foreign Language Proficiency.* London: Continuum.
3.	Alderson, J. Ch. (2007) The challenge of (diagnostic) testing: do we know what we are measuring. In J. Fox, M. Wesche, D. Bayliss, L. Cheng, C. E. Turner and Ch. Doe (eds.), *Language Testing Reconsidered.* (pp.21-40) Ottawa: Ottawa University Press.
4.	Allison, D. (1999) *Language Testing and Evaluation: An Introductory Course.* Singapore: Singapore University Press.
5.	Atkins, S., Clear, J. and Ostler, N. (1992) Corpus design criteria. In *Literary and Linguistic Computing* 7/1: 1-16.
6.	Bachman, L.F. (1990*) Fundamental Considerations in Language Testing.* Oxford: Oxford University Press.
7.	Bachman, L.F. and Palmer, A.S. (1996) *Language Testing in Practice*. Oxford: Oxford University Press.
8.	Bachman, L.F. (2004) *Statistical Analyses for Language Assessment.* Cambridge: Cambridge University Press.
9.	Bachman, L.F. (2007) What is the construct? The dialectic of abilities and contexts in defining constructions in language assessment. In J. Fox, M. Wesche, D. Bayliss, L. Cheng, C. E. Turner and Ch. Doe (eds.), *Language Testing Reconsidered.* (pp.41-66) Ottawa: Ottawa University Press.
10.	Bachman, L.F. and Palmer, A. (2010) *Language Assessment in Practice*. Oxford: Oxford University Press.
11.	Bailey, N., Madden, C. and Krashen, S.D. (1974) Is there a 'natural sequence' in adult second language learning? In *Language Learning 24*: 235-243.
12.	Berk, L. (1986) Relationship of elementary scholl children's private speech to behavioural accompaniment to task, attention and task performance. In *Developmental Psychology, 22 (5)*: 671-680.
13.	Biber, D. (1990) Methodological issues regarding corpus-based analyses of linguistic variation. In *Literary and Linguistic Computing, 5*: 257-269.
14.	Biber, D. (1993) Representativeness in corpus design. In *Literary and Linguistic Computing,* 8/4: 243-257.
15.	Biber, D., Johansson, S., Leech, G., Conrad, S. and Finegan, E. (eds.), (1999) *Longman Grammar of Spoken and Written English*. England: Pearson Education Limited.
16.	Biskri I. and Desclés J-P. (2005) Applicative and Combinatory Categorial Grammar and Subordinate Constructions in French. In *International Journal of Artificial Intelligence Tools*, vol.14 N°1 & 2: 125-136.
17.	Blokh, M.Y. (2003) *A Course in Theoretical English Grammar.* Москва: Высшая школа.
18.	Bloom, L. and Lahey, M. (1978) *Language Development and Language Disorders*. New York: John Wiley & Sons.
19.	Bornstein, D.D. (1976) *Readings in the Theory of Grammar*. Cambridge, Massachusetts: Winthrop Publishers, Inc.

20. Bowker, L. and Pearson, J. (2002) *Working with Specialized Language*: *A practical guide to using corpora*. London and New York: Routledge.
21. Brown, R. (1973) *A First Language*. Cambridge, MA: Harvard University Press.
22. Canale, M. and Swain, M. (1980) Theoretical bases of communicative approaches to second language teaching and testing. In *Applied Linguistics*, I (1): 1-47.
23. Carnie, A. (2007) *Syntax: A Generative Introduction*. Second Edition. USA, UK, Australia: Blackwell Publishing.
24. Carroll, J.B. (1961) Fundamental considerations in testing for English Proficiency of foreign students. In *Testing the English Proficiency of Foreign Students*. (pp.30-40) Washington, DC: Center for Applied Linguistics.
25. Carroll, J.B. (1968) The psychology of language testing. In A.Davies (ed.) *Language Testing Symposioum: A Psycholinguistic Approach.* (pp.46-69) London: Oxford University Press.
26. Carter, R. and McCarthy, M. (2006) *Cambridge Grammar of English*. Cambridge: Cambridge University Press.
27. Chevalier, J-C. and Blanche-Benveniste, C. (1964) *Grammaire Larousse du français contemporain*. Paris: Librairie Larousse.
28. Chaquet, M. and Paillard, M. (1989) *Approche Linguistique des Problèmes de Traduction Anglais >< Français*. Paris: Ophrys.
29. Chomsky, N. (1965) *Aspects of the Theory of Syntax.* Cambridge, Massachusetts: MIT Press
30. Chomsky, N. (1980a) *Rules and Representations*. Oxford: Blackwell
31. Chomsky, N. (1980b) On Cognitive Structures and Their Development. In M.Piattelli-Palmarini (ed.), *Language and Learning: The Debate between Jean Piaget and Noam Chomsky*. London: Routledge & Kegan Paul.
32. Chomsky, N. (1981) *Lectures on Government and Binding.* Dortrecht: Foris.
33. Chomsky, N. (1986) *Knowledge of Language: Its Nature, Origin and Use.* New York: Praeger.
34. Chomsky, N. (1995) *The Minimalist Program*. Massachusetts: Massachusetts Institute of Technology.
35. Cizek, G.J. and Bunch, M.B. (2007) *Standard Setting*. Thousand Oaks, CA: Sage.
36. Cobb, T. (2009) *Compleat Lexical Tutor v.6.0* Available from *http://www.lextutor.ca* [Accessed September 19, 2013].
37. Comrie, B. (1989) *Language Universals and Linguistic Typology*. Second edition. Chicago: the University of Chicago Press.
38. Cook, V. (1993) *Linguistics and Second Language Acquisition*. New York: St.Martin's Press.
39. Cook, V. J. and Newson, M. (1996) *Chomsky's Universal Grammar*. An introduction. Second edition. Oxford: Blackwell publishers.
40. Cosme, Ch. (2006) A corpus-based perspective on clause linking in English, French and Dutch. In Fabricius-Hansen, Catherine, Ramm and Wiebke (eds.), *'Subordination' versus 'Coordination' in Sentence and Text: A cross linguistic perspective.* Amsterdam/Philadelphia: John Benjamins Publishing Company.
41. Council of Europe (2001) *A Common European Framework of Reference for Languages: Learning, Teaching, Assessment.* Cambridge: Cambridge University Press.
42. Council of Europe (2003) *Relating Language Examinations to the Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEF).*

Manual. Overview of Preliminary Pilot Version. Strasbourg: Language Policy Division.

43. Council of Europe (2006) *Plurilingual Education in Europe. Strasbourg: Language Policy Division.* Available from http://www.coe.int/t/dg4/linguistic/Source/PlurinlingalEducation_En.pdf [Accessed October 18, 2013].

44. Council of Europe (2009) *Relating Language Examinations to the Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEFR): A Manual.* Strasbourg: Language Policy Division.

45. Cristofaro, S. (2003) *Subordination. Oxford Studies in Typology and Linguistic Theory*. New York: Oxford University Press.

46. Croft, W. (1990) *Typology and Universals*. Cambridge: Cambridge University Press.

47. Croft, W. (2001) *Radical Construction Grammar. Syntactic Theory in Typological Perspective.* Oxford: Oxford University Press.

48. Crystal, D. (2007) *How Language Works*. London: Penguin Books Limited.

49. De Hann, P. (1984) Problem-oriented tagging of English corpus data. In J. Aarts and W. Meijs (eds.), *Corpus Linguistics.* (pp.123-139) Amsterdam: Rodopi.

50. De Rivarol, A. (1784) *Discours de l'Université de la language Française.* Berlin, Paris.

51. De Villiers, J. G., Tager Flusberg, H. B., Hakuta, K. and Cohen, M. (1979) Children's Comprehension of Relative Clauses. In *Journal of Psycholinguistic Research*, Vol.8, No 5: 499-518.

52. Delatour, Y., Jennepin, D., Léon-Dufour, M., et Teyssier, B. (2004) *Nouvelle Grammaire du Français.* Paris: Hachette.

53. Dik, S.C. (1991) Functional Grammar. In F.Droste and J.Joseph (eds.), *Linguistic. Theory and Grammatical Description*. (pp.247-74) Amsterdam: Benjamins.

54. Dik, S.C. (ed.by Kees Hengveld), (1997) *The Theory of Functional Grammar. Part 1: The Structure of the Clause*. Berlin, New York : Mouton de Gruyer.

55. Dik, S.C. (ed.by Kees Hengveld), (1997) *The Theory of Functional Grammar. Part 2: Complex and Derived Constructions*. Berlin, New York : Mouton de Gruyer.

56. Doughty, C. (1988) *The effect of instruction on the acquisition of relativization in English as a second language*. Ph.D. dissertation, University of Pennsylvania.

57. Doughty, C. (1991) Second language instruction does make a difference: evidence from an empirical study of second language relativisation. In *Studies in Second Language Acquisition,* 13/4: 431-469.

58. Dulay, H., and Burt, M. (1974) Natural sequence in child second language acquisition. In *Language learning*, 24: 37-53.

59. Ellis, N. C. (2006) Selective attention and transfer phenomena in L2 acquisition: Contingency, cue competition, salience, interference, overshadowing, blocking, and perceptual learning. In *Applied Linguistics, 27 (2):* 164-194.

60. Ellis, R. (1997) Series Editor H.G.Widdowson. *Second Language Acquisition*. Oxford: Oxford University Press.

61. *English Profile: Introducing the CEFR for English* (2011). Cambridge: UCLES/Cambridge University Press.

62. Farrell, P. (2005) *Grammatical Relations*. Oxford: Oxford University Press.

63. Ferrara, S., Perie, M., and Johnson, E. (2002). *Matching the Judgmental Task With Standard Setting Panelist Expertise: The Item-Descriptor (ID) Matching Procedure*.

Presented to the National Research Council's Board of Testing and Assessment (BOTA), Washington, DC.

64. Figueras, N., North, B., Takal, S., Verhelst, N. And Van Avermaet, P. (2005) Relating examination to the Common European Framework: A manual. In *Language Testing 22*: 261-179.

65. Firbas, J. (1992) *Functional sentence perspective in written and spoken communication*. Cambridge: Cambridge University Press.

66. Fischer O., Van Kemenade, A., Koopman, W. and Van der Wurff, W. (2000) *The Syntax of Early English*. Cambridge: Cambridge University Press.

67. Foley, W. A. and Van Valin, R. D. Jr. (1984) *Functional Syntax and Universal Grammar*. Cambridge: Cambridge University Press.

68. Gass S. M. and Selinker L. (2008) *Second Language Acquisition. An Introductory Course*. New York: Routledge

69. Geoffrey, L. (2011) Frequency, Corpora and Language Learning. In F. Meunier, S. De Cock, G. Gilquin and M. Paquot (eds). *A Taste for Corpora: In honour of Sylviane Granger.* Amsterdam and Philadelphia: John Benjamins.

70. Gilquin, G. (2003) Causative *get* and *have*. In *Journal of English Linguistics* 31/2: 125-148.

71. Givón, T. (1979) From discourse to syntax: Grammar as a processing strategy. In T. Givón (ed.), *Discourse and Syntax: Syntax and Semantics, Vol.12*. (pp. 81-112) New York: Academic Press.

72. Givón, T. (1990) *Syntax: A Functional-Typological Introduction.* Vol.2 Amsterdam and Philadelphia: John Benjamins.

73. Goldschneider, J. and DeKeyser, R. (2001) Explaining the 'natural order of L2 morpheme acquisition' in English: A meta-analysis of multiple determinants. In *Language Learning, 51*: 1-50.

74. Grabe, W. and Kaplan, R.B. (1996) *Theory and Practice of Writing.* New York: Longman.

75. Grabowski, J. (1996) Writing and speaking: Common grounds and differences toward a regulation theory of written language productin. In C.M.Levy and S.Ransdell (eds.), *The Science of Writing.* NJ: Lawrence Erlbaum Associates.

76. Granger, S. (2003) The corpus approach: A common way forward for contrastive linguistics and translation studies. In S.Granger, J.Lerot and S.Petch-Tyson (eds.), *Corpus-based approaches to contrastive linguistics and translation studies*. (pp.17-29) Amsterdam: Rodopi.

77. Graves, D. (1984) *A Researcher Learns to Write.* Portsmouth, NH: Heinmann.

78. Greenberg, J.H. (1963) Some universals of grammar particular reference to the order of meaningful elements. In J.Greenberg (ed.), *Universals of Grammar.* (pp.73-113) Cambridge, MA: MIT Press.

79. Grevisse, M. (1969) *Cours d'analyse grammaticale*. Paris: Duculot.

80. Grevisse, M. et Goosse, A. (1993) *Le bon usage, grammaire française*. Paris: Ducuclot. Édition par André Goosse.

81. Guiraud, P. (1970) *La Syntaxe du Français.* Paris : Presses Universitaires de France.

82. Haegeman, L. (2006) *Thinking Syntactically*. *A Guide to Argumentation and Analysis*. USA: Blackwell Publishing.

83. Hakulinen, A., Karlsson, F. and Vilkuma, M. (1980) *Suomen tekstilauseiden piirteitä: kvantitatiivinen tutkimus.* Department of General Linguistics, University of Helsinki. Publications No.6.

84. Halliday, M. A. K. Revised by C. Matthiessen (2004) *An Introduction to Functional Grammar.* Third edition. London: Hodder Arnold.

85. Hamp-Lyons (2003) *Writing Teachers as Assessors of Writing.* Available from www.hamp-lyons.com/writing_teachers_as_assessors_of_writing.doc [Accessed July 5, 2013].

86. Hamp-Lyons, L. (1990). Second language writing: assessment issues. In B. Kroll (Ed.), *Second Language Writing: Research insights for the classroom.* (pp. 69-87) Cambridge and New York: Cambridge University Press.

87. Haskell, R. E. (2001) *Transfer of Learning: Cognition, Instruction, and Reasoning.* The United States of America: Academic Press.

88. Haspelmath, M. (1995) The converb as a cross-linguistically valid category. In M.Haspelmath and E.Köning (eds.), *Converbs in Cross-linguistic Perspective.* (pp.1-55) Berlin: Mouton de Gruyter.

89. Hasselgård, H. and Johansson, S. (2011) Learner corpora and cotrastive interlanguage analysis. In F.Meunier, G.Gilquin, M.Paquot (eds.), *A Taste for Corpora: In Honour of Sylvian Granger.* (pp.33-61) Amsterdam: John Benjamins Publishing Company.

90. Hawkins, J.A. (1994) *A Performance Theory of Order and Constituency.* Cambridge: Cambridge University Press.

91. Hawkins, J.A. and Buttery, P. (2010) Criterial features in learner corpora: Theory and illustrations. In *English Profile Journal* 1 (1), e5.

92. Hawkins, J.A. and Filipović, L. (2012) *Criterial Features in L2 English: Specifying the Reference Levels of the Common European Framework (English Profile Studies).* Cambridge: Cambridge University Press.

93. Hawkins, R. (2006) *Second Language Syntax.* Oxford: Blackwell Publishing.

94. Hawkins, R. and Towell, R. (2001) *French Grammar and Usage.* London: Arnold.

95. Hayes, J. R. and Flower, L. (1980) A cognitive process theory of writing. In *College Composition and Communication, 32*: 365-387.

96. Heigham, J. and Goker, A. (eds.), (2009) *Qualitative Research in Applied Linguistics.* England: Palgrave MacMillan.

97. Hendriks, H. (2008) Presenting the English Profile Programme: In search of criterial features. *In Cambrigdge ESOL: Research Notes 33*: 7-10.

98. Henning, G. (1987) *A Guide to Language Testing, Development, Evaluation, Research.* London: Newbury House.

99. Hilles, S. (1986) Interlanguage in the pro-drop parameter. In *Second Language Research*, 2: 33-52.

100. Holvūts, A. (2012) *Vispārīgās sintakses pamati.* Rīga: Latviešu valodas aģentūra.

101. Huddleston, R. D. (1984) *Introduction to the Grammar of English.* Cambridge: Cambridge University Press.

102. Hunston, S. (2002) Pattern grammar, language teaching, and linguistic variation. In R.Reppen, S.M.Fitzmaurice and D.Biber (eds.), *Using Corpora to Explore Linguistic Variation.* (pp.167-183) Amsterdam: John Benjamins.

103. Hyland, K. (2003) *Second Language Writing.* Cambridge: Cambridge University Press.

104. Hyland, K. (2004) *Genre and Second Language Writing.* Michigan: The University of Michigan Press.
105. Hymes, D. H. (1972) On communicative competence. In J.B. Pride and J.Holmes (eds.), *Sociolinguistics*. Harmondsworth, England: Penguin Books.
106. Jackson, H. (2007) *Key Terms in Linguistics*. London: Continuum.
107. Johansson, S. and Hasselgård, H. (1999) Corpora and cross-linguistic research in the Nordic countries. In S.Granger, L.Beheydt and J.-P.Colson (eds.), *Contrastive linguistics and translation, a special issue of Le Langage et l'Homme 34.I.* (pp.145-162) Leuven: Peeters.
108. Jones, M.A. (1996) *Foundations of French Syntax*. Cambridge: Cambridge University Press.
109. Jones, N. (2009) A comparative approach to constructing a multilingual proficiency framework: contraining the role of standard-setting. In N.Figueras and J.Noijons (eds.), *Linking to the CEFR levels: Research perspectives*. (pp.35-43) Arnhem: Cito, EALTA, Council of Europe.
110. Kaftandjieva, F. (2004) Standard setting. In *Reference Supplement B to the Preliminary Pilot version of the Manual for Relating Language Examinations to the Common European Framework of Reference for Languages: learning, teaching, assessment*. Strasbourg: Council of Europe.
111. Kalnberzina, V. (2007) *Impact of Relation of Year 12 English Language Examination to CEFR on the Year 12 Writing Test.* Paper presented at the International Conference of the FIPLV Nordic-Baltic Region 'Innovations in Language Teaching and Learning in the Multicultural Context'. Riga: Latvia.
112. Kannas, C. (2003) *Le Becherelle Pratique de la Language Française*. Paris: Hatier.
113. Keenan, E.L. and Comrie B. (1977) Noun phrase accessibility and universal grammar. In Linguistic Inquiry 8: 63-99.
114. Keil, H. (ed.), (1855) *Grammatici Latini,* 2, p.53. Leipzig.
115. Kenning, M.M. (2010) What are parallel and comparable corpora and how can we use them? In A.O'Keeffe and M.McCarthy (eds.), *The Routledge Handbook of Corpus Linguistics*. (pp.487-499) New York: Routledge.
116. Kies, D. (2005) *The Clause*. Available from http://papyr.com/hypertextbooks/grammar/cl_intro.htm [Accessed November 3, 2013]
117. Koptojevskaja-Tamm, M. (1993) *Nominalizations.* London/New York: Routledge.
118. Krashen, S. (1977) Some issues relating to the monitor model. In H.Brown, C.Yorio and R.Crymes (eds.), On TESOL '77: *Teaching and Learning English as a Second Language : Trends in Research and Practice*. (pp. 144-158) Washington, DC: TESOL.
119. Kroeger, P.R. (2004) *Analysing Syntax*. Cambridge: Cambridge University Press.
120. Kunda, T. (2011) *Relating Latvian Year 12 Examinaation in English to the CEFR*. Available from http://visc.gov.lv/eksameni/vispizgl/dokumenti/20110920_petijums_en.pdf [Accessed January 14, 2012].
121. Lado, R. (1957) *Linguistics across Cultures: Applied Linguistics for Language Teachers*. Ann Arbor MI: University of Michigan Press.
122. Lado, R. (1961) *Language Testing: The Construction and Use of Foreign Language Tests*. London: Longman.
123. Langacker, R. W. (1972) *Fundamentals of Linguistic Analysis*. New York: Harcourt Brace Jovanovich, Inc.

124. Langacker, R. W. (1991*) Foundations of Cognitive Grammar. Vol.II: Descriptive Applications.* Stanford: Stanford University Press.
125. Larsen-Freeman, D. (1976) An explanation for the morpheme acquisition order of second language learners. In *Language Learning*, 26: 125-134.
126. Latvijas Republika (1992) *Valsts valodas likums*. Available from http://likumi.lv/doc.php?id=14740 [Accessed October 10, 2013].
127. Le Goffic, P. (1993) *Grammaire de la Phrase Française*. Paris: Hachette.
128. Leech, G. (1991) The state of art in corpus linguistics. In K. Aijmer and B. Altenberg (eds.), *English Corpus Linguistics*. (pp.8-29) London: Longman.
129. Leech, G. (1992) Corpora and theories of linguistic performance. In J. Svartvik (ed.), *Directions in Corpus Linguistics*. (pp.105-122) Berlin: Mouton de Gruyter.
130. Leech, G. (1997) Introducing corpus annotation. In R Garside, G. Leech and A. McEnery (eds.), *Corpus Annotation*. (pp. 1-18) London: Longman.
131. Leech, G. (2011) Frequency, corpora and language learning. In F. Meunier, G. Gilquin, M. Paquot (eds.), *A Taste for Corpora: In Honour of Sylvian Granger*. (pp.7-32) Amsterdam: John Benjamins Publishing Company.
132. Liddicoat, A. J. (2007) *An Introduction to Conversation Analysis.* London/New York: Continuum.
133. Lightbown, P. M. (1983) Exploring relationships between developmental and instructional sequences in L2 acquisition. In H. Selinger and M. H. Long (eds.), *Classroom Oriented Research in Second Language Acquisition.* (pp.217-243) Rowley, MA: Newbury House.
134. Lightbown, P. M. and Spada, N. (1999) *How Languages are Learned*. Second edition. Oxford: Oxford University Press.
135. Lindquist, H. (2009) *Corpus Linguistics and the Description of English.* Edinburgh: Edinburgh University Press.
136. Littlewood, W. T. (2004) *Foreign and Second Language Learning.* Cambridge: Cambridge University Press.
137. Lyons, J. (1968) *Introduction to Theoretical Linguistics.* Cambridge: Cambridge University Press.
138. Mairal, R. and Gil, J.(eds.), (2006) *Linguistic Universals*. Cambridge: Cambridge University Press.
139. Manning, C. and Schütze, H. (1999) *Foundations of Statistical Natural Language Processing*. Cambridge, MA: The MIT Press.
140. Matthews, P.H. (1981) *Syntax*. Great Britain: Cambridge University Press.
141. Matthews, P.H. (1997) *The Concise Oxford Dictionary of Linguistics*. Oxford: Oxford University Press.
142. McEnery, T. (2003) Corpus linguistics. In R.Mitkov (ed.), *The Oxford Handbook of Computational Linguistics.* (pp.448-463) Oxford: Oxford University Press.
143. McEnery, T. and Hardie, A. (2012) *Corpus Linguistics: Method, Theory and Practice*. Cambridge: Cambridge University Press.
144. McEnery, T. and Wilson A. (2001) *Corpus Linguistics*. 2nd edition. Edinburgh: Edinburgh University Press.
145. McEnery, T. and XIAO, Z. (2007) Parallel and comparable corpora – The state of play – In Y.Kawaguchi, T.Takagaki, N.Tomimori and Y.Tsuruga (eds.), *Corpus-based Perspectives in Linguistics*. (pp.131-146) Amsterdam: John Benjamins Publishing Company.

146. McEnery, T., Xiao, R. and Tono, Y. (2006) *Corpus-Based Language Studies: and advanced resource book.* London and New York: Routledge

147. McNamara, T. (1996) *Measuring Second Language Performance.* London and New York: Longman.

148. Messick, S. (1989). Validity. In R. L. Linn (Ed.) *Educational Measurement* (3rd ed.). (pp. 13-103) New York: Macmillan.

149. Messick, S. (1994) The interplay of evidence and consequences in the validation of performance assessments. In *Educational Researcher 23,* 2: 13-23.

150. Meyer, C.F. (2002) *English Corpus Linguistics*. Cambridge: Cambridge University Press.

151. Ministru kabineta noteikumi Nr.715 (2008) *Noteikumi par valsts vispārējās vidējās izglītības standartu un vispārējās vidējās izglītības mācību priekšmetu standartiem.* Available from http://likumi.lv/doc.php?id=181216 [Accessed June 5, 2010]

152. Mowrer, O.H. (1957) *Learning Theory and Behavior*. Urabana III: University of Illionois Press.

153. North, S. (1987) *The Making of Knowledge in Composition*. Portsmouth, NH: Henmann.

154. O'Grady, W. (2005) The Radical Middle: Nativism without Universal Grammar. In Catherine J. Doughty and Michael H. Long (eds.), *The Hanbook of Second Language Acquisition*. (pp.43-62) USA: Blackwell Publishing.

155. O'Grady, W. and Archibald, J. (2004) *Contemporary Linguistic Analysis*. Fourth Custom Edition for the University of Alberta, Pearson Custom Publishing.

156. Osgood, Ch. E. (1957) *The Measurement of Meaning*. Illionois: University of Illionois Press.

157. Owens, R. E. JR. (1992) *Language Development*: an introduction. New York: Merill, Macmillan Publishing Company.

158. Papageorgiou, S. (2009) *Setting Performance Standards in Europe: The Judges Contribution to Relating Language Examinations to the Common European Framework of Reference*. Frankfurt am Mein: PETER LANG.

159. Park, H. (2004) A minimalist approach to null subjects and objects in second language acquisition. In *Second Language Research*, 20: 195-223.

160. Pienemann, M. (1985) Learnability and syllabus construction. In K.Hyltenstam and M.Pienemann (eds.), *Modelling and assessing second language acquisition*. (pp. 23-77) San Diego, CA: College Hill Press.

161. Pienemann, M. (1998) *Language Processing and Second Language Development: Processability Theory*. Amsterdam/Philadelphia: John Benjamins Publishing Company.

162. Pienemann, M. (2008) Processability Theory. In B. VanPattern and J. Williams (eds.), *Theories in Second Language Acquisition*. (pp.137-154) New York: Routledge.

163. Pienemann, M., Johnston, M. and Brindley, G. (1988) Constructing an acquisition-based procedure for scond language assessment. In *Studies in Second Language Acquisition 10*: 217-224.

164. Purpura, J.E. (2004) *Assessing Grammar*. Cambridge: Cambridge University Press.

165. Quirk, R., Greenbaum, S., Leech, G. and Svartvik, J. (1985) *A Comprehensive Grammar of the English Language.* Index by David Crystal. England: Longman.

166. Riegel, M., Pellat, J-C. et Rioul, R. (1994) *Grammaire Méthodique du Français*. Paris: Press Universaitaire de France.

167. Rowlett, P. (2007) *The Syntax of French*. Cambridge: Cambridge University Press.
168. Saville-Troike, M. (2006) *Introducing Second Language Acquisition.* New York: Cambridge University Press.
169. Schmied, J. (1993) Qualitative and quantitative research approaches to English relative constructions. In C.Souter and E.Atwell (eds.), *Corpus Based Computational Linguistics*. Amsterdam: Rodopi.
170. Scott, M. (2005) *WordSmith Tools 5.0* Available from http://www.lexically.net/wordsmith/version5/index.html [Accessed May 14, 2010].
171. Silva, T. (1993) Towards an understanding of the distinct nature of L2 writing: The ESL research and its implications. *TESOL Quarterly*, 27: 657-676.
172. Sinclair, J. (1991) *Corpus, Concordance and Collocation*. Oxford: Oxford University Press.
173. Sinclair, J. (1996) EAGLES preliminary recommendations on corpus typology. In *EAG-TCWG-CTYP/P*. Pisa: ILC-CNR.
174. Sinclair, J. (2004) *Trust the Text: language, corpus and discourse.* New York: Routledge.
175. Skinner, B. F. (1957) *Verbal Behaviour*. Acton, MA: Copley.
176. Stanford Natural Language Processing Group (2006 [2010]) *Stanford NLP Parser.* Available from http://nlp.stanford.edu/software/lex-parser.shtml [Accessed January 6, 2013].
177. SurveyLang (2011) *Standard setting*. Available from http://www.surveylang.org [Accessed November 4, 2013].
178. Svartvik, J. (ed.) (1992) *Directions in Corpus Linguistics*. Berlin: Mouton de Gruyter.
179. Tallerman, M. (2005) *Understanding Syntax*. Second edition. Malta: Hodder Arnold.
180. Teubert, W. and Cermáková, A. (2007) *Corpus Linguistics: A Short Introduction. London*: Continuum.
181. Tognini-Bonelli, E. (2001) *Corpus Linguistics at Work*. Amsterdam: John Benjamins.
182. Tono, Y. (2002) *Learner corpora: design, development and applications.* Graduate School of Applied Linguistics, Meikai University, JAPAN. Available from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.115.6849&rep=rep1&type=pdf [Accessed September 5, 2013].
183. Van Ek, J. A. and Trim, J. L. M. (1991) *Threshold Level 1990*. Strasbourg: Council of Europe Press.
184. Van Ek, J. A. and Trim, J. L. M. (2001) *Vantage*. Cambridge: Cambridge University Press.
185. Van Valin, R. D. Jr. and LaPolla, R. J. (1997) *Syntax: Structure, Meaning and Function.* Cambridge: Cambridge University Press.
186. VanPatten, B. and Williams, J. (eds.), (2008) *Theories in Second Language Acquisition. An Introduction.* New York: Routledge.
187. VISC (2011a) Centralizētais eksāmens par vispārējās vidējās izglītības apguvi. Angļu valoda. Examination material. Available from http://visc.gov.lv/eksameni/vispizgl/uzdevumi/2011/vidussk/12_eng_term.pdf [Accessed February 29, 2012].
188. VISC (2011b) Arī šajā mācību gadā centralizēto eksāmenu vērtējumi tiks izteikti līmeņos. Available from: http://visc.gov.lv/eksameni/vispizgl/info.shtml [Accessed November 16, 2011].
189. VISC (2012) Par centralizēto eksāmenu paraugiem svešvalodās (angļu, franču, krievu un vācu valodā). Available from

http://visc.gov.lv/eksameni/vispizgl/dokumenti/20110315_par_ce_paraugiem_svesvalodas.pdf [Accessed February 15, 2012].

190. Weigle, S. C. (2002) *Assessing Writing.* Cambridge: Cambridge University Press.
191. Weir, C.J. (1990) *Communicative Language Testing*. Prentice Hall International.
192. Weir, C.J. (2005) *Language Testing and Validation: An Evidence-Based Approach.* New York: Palgrave Macmillan.
193. White, L. (1985) The pro-drop parameter in adult second language acquisition. In *Language Learning 35*: 47-62.
194. White, L. (2003) *Second Language Acquisition and Universal Grammar.* Cambridge*:* Cambridge University Press.
195. Widdowson, H. (1993) Innovation in teacher development. In *Annual Review of Applied Linguistics 13*: 260-275.
196. Wilmet, M. (1997*) Grammaire critique du français*. Paris: Hachette superieur, Duculot.

Internet sources

1. Available from http://homepages.uwp.edu/canary/grammar_text/82-extraposed.html [Accessed March 18, 2012].
2. Available from http://www.cambridge.org/elt/corpus/ [*Accessed on August 5, 2012]*

Appendices

Appendix 1
Acquisition of Sentence Forms within Brown's Stages of Development (adapted from Owen, 1992)

Stage	Age (in months)	Declarative	Negative	Interrogative	Embedding	Conjoining
Early I	12 – 22	Agent + action; Action + object	Single word – *no, all gone, gone;* negative + X	*Yes/no* asked with rising intonation on a single word; *what* and *where*		Serial naming without *and*
Late I	22 – 26	Subj + verb + obj appears	*No* and *not* used interchangeably	*That + X; what* + noun phrase + *(doing)?* *Where* + noun phrase + *(going)?*	Prepositions *in* and *on* appear	*And* appears
Early II	27 – 28	Subj + copula + compl appears				
Late II	28 – 29	Basic subject-verb-object used by most children	*No, not, don't* and *can't* used interchangeably; negative element placed between subject and predicate	*What* or *where* + subj + pred; Earliest inversion appears with copula in *What/where* + copula + subj	*Gonna, wanna, gotta*, etc. appear	
Early III	31 – 32	Subj + aux + verb + obj appears; auxiliary verb forms *can, do, have, will* and *be* appear				*But, so, or,* and *if* appear
Late III	33 – 34	Auxiliary verb appears with copula in subj + aux + copula + X	*Won't* appears	Auxiliary verbs *do, can*, and *will* begin to appear in questions; inversion of subject and auxiliary verb appears in yes/no questions		
Early IV	35 – 37		Negative appears with auxiliary verbs (subj + aux + neg + verb)	Inversion of auxiliary verb and subject in *wh-* questions	Object noun phrase complements appear with verbs such as *think, guess*, and *show*; embedded *wh-* questions	Clausal conjoining with *and* appears (most children cannot produce this form until late stage V); *because* appears
Late IV	38 – 40	Double auxiliary verbs appear in subj+ aux + aux + verb + X	Adds *isn't, aren't, doesn't* and *didn't*	Inversion of copula and subject in yes/no questions; adds *when* and *how*	Infinitive phrases appear at the ends of sentences	
Stage V	41 – 46	Indirect object appears in subj + aux + verb + ind obj + obj	Adds *wasn't, wouldn't, couldn't*, and *shouldn't*; negative appears with copula in subj + copula + neg.	Adds modals; stabilizes inverted auxiliary; some adultlike tag questions appear	Relative clauses appear in object position; multiple embeddings appear by late stage V; infinitive phrases with same subject as the main verb	Clausal conjoining with *if* appears; three-clause declaratives appear
Post - V	47+		Adds indefinite forms *nobody, no one, none*, and *nothing;* has difficulty with double negatives	Questions other than one-word *why* questions appear; negative interrogative beyond age 5	Gerunds appear; relative clauses attached to the subject; embedding and conjoining appear with same sentence	Clausal conjoining with *because* appears with *when, but,* and *so;* embedding and conjoining appear within the same sentence

Appendix 2
Components of Language Competence (Bachman, 1990: 87)

Proficient User	**C2**	Can understand with ease virtually everything heard or read. Can summarise information from different spoken and written sources, reconstructing arguments and accounts in a coherent presentation. Can express him/herself spontaneously, very fluently and precisely, differentiating finer shades of meaning even in more complex situations.
	C1	Can understand a wide range of demanding, longer texts, and recognise implicit meaning. Can express him/herself fluently and spontaneously without much obvious searching for expressions. Can use language flexibly and effectively for social, academic and professional purposes. Can produce clear, well-structured, detailed text on complex subjects, showing controlled use of organisational patterns, connectors and cohesive devices.
Independent User	**B2**	Can understand the main ideas of complex text on both concrete and abstract topics, including technical discussions in his/her field of specialisation. Can interact with a degree of fluency and spontaneity that makes regular interaction with native speakers quite possible without strain for either party. Can produce clear, detailed text on a wide range of subjects and explain a viewpoint on a topical issue giving the advantages and disadvantages of various options.
	B1	Can understand the main points of clear standard input on familiar matters regularly encountered in work, school, leisure, etc. Can deal with most situations likely to arise whilst travelling in an area where the language is spoken. Can produce simple connected text on topics which are familiar or of personal interest. Can describe experiences and events, dreams, hopes and ambitions and briefly give reasons and explanations for opinions and plans.
Basic User	**A2**	Can understand sentences and frequently used expressions related to areas of most immediate relevance (e.g. very basic personal and family information, shopping, local geography, employment). Can communicate in simple and routine tasks requiring a simple and direct exchange of information on familiar and routine matters. Can describe in simple terms aspects of his/her background, immediate environment and matters in areas of immediate need.
	A1	Can understand and use familiar everyday expressions and very basic phrases aimed at the satisfaction of needs of a concrete type. Can introduce him/herself and others and can ask and answer questions about personal details such as where he/she lives, people he/she knows and things he/she has. Can interact in a simple way provided the other person talks slowly and clearly and is prepared to help.

"BUILDING AN ARGUMENT"

Recommended procedures encourage alignment of examinations to CEFR at differing degrees of rigour appropriate to different testing contexts and to the extent permitted by resources and expertise available.

FAMILIARISATION WITH CEFR

SPECIFICATION	STANDARD SETTING	VALIDATION
Description et analyse de la qualité du test : o contenu général de l'examen o processus d'élaboration o correction, **notation**, résultats o analyse de **l'examen** et révision	Standardisation of judgments Training in assessing performances in relation to CEFR levels (using illustrative samples) / Training in judging the difficulty of test items in relation to CEFR illustrative items Judgment sessions	Test validity ° Content validity ° Operational aspects (in pretesting, piloting) ° Psychometric aspects
Description of the test in relation to the CEFR ° Overall estimation of examination level ° Communicative activities tested ° Aspects of communicative language competence tested ° Graphic profile of relationship of the test to CEFR	Benchmarking local performance samples to CEFR levels / Assigning local items and tasks to CEFR levels Establishing cut-off scores	Standard setting validity ° Procedural validity ° Internal validity ° External validity
Documented CLAIM of link to CEFR on the basis of specification)	Documented CLAIM (on the basis of specification and standardisation)	Documented CLAIM (confirmation on basis of empirical verific

230

Appendix 5
Marking Scale for Assessing Writing in English (2009)

	Contents	Organisation	Grammar	Vocabulary	Orthography
5 C1	The arguments are convincing, relevant, fully extended and supported.	Clear, smoothly flowing and well-structured text, appropriate organisational pattern, connectors and cohesive devices.	Consistently maintains a high degree of grammatical control of complex language.	A good command of a very broad lexical repertoire, consistently used, good command of idiomatic expressions.	Clearly intelligible writing, spelling and capitalisation consistently accurate.
4 B2	The given points are well highlighted and appropriately illustrated.	Clearly intelligible continuous writing with a variety of connectors to mark the relationships.	Shows a relatively high degree of grammatical control, minor flaws can occur, but they are difficult to spot..	A good command of a broad lexical repertoire, appropriately used.	Spelling, accurate, writing intelligible, may contain contractions.
3 B1	The main ideas are relevant, there may be a lack of focus or tendency to generalize.	Mostly continuous writing with appropriate connectors and paragraphing.	Shows some degree of grammatical control in complex structures. .	Has a good range of vocabulary that allows to discuss the problems, occasionally faulty collocation or meaning.	Intelligible writing, spelling accurate, occasional or non-systematic errors may persist.
2 A2	Develops a general description of the situation in Latvia, the arguments appear irrelevant or repetitive.	Links sentences into a connected text with simple connectors like 'and', 'but' and 'because'. Ideas lack coherence and paragraphing.	Simple structures predominate, inappropriateness reduces the communicative effect.	Limited vocabulary, there may be frequent simplifications or repetition.	Generally intelligible, spelling reasonably accurate.
1 A1	The text bears some relevance to the topic.	Links words with simple connectors like 'and', 'but' and 'because'.	Few correct grammatical structures and sentence patterns.	Has basic vocabulary to describe the reasons.	Basic and copied words accurate.
0	Not enough to evaluate.	Not enough to evaluate.	Not enough to evaluate.	Not enough to evaluate.	Not enough to evaluate.

Marking Scale for Assessing Writing in French (2009)

Contenu*	pt	Vocabulaire*	pt	Grammaire**	pt	Orthographe d'usage	pt	Pertinence, cohérence	pt
Sujet traité entièrement et développé avec originalité	8	Emploi très adapté. Richesse et variété	6	Richesse de l'usage des structures	7	Absence de fautes d'orthographe	4	Argumentation logique et convaincante	5
Développement détaillé	7	Emploi adapté. Lexique assez varié	5	Emploi correct des structures simples et complexes	6	Quelques fautes d'inattention	3	Argumentation assez convaincante	4
Développement suffisant	6	Emploi correct du vocabulaire	4	Emploi correct des structures simples. Des erreurs rares dans les structures complexes	5	Nombre de fautes limité	2	Manque d'argumentation convaincante	3
Développement suffisant mais incomplet	4	Emploi inégal du vocabulaire, quelques imprécisions	3	Emploi correct des structures simples. Des erreurs fréquentes dans les structures complexes	3	De nombreuses fautes	1	Manque de logique et d'argumentation convaincante	2
Développement insuffisant	3	Vocabulaire limité, souvent employé incorrectement	2	Structures de base connues mais erreurs fréquentes	2	De très nombreuses fautes	0	Absence d'argumentation	1
Longueur du devoir insuffisante. Le sujet n'est pas traité	1	Très pauvre. Répétitions trop fréquentes	1	Connaissances insuffisantes de la morphosyntaxe	1			Texte illogique	0
Maximum	8		6		7		4		5
Total général 30 points									

*Il est possible d'attribuer 4 points pour le **contenu** si l'étudiant présente un niveau intermédiaire (entre 3 et 5 points).

** Il est possible d'attribuer 4 points pour la **grammaire** si l'étudiant présente un niveau intermédiaire (entre 3 et 5 points).

Appendix 7
Descriptive Statistics of Subordinate Clauses

In English

	A		B		C		D		E		F	
Total number	105		240		219		212		101		40	
Mean number of sub-clauses per sentence	1.46		1.46		1.3		1.38		1.3		1.3	
	Nr.	%	Nr.	%	Nr.	%	Nr.	%	Nr.	%	Nr.	%
Total number of mono-sub-clauses	**72**	**69**	**161**	**67**	**157**	**72**	**145**	**68**	**74**	**73**	**29**	**73**
Noun clauses	27	38	47	29	52	33	57	39	26	35	11	38
Adverbial clauses	16	22	50	31	64	41	35	24	33	46	15	52
Adjectival clauses	29	40	64	40	41	26	53	37	15	20	3	10
Total number of multi sub-clauses	**33**	**31**	**79**	**33**	**62**	**28**	**67**	**32**	**27**	**27**	**11**	**27**
Number of sub-clauses in multi sub-clauses	82		190		137		148		60		23	

In French

	A		B		C		D		E		F	
Total number	53		234		183		4					
Mean number of sub-clauses per sentence	1.4		1.5		1.3		1					
	Nr.	%	Nr.	%	Nr.	%	Nr.	%	Nr.	%	Nr.	%
Total number of mono-sub-clauses	**34**	**64**	**147**	**63**	**142**	**78**	**4**	**100**				
Noun clauses	10	29	50	34	46	32	4	100				
Adverbial clauses	15	44	68	48	65	46						
Adjectival clauses	9	26	29	20	31	22						
Total number of multi sub-clauses	**19**	**36**	**87**	**37**	**41**	**22**						
Number of sub-clauses in multi sub-clauses	41		203		92							

	Contents	Organisation	Grammar	Vocabulary	Orthography
5 C1	The arguments are convincing, relevant, fully extended and well supported.	Clear, smoothly flowing and well-structured text, appropriate organizational pattern, connectors and cohesive devices.	Consistently maintains a high degree of grammatical control of complex language.	A good command of a very broad lexical repertoire, consistently used, good command of idiomatic expressions.	Clearly intelligible writing, spelling and punctuation consistently accurate. May occur a slip of the pen.
4 B2.2	The given arguments are varied, detailed, well highlighted and appropriately supported.	Clearly intelligible continuous writing with a variety of connectors to mark the relationships and paragraphs.	Shows a relatively high degree of grammatical control, minor flaws may occur, but they are difficult to spot.	A good command of a broad lexical repertoire, appropriately and flexibly used to avoid repetition.	Spelling accurate, writing intelligible, may contain a few mistakes which are difficult to spot.
3 B2.1	The main ideas are relevant; there may be lack of focus or precision in supporting the material.	Mostly continuous writing with appropriate connectors and paragraphing.	A mix of simple and complex sentence forms is used with occasional inaccuracies.	A good range of vocabulary, appropriate for the task, occasionally faulty collocation or meaning.	Intelligible writing, spelling mostly accurate, occasional or non-systematic errors may persist.
2 B1	Some main ideas are put forward but incompletely developed, the arguments may appear irrelevant, unclear or repetitive.	Links sentences into a connected text with simple connectors like 'and', 'but' and 'because'. Ideas may lack coherence and appropriate paragraphing.	Simple structures with errors predominate; errors and inappropriacies do not reduce the communicative effect.	Sufficient vocabulary to complete the task, there may be inappropriate word choice, frequent simplifications or repetition.	Spelling is sufficiently accurate to be understood most of the time.
1 A2	The text bears some relevance to the topic.	Ideas partly organized, generally intelligible.	Simple structures with errors predominate, occasionally reduce the communicative effect.	Limited vocabulary with faulty collocations and word choice occasionally reduce the communicative effect.	Spelling is often inaccurate but the ideas can be followed throughout the text.
0	Not enough to evaluate.				

VISC   Vaļņu ielā 2, Rīga, LV - 1050

234

Marking Scale for Assessing Writing in French (2013)

	Réalisation de la tâche	Organisation	Vocabulaire	Morphosyntaxe	Orthographe
5 C1	Ecrit bien argumenté en confirmant un point de vue de manière élaborée par l'intégration d'arguments secondaires, de justifications et d'exemples pertinents.	Ecrit clair, fluide et bien structuré, démontrant un usage contrôlé des outils d'organisation, d'articulation et de cohésion du discours pour parvenir à une conclusion appropriée.	Vaste répertoire lexical permettant de surmonter sans recherche apparente ses lacunes. Bonne maîtrise d'expressions idiomatiques. De petites bévues occasionnelles.	Richesse de structures permettant de varier la formulation. Maintenance constante d'un haut degré de correction. Erreurs sont rares et difficiles à repérer.	Exactitude orthographique à l'exception de quelques lapsus.
4 B2.2	Ecrit argumenté soulignant d'une manière appropriée des points importants et des détails pertinents en expliquant des avantages ou les inconvénients de différentes opinions.	Ecrit structuré reliant clairement les idées exprimées sous forme d'un texte fluide et cohérent. Variété des articulateurs logiques.	Gamme étendue en dépit de lacunes lexicales ponctuelles entraînant l'usage de périphrases.	Constructions variées. Bon contrôle morphosyntaxique employant des phrases complexes en dépit d'erreurs non systématiques.	Bon contrôle orthographique. Des erreurs dues à l'inattention.
3 B2.1	Ecrit argumenté soulignant d'une manière appropriée des points importants et des détails pertinents.	Ecrit structuré reliant clairement les idées exprimées sous forme d'un texte fluide et cohérent. Des articulateurs logiques bien présents.	Vocabulaire généralement approprié bien que des confusions et le choix de mots incorrects se produisent sans gêner la compréhension	Bon contrôle morphosyntaxique employant des phrases simples, des erreurs sans conduire à des malentendus	Bon contrôle orthographique, des erreurs non systématiques, dues à l'influence de la langue maternelle.
2 B1	Ecrit adéquat au sujet proposé. Faits et événements exposés de manière claire.	Ecrit structuré et cohérent reliant une série d'éléments courts, simples et distincts en un discours qui s'enchaîne.	Bonne maîtrise du vocabulaire élémentaire, employé généralement correctement dans l'expression de pensées complexes.	Bonne maîtrise de la structure de la phrase simple et complexe les plus courantes malgré de nettes influences de la langue maternelle.	Assez juste pour être suivie facilement le plus souvent.
1 A2	Ecrit adéquat au sujet, rédigé d'une manière simple en communiquant des informations assez développées mais avec une précision limitée.	Ecrit simple et cohérent. Des énoncés liés avec les articulations les plus fréquentes.	Répertoire limité mais adéquat, emploi des mots et des expressions courantes en suivant les usages de base, des répétitions peuvent se produire.	Structures et formes grammaticales simples avec la présence systématique d'erreurs élémentaires.	Relative exactitude phonétique mais pas forcement orthographique.
0	La production est insuffisante pour être évaluée.				

Different Types of Multi Sub-Clauses in English and French

Types of multi sub-clauses	English						French			
	A	**B**	**C**	**D**	**E**	**F**	**A**	**B**	**C**	**D**
<N,N>	3	3	3	4			1	5		
<N,ADV>	8	11	15	16	8	4	6	12	14	
<ADV,N>	1	8	1	2		1		6	3	
<N,N,N>	1					1				
<N,ADV,ADV>	2		2					3	2	
<ADV,ADV,N>								1		
<ADV,N,ADV>	2									
<N,ADV,N>	1	1	1		2			2		
<ADV,ADV>	1	4		3	2		1	7	3	
<ADV,N,N>	1			2						
<N,N,ADV>		4								
<ADV,N,ADV>			1	2				1		
<N,ADV,N,N>		1						1		
<N,N,ADV,ADV>								1		
<N,ADV,ADV,ADV>		1								
<ADV,N,ADV,ADV>		1								
<ADV,ADV,ADV>			1							
<ADJ,ADV,N>	1									
<N,ADJ,ADV>			1	2				1		
<N,ADV,ADJ>			2					4	3	
<ADJ,N,ADV>					1				1	
<N,ADJ>	2	10	19	13	6	2	1	13	2	
<ADJ,N>	1	5	3	4	3	1	4	4		
<N,N,ADJ>	1		1							
<N,ADJ,ADJ>	1	2			1		1	2	1	
<ADJ,N,N>				1	1			2		
<N,ADJ,N>		1								
<N,ADV,ADV,ADJ>			1							
<N,ADV,ADJ,ADV>							1			
<N,ADJ,N,N>				1						
<ADV,N,ADV,ADJ>			1							
<N,ADJ,N,ADJ,ADV>		1								
<N,ADV,ADJ,AD,ADJ>		1								
<N,ADJ,ADJ,ADV,N,ADV>				1						
<N,ADV,ADV,ADV,ADV,N>								1		
<ADV,ADJ,N,ADJ,ADJ>		1								
<ADV,N,ADJ,ADJ,ADV>								1		
<ADJ,ADJ>	2	7	2	2			2	3	3	
<ADJ,ADV>	4	6	2	6	1	1		4	2	
<ADV,ADJ>	4	5	3	7	1	1	2	11	5	
<ADJ,ADJ,ADJ>		1								
<ADV,ADJ,ADV>		1								
<ADJ,ADV,ADV>		2		1					1	
<ADV,ADJ,ADV,ADJ>									1	
<ADV,ADV,ADJ>					1					
<ADJ,ADV,ADJ>								2		

Appendix 11
Script 3

Year 12 Examination in English 2013
Task 2

You should spend about 45 minutes on this task.
Write between 200 – 250 words.
FILMS VERSUS BOOKS
A lot of young people nowadays think that it is much better to watch a film version of a book than to read it. Write an essay expressing your opinion and providing reasons to support your point of view.

Script 3 (289 words)

Nowadays many books have their film version. The films are not just for those who have not read the book, they are for everyone. But it makes a problem for young people they have to make choise what to do read the book or watch the film.

Some people think that reading a 500 page book is a waste of time if you can watch a movie in an hour and a half. There also are young people who are doing many extra activities during the day that they don't have so much time for reading a book.

There also are people who can't concentrate their selves for reading a book because mostly in the books there are pages full of letters and with no pictures. These type of people chose to watch movies because films give them visual and sound effects.

But the main reason could be that we are living in technology age and it is in our mind that everything has be understandable easy. It doesn't take much concentration to watch a movie as it takes to read a book because when you read a book you have to think about every page every chapter and then you will come to conclusion not as it is when you watch the film, watch it till the end and you get the idea not thinking about some scenes that were in the movie.

So if we want for young people to read books, not to watch movies we have to train their concentration skills. We also can remind them that not every book and film are the same, for example, 'Captain Corelli's Mandolin' the end of the book is not the same as it is in the movie.

Expert's assessment

Contents	3
Organisation	3
Grammar	3
Vocabulary	3
Orthography	3
Total	15

Appendix 12
Syntactic Criterial Feature Validation Results

Nr. of assessor	Script 3		Nr. of assessor	Script 3	
	Assessment 1	Assessment 2		Assessment 1	Assessment 2
1	3	3	60	3	3
2	3	3	61	4	4
3	3	3	62	3	3
4	3	3	63	3	3
5	3	3	64	3	3
6	4	4	65	3	3
7	4	4	66	4	3
8	3	3	67	3	3
9	4	4	68	3	3
10	3	3	69	3	3
11	3	3	70	3	3
12	2	2	71	3	3
13	2	2	72	3	3
14	2	2	73	3	3
15	3	3	74	3	3
16	4	4	75	4	4
17	3	3	76	3	3
18	3	3	77	3	3
19	3	3	78	3	3
20	3	3	79	3	3
21	3	3	80	3	3
22	3	3	81	3	3
23	3	3	82	3	3
24	3	3	83	3	3
25	3	3	84	4	4
26	3	3	85	2	2
27	2	2	86	3	3
28	3	3	87	2	2
29	4	4	88	2	2
30	3	3	89	3	3
31	3	3	90	3	3
32	3	3	91	3	3
33	3	3	92	3	3
34	4	4	93	3	3
35	3	3	94	2	2
36	3	3	95	3	3
37	5	5	96	2	2
38	4	4	97	4	4
39	3	3	98	3	3
40	4	4	99	3	2
41	4	4	100	3	3
42	4	4			
43	4	4			
44	4	4			
45	4	4			
46	5	5			
47	4	3			
48	2	2			
49	3	3			
50	3	3			
51	3	3			
52	3	3			
53	3	3			
54	3	3			
55	3	3			
56	2	2			
57	4	4			
58	3	3			
59	4	4			