



universität
wien

DIPLOMARBEIT

Titel der Diplomarbeit
„Aspects of the English Verb in CLIL Classroom
Discourse“

Verfasserin
Silvia Tiefenbacher

angestrebter akademischer Grad
Magistra der Philosophie (Mag.phil.)

Wien, 2012

Studienkennzahl lt. Studienblatt:	A 190 299 344
Studienrichtung lt. Studienblatt:	Lehramtstudium UF Psychologie und Philosophie, UF Englisch
Betreuerin:	Ao. Univ. Prof. Mag. Dr. Christiane, Dalton-Puffer

Declaration of Authenticity

I confirm to have conceived and written this paper in English all by myself. Quotations from other authors and any ideas borrowed and/or passages paraphrased from the works of other authors are all clearly marked within the text and acknowledged in the bibliographical reference.

Vienna, May 2012

Eidesstattliche Erklärung

Ich versichere,

1. dass ich die Diplomarbeit selbstständig verfasst, andere als die angegebenen Hilfsmittel nicht benutzt und mich auch sonst keiner unerlaubten Hilfe bedient habe.
2. dass ich diese Diplomarbeit bisher weder im Inland noch im Ausland in irgendeiner Form als Prüfungsarbeit vorgelegt habe.

Vienna, May 2012

Acknowledgments

I want to thank my parents who supported me and helped me to reach my goals, my boyfriend who was always there for me when I thought that I could never make it.

Furthermore, I have to thank my friends for their understanding and for putting some pressure on me, especially Florian, Michael, Nicola and Petra who proofread my thesis. Special thanks go to my supervisor Christane Dalton-Puffer for her guidance and support.

Table of Content

		Page
	List of Tables	v
	List of Diagrams	vii
	List of Abbreviations	vii
	Introduction	1
1	The English Verb	5
1.1	Verb Phrases	5
1.2	Verb Classes and Functions	7
1.2.1	Semantic Domains.....	8
1.2.2	The Lexical Verb.....	11
1.2.3	The Primary Verb.....	13
1.2.4	The Auxiliary Verb	15
1.3	Distinctions in the Verb Phrase	18
1.3.1	Tense.....	19
1.3.2	Aspect.....	29
1.3.3	Voice.....	36
2	Corpus Analysis	39
2.1	What is a Corpus?	39
2.2	Corpus Linguistics	41
2.3	Relevant Corpus Use	42
2.4	British National Corpus	43
2.4.1	Frequency of Verbs in General.....	45
2.4.2	Frequency of Grammatical Classes.....	46
2.4.3	Differences between Conversation and Task-oriented Speech.....	49
2.5	Longman Spoken and Written English Corpus	52
2.5.1	Distribution of Lexical, Modal and Primary Auxiliary Verbs.....	53
2.5.2	LSWE Findings on Semantic Domain Distribution.....	54
2.5.3	Most common Lexical Verbs.....	55
2.5.4	Frequency of Tense and Modality.....	56
2.5.5	Corpus Findings on Aspect.....	57
2.5.6	Frequency of Active and Passive Voice.....	58
3	Research Methodology in Applied Linguistics	60
3.1	An Introduction to Research Methodology	60
3.2	The Roots of Quantitative Research	61
3.3	Main Features of Quantitative Data Research	62
3.4	The Advantages and Disadvantages of Quantitative Research	64
3.5	How to do Empirical Investigation using Quantitative Research	65
3.5.1	Data Collection.....	66
3.5.2	Data Analysis.....	67

	3.5.4	Reliability and Validity.....	68
4		Data Analysis.....	70
	4.1	Objectives.....	70
	4.2	HIST 1.....	71
	4.2.1	Word Class Distribution.....	71
	4.2.2	Primary Verb Distribution.....	72
	4.2.3	Modal Verb Distribution.....	81
	4.2.4	Distribution of Modal Verb	82
	4.2.5	Summary of Findings.....	83
	4.3	HIST 2.....	85
	4.3.1	Word Class Distribution.....	85
	4.3.2	Primary Verb Distribution.....	86
	4.3.3	Distribution of Modal Verb	96
	4.3.4	Lexical Verb Distribution.....	97
	4.3.5	Semantic Meaning and Lexical Verb Distribution.....	98
	4.3.6	Summary of Findings.....	100
	4.4	MARK 1.....	103
	4.4.1	Word Class Distribution.....	103
	4.4.2	Primary Verb Distribution.....	104
	4.4.3	Distribution of Modal Verb	112
	4.4.4	Lexical Verb Distribution.....	113
	4.4.5	Summary of Findings.....	114
	4.5	MARK 2.....	115
	4.5.1	Word Class Distribution.....	115
	4.5.2	Primary Verb Distribution.....	116
	4.5.3	Distribution of Modal Verb	124
	4.5.4	Lexical Verb Distribution.....	125
	4.5.5	Semantic Meaning and Lexical Verb Distribution	127
	4.5.6	Summary of Findings.....	128
	4.6	Differences between History and Marketing Lessons.....	131
		Conclusion	133
		Bibliography	134
		Zusammenfassung	137
		Abstract	139
		Curriculum Vitae	141

List of Tables

Table Number	Page
Table 1. Categorization of Semantic Domains.....	11
Table 2. English Verb Forms.....	12
Table 3. Regular and Irregular Verb Forms.....	12
Table 4. Classification of Verb Forms.....	13
Table 5. Auxiliary Verb List.....	16
Table 6. Modal Auxiliary Combinations.....	17
Table 7. Representation of the Past Tense on a Timeline.....	25
Table 8. Representation of the Present Progressive with Future Meaning.....	30
Table 9. Representation of Present Progressive with Future Meaning and Past Tense Implication.....	31
Table 10. Representation of Past Progressive with no Reference Point.....	32
Table 11. Representation of the Past Tense including a Reference Point.....	33
Table 12. Representation of the Past Progressive use in a Timeframe.....	33
Table 13. Representation of the Present Perfect Simple.....	34
Table 14. Distribution of Verb classes in HIST 1.....	71
Table 15. Primary Verb Distribution in HIST 1.....	72
Table 16. Distribution of BE, DO and HAVE in HIST 1.....	74
Table 17. Distribution of BE as Auxiliary and Main verb in HIST 1.....	74
Table 18. Verb form Distribution of BE in HIST 1.....	75
Table 19. Grammatical Functions of BE in HIST 1.....	76
Table 20. Distribution of DO as Auxiliary and Main Verb in HIST 1.....	77
Table 21. Verb Form Distribution of DO in HIST 1.....	77
Table 22. Grammatical Functions of DO in HIST 1.....	78
Table 23: Distribution of HAVE as Auxiliary and Main Verb in HIST 1.....	79
Table 24: Distribution of HAVE's Verb Forms in HIST 1.....	80
Table 25: Grammatical Functions of HAVE in HIST 1.....	81
Table 26: Distribution of Modal Verbs in HIST 1.....	82
Table 27: Distribution of Main Verb Forms in HIST 1.....	83
Table 28: Distribution of Grammatical Functions in HIST 1.....	84
Table 29: Verb Class Distribution in HIST 2.....	86
Table 30: Auxiliary and Main Verb Distribution of Primary Verbs in HIST 2.....	87
Table 31: Distribution of BE, DO and HAVE in HIST 2.....	87
Table 32: Share of Auxiliary and Main Verb Use of BE in HIST 2.....	88
Table 33: Verb Form Distribution of BE in HIST 2.....	88
Table 34: Grammatical Functions of BE in HIST 2.....	90
Table 35: Share of DO as an Auxiliary and Main Verb in HIST 2.....	91
Table 36: Verb Form Distribution of DO in HIST 2.....	91
Table 37: Grammatical Functions of DO in HIST 2.....	93

Table 38: Auxiliary and Main Verb Share of HAVE in HIST2.....	94
Table 39: Verb Form Distribution of HAVE in HIST 2.....	94
Table 40: Grammatical Functions of HAVE in HIST 2.....	95
Table 41: List of Modal Verbs and their Distribution in HIST 2.....	96
Table 42: Lexical Verb Form Distribution in HIST 2.....	97
Table 43: Semantic Domain Distribution in History Lessons Compared to the LSWE.....	98
Table 44: Most frequently used Verbs in History Lessons, LSWE and BNC.....	99
Table 45: Distribution of Tense, Aspect, Mood and Passive Voice in HIST 2.....	101
Table 46: Verb Class Distribution in MARK 1.....	103
Table 47: Auxiliary and Main Verb Distribution of Primary Verbs in MARK 1.....	104
Table 48: Distribution of BE, Do and HAVE in MARK 1.....	105
Table 49: Auxiliary and Main Verb Share of BE in MARK 1.....	106
Table 50: Verb Form Distribution of BE in MARK 1.....	106
Table 51: Grammatical Functions of BE in MARK 1.....	107
Table 52: Auxiliary and Main Verb Share of DO in MARK 1.....	108
Table 53: Verb Form Distribution of DO in MARK 1.....	108
Table 54: Grammatical Functions of DO in MARK 1.....	109
Table 55: Auxiliary and Main Verb Share of HAVE in MARK 1.....	110
Table 56: Verb Form Distribution of HAVE in MARK 1.....	110
Table 57: Grammatical Functions of HAVE in MARK 1.....	111
Table 58: Modal Verb Distribution in MARK 1.....	112
Table 59: Verb Form Distribution of Lexical Verbs in MARK 1.....	112
Table 60: Distribution of Tense, Aspect, Mood and Passive Voice in MARK 1.....	115
Table 61: Verb Class Distribution in MARK 2.....	116
Table 62: Auxiliary and Main Verb Distribution in MARK 2.....	117
Table 63: Distribution of BE, DO and HAVE in MARK 2.....	118
Table 64: Auxiliary and Main Verb Distribution of BE in MARK 2.....	118
Table 65: Verb Form Distribution of BE in MARK 2.....	119
Table 66: Grammatical Functions of BE in MARK 2.....	120
Table 67: Auxiliary and Main Verb Distribution of DO in MARK 2.....	121
Table 68: Verb Form Distribution of DO in MARK 2.....	121
Table 69: Grammatical Functions of DO in MARK 2.....	122
Table 70: Auxiliary and Main Verb Distribution of HAVE in MARK 2.....	123
Table 71: Verb Form Distribution of HAVE in MARK 2.....	123
Table 72: Grammatical Functions of HAVE in MARK 2.....	124
Table 73: List of Modal Verbs and their Distribution in MARK 2.....	125
Table 74: Lexical Verb Form Distribution in MARK 2.....	126
Table 75: Semantic Domain Distribution in HIST 1 and 2 Compared to the LSWE.....	127
Table 76: Most Common Lexical Verbs in MARK 1 and 2 Compared to the LSWE and BNC.....	128
Table 77: Tense, Aspect, Mood and Passive Voice in MARK 1 and MARK 2.....	130

List of Diagrams

Diagram Number	Page
Diagram 1: List of most frequently used Verbs (BNC).....	45
Diagram 2: Frequency of the Forms of BE.....	46
Diagram 3: Frequency of the Forms of DO and HAVE.....	47
Diagram 4: Frequency of Lexical Verb Forms (BNC).....	48
Diagram 5: Difference of the Uses of BE in Task-oriented speech and conversations (BNC).....	49
Diagram 6: Difference of the Uses of DO in Task-oriented speech and conversations (BNC).....	50
Diagram 7: Difference in the Use of HAVE forms in task-oriented speech and conversations	51
Diagram 8: Difference in the Use of Lexical Verb forms in Task-oriented speech and conversations.....	51
Diagram 9: Frequency of Verb Classes in general and in conversations (LSWE).....	54
Diagram 10: Semantic Domain Distribution (LSWE).....	55
Diagram 11: List of the most Common Lexical Verbs (LSWE).....	55
Diagram 12: Distribution of Tense and Mood (LSWE).....	56
Diagram 13: Distribution of Aspect (LSWE).....	57
Diagram 14: Voice Distribution (LSWE).....	59
Diagram 15: Illustration of Quantitative Research.....	65

List of Abbreviations

HIST 1	First History Lesson analyzed
HIST 2	Second History Lesson analyzed
MARK 1	First Marketing Lesson analyzed
MARK 2	Second History Lesson analyzed
-s form	Verb form with –s inflection
-ing form	-ing participle
-ed form	Past tense form
-en form	Past participle

Introduction

Next to nouns and adjectives verbs belong to the most important word classes. Verbs are used to express actions, facts and feelings. English language learning has to ensure that the learners are able to articulate their thoughts. Furthermore they should know and be able to use those verbs which are frequently used in speech. Which verbs and grammatical structures are most frequent in the English language can be learned by investigating corpus findings. By comparing these to the results of classroom discourse analysis one can see if students know and use the most common verbs and grammatical functions. The thesis 'The English Verb in CLIL Classroom Discourse' is an empirical and theoretical analysis of English Verb occurrence in CLIL lessons.

The thesis aims at an analysis of speech events in CLIL classroom discourse. CLIL is the acronym for Content and Language Integrated Learning. It is an educational approach in which content is taught by using a different language than the student's mother tongue. In the field of language teaching CLIL is an approach which seeks to improve the language skills of students not by language teaching but by using the second language (L2) as a tool to convey the content of other subjects, through this students have the opportunity to acquire and not to learn the L2. (Dalton-Puffer 2007:10) CLIL is not yet fully defined as an educational model. In Lorenzo and Moore's article "On the natural emergence of language structure in CLIL" rationales for linguistic development are stated. These are based on Wesche and should, in my opinion, be the starting point of language learning. They are also taken into consideration when articulating CLIL aims. These rationales say that children and teenagers are able to learn languages by accident. In order to become-fluent it is necessary to be exposed to the L2 for a longer period of time, and language should not be taught in L2 language class but it should be made available to learners in other classes. (2010:23-24)

According to the CLIL Compendium cited in Dalton-Puffer (2007: 10) CLIL aims at:

- an improvement of competence in the used language,
- an improvement of speaking skills,
- a deeper consciousness of L1 and L2 respectively,
- a development of plurilingual interest and attitude.
- and the introduction of a target language.

From Wesche's rationales and the aims constituted by the CLIL compendium it can be deduced that this method is based on the assumption that language proficiency is easier obtained by learning in context and communication when utilizing it. For example, students who spend a semester abroad are not taught how to speak but are forced to apply the L2 in order to communicate. CLIL tries to simulate such a setting by teaching content in another language than the student's mother tongue. (Dalton-Puffer 2007: 2-4) Knowledge and skills are gained because they are essential for communication and as they are trained. In social interaction language is required and practiced at the same time. (Dalton-Puffer 2007:8) Dalton-Puffer (2007:8) mentions Brown, Collins and Duguid's concept of cognitive apprenticeship which claims that a student's learning process is supported by authenticity in gaining, developing and using cognitive tools. Learning is, thus, achieved by social interaction and the construction of knowledge. This theory and the one established by Vigotsky share some similarities. Vigotsky is the inventor of the socio-cognitive development theory. He says that "social interaction plays a fundamental role in the development of cognition" (qtd. in Dalton-Puffer 2007:8). According to Vigotsky learning only takes place after an interaction with one's social environment. The main point in both theories, therefore, is that language learning is only possible by using the language and not by learning grammatical or syntactic rules.

Using English as a tool to teach another subject seems to belong to the field of bilingualism. The difference between CLIL and bilingualism is that bilingual education is necessary in countries, such as Canada, where people are to be proficient in two languages, but CLIL aims at using a foreign language which is

not spoken in the country as the L2. Students only come across the L2 at school or when on vacation or on a semester abroad. The teacher in CLIL lessons is neither a native speaker nor a foreign language expert but a content expert. While the content of language classes is to teach grammatical, syntactical and cultural issues the content of CLIL classes is the content of the content subject. Next to the CLIL classes there are separate language classes taught by foreign language specialists. Basically, this means that CLIL is content teaching and not language teaching which aims to enhance foreign language proficiency. Supporters of CLIL claim that CLIL intends to fuse language and subject learning.

The approach of *Content Integrated Language Teaching*, which after a rise of importance, is today also encouraged by the European Union, based on its proclamation that every citizen should be able to be proficient in at least two languages spoken in the European Union. (qtd. in Dalton-Puffer 2010: 4-5) The European Commission states that pupils on the secondary level should have the possibility to be taught in the first L2 they learned. But the European Union did not issue an instruction and all the countries of the Union can decide for themselves if they want to implement CLIL in their schools or not. Very few countries like Spain, The Netherlands, Germany and Austria do. (Dalton-Puffer 2010:4) It is an achievement that it has been agreed that the *European CLIL program curriculum* should both be language driven, enhancing language skills, and content driven, meaning to deliver subject information. Although both fields should be equally enforced in Europe CLIL is rather content driven. (Dalton-Puffer 2010:1-2)

In Austria this approach has experienced a significant increase since the 1990's. Not only the theory that 'learning by doing' influenced this hub, but also internationalization and globalization. These have made a reform in education necessary. Proficiency in another language than the learner's L1 is an asset in the competition on the job market. (Dalton-Puffer 2007:1, 2010:4) The ability to speak different languages has ever since the antiquity been a sign of belonging to the elite. This goes hand in hand with being taught in another, especially a prestigious, language. In former times the elite was taught in Latin. With the English language having become one of the most spoken languages and most

influential languages it has replaced Latin as the language one should be educated in. (Dalton-Puffer 2010:3) By introducing some research on CLIL the benefits of this approach are emphasized.

Previous research has shown that CLIL undoubtedly has its advantages. In Tarja Nikula's article "Effects of CLIL on a teacher's classroom language use" one learns which positive effects CLIL has on the student's classroom behavior. In Nikula's article also the findings of a research conducted in a Finish biology class are discussed. It has been found out that in CLIL turn-taking is more common than in traditional lessons. Students even show more initiative to take part in classroom discourse in CLIL lessons. (Dalton-Puffer 2010:110) Another interesting research was done by Hüttner and Rieder-Bühneman it is called "A cross-sectional analysis of oral narratives by children with CLIL and non-CLIL instruction". It has shown that CLIL students have both quantitative and qualitative advantages in language competence. There is less switching to the L1 and the students try to convey information by explaining it in the L2. Also their lexical and grammatical accuracy is higher as the one of students not being instructed in the L2. (Dalton-Puffer 2010:77) The introduced researches have shown the benefits of CLIL on classroom participation and advantages in competence. This thesis aims at giving evidence of the advantages on language proficiency by investigating the use of the English verb. The thesis is divided into a theoretical part and empirical part.

The theoretical part involves a detailed description of the English verb classes as well as of verb forms and grammatical functions. The first chapter includes an investigation of modal, lexical and primary verbs. Another concern is the description of tense, aspect, mood and voice. The second chapter gives an account on the distribution of verb classes, grammatical structures, verb forms and semantic domains, which were introduced in the first chapter. The numbers are taken from the Longman Corpus of Spoken and Written English and the British National Corpus. The third chapter leads to the empirical part stating how research is done and which methods should be applied in order to make the research outcome relevant and valid. The final part is devoted to the empirical study in which the outcome of the data analysis is illustrated and compared to the corpus findings of both the LSWE and BNC.

1 The English Verb

The aim of this thesis is to extract data on verb forms used from CLIL classroom discourse transcripts. In this chapter theoretical background information on the in the data analysis investigated verb forms and grammatical forms is given. The following lines will give an account of what exactly is analyzed.

First of all the individual verb forms will be described, which are main or lexical verbs, primary verbs and modal verbs. As English verbs may have more than one meaning a chapter on their different semantic meanings is included. After this introduction a detailed theoretical account of finite verbs follows as they, in contrast to non-finite verbs, can be marked for tense and aspect. A major part of the theoretical analysis is devoted to these concepts. But not only tense and aspect are of interest in this research but also voice which is, the last point, taken into account in the data analysis.

1.1 Verb Phrase

In order to reach a full understanding of the English verb it necessary to take the analysis of the English verb phrase as a starting point. The most important element of a clause is the verb. It is most central as it expresses what is happening. It refers to actions, affairs, events and states. (Aarts 2011: 65) As verb phrases often consist of more than one verb a distinction has to be made. The first verb in a verb phrase is called operator and is crucial for the phrase's meaning. The operator is always the first or the only auxiliary verb in the phrase. The operator is important as it is used to form certain sentence structures, such as questions and negations, emphasis and it also may substitute predicates. (Greenbaum 1991: 20-21) Verbs can either be transitive or intransitive. Transitive verbs take direct objects, for example:

1. *She just wants a change* (Aarts 2011: 65)

Wants is the transitive verb and *a change* is the direct object referring to the

verb. It is impossible to utter the sentence leaving out the direct object, which holds important information for the meaning of the sentence. Intransitive verbs, such as *arrive*, *dance*, *joke*, and *walk* do not require a direct object but can stand on their own. (Aarts 2011: 65- 66)

Just as in sentence 2, from Aarts (2011:65):

2. *Tell me when you will arrive.*

Although intransitive verbs do not need a direct object or adjunct they can have one optionally. In the sentence above an optional adjunct would be where the person will arrive. (Aarts 2011: 66)

A further important fact is that verbs have finite and non-finite verb forms. The former are marked for tense but the latter are not. The verb *go* has finite and non-finite verb forms. Finite verb forms are *go*, *goes*, and *went*. Non-finite verb forms are *going* and *gone*. Most verbs have 5 forms, including finite and non-finite forms. However, some have more such as the verb *be* which functions as lexical, primary and modal verb. Modal auxiliary verbs like *must*, *ought to*, *dare* and *need* have only one form, which is finite. (Palmer 1974:12-13)

Before giving account of verb phrase construction it has to be stated that “there are certain very limited features of concord or agreement of verbal form with the subject of the sentence.” (Palmer 1974: 14) Palmer states that there are only three types of concord and only one of them applies for all verbs. The concord which is valid for all verbs says that all verbs of the English language, except modal verbs, have two present tense forms, one of them, the –s form, is used after *he*, *she* *it* and singular noun phrases. The simple form is used after all the other pronouns and plural noun phrases. These two verbal forms cannot be defined as being exclusively singular or plural. The verb *be* to the second and third concord. These concords say, firstly that the verb *be* has two past tense forms, *was* and *were*, *was* is singular and *were* plural. Secondly that *be* is the only verb in the English language with a form for the first person singular of the present tense, namely *am*. (1974:14)

The English verb phrase consists of a main verb, which is either a lexical verb or a primary verb (Biber, Johansson, Leech, Conrad & Finegan 1999: 99), and no more than four auxiliary verbs, which belong to different subclasses of auxiliaries. It is uncommon that all four possible auxiliary verbs appear in one verb phrase. Further information on the appearance of auxiliaries will be given in 1.1.2. For now it has to suffice that auxiliaries always precede the main verb and follow a certain sequence. (Greenbaum 1991: 48) Verb phrases can either be marked for tense or not. If they are tensed then verb phrases are finite, if they are not marked for tense, then they are referred to as being non-finite. A finite verb phrase can additionally be marked for voice and aspect. Non-finite verb phrases cannot be modal but can be marked for perfective aspect and may have passive forms. Verb phrases consist of the main verb and elements which belong to it like objects or predicatives. (Biber, Johansson, Leech, Conrad & Finegan 1999: 99-100) Detailed information on the main and primary verb, the functions of the auxiliary and semantic meaning of lexical verbs are given in the following chapters.

1.2 Verb classes and functions

Verbs not only have different functions but also can be classified into three major categories, lexical or full verbs, primary verbs and modal verbs. These are classified by their role as either main or auxiliary verbs. Lexical verbs consist of an open class of words whose only function is that of a main verb. There are three primary verbs, *have*, *do* and *be*, which can either be main or auxiliary verbs. Finally, modal verbs can only be used as auxiliary verbs helping the main verb to express meaning or mark aspect and voice. Further distinctions between verbs which are distinguished by their semantic domains can be made as well as distinctions between simple lexical verbs and multi-word verbs like phrasal verbs. (Biber, Johansson, Leech, Conrad & Finegan 1999: 358) For this thesis only main and modal verbs as well as their semantic role are of importance. Multi-word verbs are left out as the data analysis does not aim at giving an account of the distribution of single- word and multi-word verb occurrences. Also a theoretical discussion on why verbs collocate with certain prepositions and what phrasal verbs are is of no interest here.

1.2.1 Semantic domains

Semantics is the study of meaning of not only single verbs but also whole sentences. It is a field of linguistic research which is central to studying the human mind and communication. People convey meaning and express their experiences through using a language. (Leech 1974: 1-5) For this thesis only the semantic meaning of verb forms is of interest. Biber et al subdivide verbs into seven major semantic domains. Their classification is based on the core meaning of verbs. Those domains are: *activity verbs*, *communication verbs*, *causative verbs*, *verbs of simple occurrence*, *verbs of existence or relationship* and *aspectual verbs*. It has however to be stated that many English verbs have a whole variety of meanings they can occupy. Therefore many verbs can not be assigned to only one semantic field and their classification is difficult. This is especially common in activity verbs. For example the verbs *contact* and *raise*, they can either refer to physical activities or communicative acts. The verbs *admit* and *consult* can even belong to three different domains: physical, communicative and mental. (Biber, Johansson, Leech, Conrad & Finegan 1999: 361) The following account on the semantic meaning of verbs is based on information found in Biber, Johansson, Leech, Conrad, and Finegan (1999: 361-364):

Verbs like *bring*, *buy*, *go*, *run*, *leave*, *move*, *work*, *take*, *give*, and *come* are *activity verbs*. The subject has the semantic role of an agent. They are used to refer to actions and events which hold the meaning of choice and can be used in both transitive and intransitive clauses. The following example is taken from the analyzed data of the second history lesson HIST2:

3. T: Camilla, move to your seat please.

[...]

S: I'm moving (HIST2: 1)

Activity verbs can also express volitional activities as well as non-volitional actions, events or static relations. (Biber, Johansson, Leech, Conrad & Finegan 1999: 361-362) The examples are taken from Biber and HIST2:

4. [...] i'd like to... make this list. (volitional) (HIST2: 2)
5. This will give the electron the chance to get over the lamp. (non-volitional) (1999:362)

Communication verbs belong to the category of activity verbs but they specify communication activities like speaking and writing. Communicational verbs are: *tell, say, shout, ask, write, suggest, state*, etc. Some examples are:

6. T: an illness....and Hippocrates said... (H2: 3)
7. T: and then he writes a prescription (H2: 3)

Verbs which are not used to talk about physical action and which are not necessarily volitional are *mental verbs*. These verbs give an account of any activities experienced by a person. The subject in mental verb clauses functions as a recipient. Not only do mental verbs have cognitive meaning but also emotional meanings which may refer to attitudes and desires, but also perception and receipt of communication. *Think* and *know* are words with cognitive meaning, *love, like, and want* express emotions, *see* and *taste* refer to perception and *read, hear, listen* to communication.

8. Doris: ahm...they could work, i think. (H1: 6)
9. T: but if they if they didn't ah didn't want to (H1: 6)

Mental words also are used to convey information about mental activities like *calculate, consider, learn, read, solve, and study* which are dynamic verbs.

10. T: [...] could Ennie just read what it says [...] (H1: 9)

However, not only dynamic verbs belong to the group of mental verbs but also verbs with stative meaning. Such verbs may describe cognitive states as *know, doubt, believe, understand* and *remember*.

11. T: [...] do you know the word league (H1:14)

And last but not least words like *love, want, suspect, fear, feel, enjoy, hate, and like*

refer to emotional states or attitudes.

12. S2: [...] in future the younger people would would [sic] care about it (H4: 9)

Causative verbs do not describe static states or activities but they refer to advancement or causation. A person or an inanimate item is the reason for change. Verbs which have such meaning are: *allow, cause, enable, help, let, permit, require* and *force*. Causative verbs are followed by either a nominalized direct object or complement clauses.

13. [...] and the world war influenced the life of thousands (H4: 10)

Another category of verbs are *occurrence verbs* referring to simple occurrences used to report events but only those which happen apart from volitional activities. The subject often is affected by the verb. *Become, change, develop, grow, increase, occur* and *happen* are examples for occurrence verbs.

14. S: then they developed a lot of... rockets (H4: 10)

Existence verbs do not only refer to states of existence but also to relationship, such verbs are not copular but they give information about a particular state of existence (*exist, live, stay*) or relationship (*include, involve, represent*).

15. [...] the underground factory is ah was not a area where people live at that time (M2: 5)

The last group of verbs discussed, in this thesis are aspectual verbs which refer to the progress of an event or activity. Generally this is done by verb phrases which are followed by a complement clause. Verbs such as *begin, continue, finish, keep, start* and *stop* are examples for expressing aspectuality.

16. T: finished Nadine let's start please (DAT 53, page 1)

This short introduction into a verb's various semantic meanings shows that single verbs can be categorized into a semantic field but also that it is not always

possible to account a verb to only one semantic field. Semantic meaning is closely linked to context. Only by examining the context it is possible to identify a verb's semantic meaning. The table below illustrates the various fields as well as their subcategories, also some examples of the semantic fields are added.

Illustration of Semantic Domains and their Categories

Domain	Subcategories	Examples
Activity Verb		bring, buy, run, leave...
Communication Verbs		tell, say, shout, talk...
Mental Verbs	Emotional meaning	love, hate, like...
	Cognitive meaning	think, know...
	Perception	see, taste, feel...
	Communication	read, hear, listen...
	Stative meaning	know, doubt, believe...
Causative Verbs		allow, cause, enable...
Occurrence Verbs		become, change...
Existence Verbs	State	exist, live, stay...
	Relationship	include, involve...
Aspectual Verbs		begin, finish, keep...

Table 1. Categorization of Semantic Domains

1.2.2 The Lexical Verb

The lexical verb expresses what is happening and therefore expresses „the principal action or event in a clause” (Aarts 2011, 67). The main verb can either be regular or irregular. While the regular main verb has four forms the base form, *-s form*, *-ing participle* and *-ed form*, the irregular main verb has less or more forms. (Greenbaum 1991: 48-49) The *-ed form* as Greenbaum calls it is the equivalent of Palmer’s term *-en form* for the past participle.

The forms of the regular main verb are:

Firstly, the base form which may also be called the first form of the verb in language teaching. This form of the verb is found in dictionary entries, for example *call* or *watch*. Secondly, the *-s form* of the verb would be *calls* or *watches*, the *-s* is added to the base form as an ending. The *-s form* is only used in present tense after the pronouns *he*, *she* or *it* and consequently after nouns and names which may be substituted by one of these pronouns. Thirdly, the *-ing participle* is a combination of

the base form and the suffix *-ing*, such as calling and watching. Fourthly, adding *-ed* to a base form changes the verb into a past tense verb or into an *-en participle*. (Greenbaum 1991: 48)

English Verb Forms

	base form	-s form	-ing participle	past form	-ed form
CALL	call	calls	calling	called	called
WATCH	watch	watches	watching	watched	watched

Table 2. English Verb Forms

In contrast to regular main verbs irregular main verbs have either more or fewer forms. While regular verbs have the *-ed form* for past and *-ed participle*, the irregular verbs have different forms. For example *hit* has only three forms: *hit*, *hits* and *hitting*, as its past tense and participle are also *hit*. However, *speak* has more forms than the regular verbs: *speak*, *speaks*, *spoke*, *spoken*, *speaking*. The most forms has the irregular verb *be*, it has eight forms. Its base form is *be*, but it has three present tense forms *am*, *is* *are*, and two past tense forms *was*, *were*, and *-ing participle being* and an *-ed participle been*. (Greenbaum 1991: 49) Next to the highly irregular verb *be*, seven major groups of irregular verbs exist. These are categorized by their principal parts, which are the base form, past form and *-ed* participle.

Regular and Irregular Verb Forms

	base form	-s form	-ing participle	past form	-ed participle
Regular Verbs	<i>call</i>	<i>calls</i>	<i>calling</i>	<i>called</i>	<i>called</i>
Irregular Verbs	<i>see</i>	<i>sees</i>	<i>seeing</i>	<i>saw</i>	<i>Seen</i>
	<i>go</i>	<i>goes</i>	<i>going</i>	<i>went</i>	<i>gone</i>
	<i>hit</i>	<i>hits</i>	<i>hitting</i>	<i>hit</i>	<i>hit</i>

Table 3. Regular and Irregular Verb Forms

There are three possibilities according to which irregular verbs can be classified. Firstly, the past and *-ed* participle are the same. Secondly, the base vowel does not change in past or *-ed* participle. Thirdly, past and *-ed* participle have an inflectional ending. (Greenbaum 1991: 76-77) This is illustrated in table 4 taken from Greenbaum (1991: 78):

Classification of Irregular Verbs

		Past = Participle	All vowels identical	Inflection
1.	burn, burnt, burnt	+	+	+
2.	saw, sawed, sawn	-	+/-	+
3.	keep, kept, kept	+	-	+
4.	speak, spoke, spoken	-	-	½
5.	cut, cut, cut	+	+	-
6.	feed, fed, fed	+	-	-
7.	drink, drank, drunk	-	-	-

Table 4. Classification of Verb Forms

The '+' shows that a characteristic implies and the '-' that it does not. '+/-' depict that one word does not have the feature. '1/2' is used for verbs which have an inflectional ending in the –ed participle but not in the past tense form. (Greenbaum 1991: 77)

1.2.3 Primary Verbs and their Functions

The English language distinguishes three primary verbs which may either function as main or auxiliary verbs. These primary verbs are *be*, *do* and *have*. The information below is based on Biber (1999: 428- 431) if not otherwise indicated.

Be is the most frequently used primary verb. It is used to connect the subject noun phrase with a subject predicative or adverbial.

17. T: [...] women and children are not citizens (H1: 1)

The auxiliary *be* is used to express progressive aspect and passive voice and as a component of the semi-modal verb *to be going to*.

Progressive aspect:

18. What is going on in the introduction phase (H1: 1)

Passive voice:

19. What are they called? (H1: 1)

The primary verb *have* can be used as a main verb or primary verb, which is

used to form grammatical functions as perfect and progressive aspect.

The transitive main verb *have* is the most frequently used lexical verb. One reason for its prevalent occurrence is that it has various different meanings and refers to numerous logical relations, such as *existentiality, family connections, food consumption, linking an inanimate subject to some abstract quality, linking a person to some abstract quality, marking causation and physical possession*. An example for linking a person to a quality is:

20. *T: yeah that's right do you have any idea (DAT 52, page 8)*

In addition to that the auxiliary *have* marks perfect aspect, which is very uncommon in speech especially in classroom discourse. Still in a marketing lesson some examples could be found, such as:

21. *These products have been launched recently (M1: 10)*

Finally, the most common primary verb *do* can function as a simple main verb, pro-verb, emphatic verb, and auxiliary verb in negations and questions.

As a main verb *do* does not necessarily express lexical content it mostly refers to activities which are relevant. In transitive constructions it refers to activities, but mostly is used for idiomatic expressions when it is combined with a noun phrase.

Activity meaning:

22. *Do me a favor.*

Idiomatic expression:

23. *I do the dishes.*

Do is also often used as a transitive in combination with pronouns like *it, this* and *that*. It functions then as a proverb and such constructions are called transitive pro-

verb forms. *Do* can even substitute a lexical verb in such a context. It is often used to avoid repetitions of a previous verb. Transitive pro-verb constructions are similarly used as an alternative for a series of actions and/or events, which is common in spoken language. In its intransitive function *do* is used instead of a verb or predicate. Often it is then considered to be an auxiliary, as it is often followed by *not*. If this is the case *doesn't* and *don't* replace negative predicates. Grammatically speaking it then is an auxiliary verb. A special case of the intransitive *do* is the question tag. In British English conversation it then functions as a main verb. However ellipsis – leaving out the question tag- are more common. (Biber, Johansson, Leech, Conrad & Finegan 1999: 428- 431)

Do, in such contexts, is called *dummy do* as it has no meaning and is used as a filler in negative and interrogative sentences as well as to express code and emphasis. The term used for this function of *do* is *do-support*. (Aarts 2011: 76) Greenbaum calls this *dummy operator do* (1991:21). The dummy *do* is used to express negations, inversions, code and emphasis. It is also used whenever grammatical rules ask for an auxiliary and if there is no other auxiliary verb given. (Palmer 1974: 21)

All in all the primary verbs' functions are twofold. They can either be main verbs, which can be marked for tense or auxiliary verbs which specify the main verb's meaning, and which can mark the main verb for aspect and voice. As they mark verb phrases for aspect and voice these verbs are the most frequently used verbs of the English language. Further information on how

1.2.4 The Modal Auxiliary Verb

Apart from the above discussed primary verbs which can function as auxiliary verbs there are verbs which only have the function of an auxiliary verb. Auxiliary verbs are used to add specific meaning to the lexical verb. For example they express intention. (Aarts 2011: 67) According to Greenbaum there are four types of auxiliaries namely, modal, perfect, progressive and passive auxiliaries. (1991:48) In fact only the first category belongs to the modal auxiliary verb class. The other three categories are primary verbs which function as auxiliaries. The focus is on modal auxiliaries, as these attach specific meaning to the main verb, while the other auxiliaries are used to

form grammatical functions. The last point made is that auxiliary verbs are marked by NICE properties. (Palmer 1974:14-15)

According to Greenbaum there are eleven auxiliaries with twenty-eight forms. The table below only illustrates finite and non-finite auxiliary verb forms:

List of all Auxiliary Verbs

	FINITE		NON-FINITE		
	PRESENT TENSE	PAST TENSE	INFINITIVE	-ing PARTICIPLE	-en PARTICIPLE
BE	am, is, are	was, were	be	being	been
HAVE	has, have	had	have	having	
DO	do, does,	did			
WILL	will	would			
SHALL	shall	should			
CAN	can	could			
MAY	may	might			
MUST	must				
OUGHT	ought				
DARE	dare				
NEED	need				

Table 5. Auxiliary Verb List

As already mentioned in 1.1 there are four types of auxiliary verbs, the last three (aux 2, aux 3 and aux 4) are primary verbs:

- modal auxiliary (aux 1), such as *can*, *may* and *will*,
- the perfect auxiliary *have* (aux 2),
- the progressive auxiliary *be* (aux 3), and
- the passive auxiliary *be* (aux 4). (Greenbaum 1991: 48)

According to Biber, Johansson, Leech, Conrad and Finegan the auxiliary specifies how “the action, state, or process denoted by the main verb is to be interpreted” (1999:99). Furthermore the first auxiliary is the most important one and functions as operator. (Biber, Johansson, Leech, Conrad & Finegan 1999:99-100) The term *operator* refers to the function of an auxiliary. They either function as negative contractions, inversions, emphatic polarities or post-operator ellipsis. (Collins 2000: 257)

The auxiliary verbs specify the following main verb, which means that they

mark the verb phrase for tense, aspect or voice. The modal auxiliary is always followed by the base form of the main verb, for example *can go, may sit, will change*. (Greenbaum 1991: 48, 55)

The modal auxiliaries can either be individually combined with a main verb or two auxiliaries can be combined to form grammatical functions. There are four possible verb phrases which are combinations of (aux 1) and (aux 3), (aux 2) and (aux4), (aux 2) and (aux 3) and finally (aux 1) and (aux 4). There may be gaps in the sequence but the order must not be changed. In the following diagram an example is represented as well as how the verb phrase may be labeled. (Greenbaum 1991: 55)

Modal Auxiliary Combinations

Combination	Example	Label
(aux 1) + (aux 3)	will be watching	modal progressive
(aux 2) + (aux 4)	has been watched	perfect passive
(aux 2) + (aux 3)	has been watching	perfect progressive
(aux 1) + (aux 4)	can be watched	modal passive

Table 6. Modal Auxiliary Combinations

Auxiliary verbs have various distributional characteristics in common. These are often called NICE properties. NICE is an acronym used for Negation, Inversion, Code and Emphasis. Auxiliary verbs can be followed by *not*, which means they form negative forms. This is very common except for the modal verb *may*. In contrast lexical verbs cannot be negated by *not* or the short form *n't* but need to use the dummy *do*. Similarly auxiliary verbs can invert with the subject while lexical verbs cannot but they also use the dummy *do* to do so. (Aarts 2011: 67-68) More precisely inversion means that the modal verb and the subject change place. This happens mostly in questions or in sentences which start with words like *seldom, hardly* or *never*. (Palmer 1974: 18) Code, or also called *avoidance of repetition* (Palmer 1974:19), is used to express „the property of auxiliaries that allows them to be 'stranded'"(Aarts 2011: 68). This means that the auxiliary can substitute verb phrases by ellipsis. For example a full verb is replaced by an auxiliary. This is similar to pronouns taking the place of nouns. Code is often used in questions and statements as in *Can you hand me that?- Yes, I can*. (Palmer 1974:20) Prosodic emphasis, or

emphatic affirmation as Palmer (1974:20) calls it, is used to emphasize certain utterances. This is done by highlighting the auxiliary as in the following examples found in Palmer (1974: 20):

24. *You must see him*
25. *I do not like it*
26. *You don't like it*

Prosodic emphasis also uses the dummy do to convey specific meaning. Not only have auxiliary verbs NICE properties but also primary verbs if they function as auxiliaries. (Aarts 2011: 68)

To put it into a nutshell there are four categories of auxiliary verbs. These and the so called dummy do support main verbs in specifying their meaning. While some auxiliary verbs can not be marked for aspect others are used only to mark aspect and voice. All auxiliaries however have in common that they share the following four properties: negation, inversion, code and emphasis. Last but not least verb phrases may include no more than two auxiliary verbs which have to follow a certain sequence.

1.3 Distinctions in the Verb Phrase

English verb phrases can be subdivided into finite and non-finite verb phrases. These have already been described above. For finite verb phrases there are six structural distinctions to be made: tense, aspect, voice, modality, negation and clause structure type. Tense only includes the differentiation between present and past tense. Aspect is closely connected to tense as present and past tense can have a progressive, perfect, or perfect progressive. Modality is not only restricted to deal with modal verbs but it also treats unmarked verbs. Voice shows if the agent of a sentence or clause is active or passive. All these domains can be positive or negative, interrogative or declarative. (Biber, Johansson, Conrad & Finegan 1999: 452) The following subchapters give a detailed account of the formation and function of tense, aspect and voice. In 1.3.1 the term tense will be introduced and the main tenses present, past and future tense will be discussed. 1.3.2 is devoted to aspect and 1.3.3 to voice.

1.3.1 Tense

Various linguists have engaged in discussions on the grammaticalization of tense, unsurprisingly there are numerous theories on how to define tense and on when the usage of a certain tense is appropriate and when it is not. This thesis concentrates on three main theorists and grammarians who advance a similar view. All approaches are centered around the pragmatic feature of deixis, although only Comrie (1985: 9) refers to it as such, other authors like Palmer (1974:39), Aarts (2011:261) and Biber (1999:454): call it the event's reference point to the present moment and go into more detail in giving information on rules and exceptions when to use or not use a certain tense. The following chapter provides an account of Comrie's, Aarts', Biber's and Greenbaum's findings. It also includes morphological, grammatical, syntactic, and lexical information on verb phrases.

According to F. R. Palmer (1974:43) tense has three functions

- to mark purely temporal relations of past and present time,
- for reported speech,
- to mark 'unreality', particularly in conditional clauses and wishes.

In this chapter only the first point is of interest, reported speech will be left out as the empirical research is exclusively on spoken language in CLIL classroom discourse and reported speech is not of relevance there. Also expressing 'unreality' is left out as, for example conditional structures are very uncommon in classroom discourse in which facts are stated.

Tense is dependent on expressions of locations in time, by verbs or adjuncts. This location of time is called reference point. The reference point for tenses is always the present moment. Tenses in general circle around this point, which means a situation occurs at precisely the present moment, before or after it. As a synonym for reference point Comrie uses the term deictic centre. The deictic centre and what happened before, after or at that point of time is a reference point to tense, according to which it can be categorized. Per definition a deictic system is any system in which units refer to one certain point. Tense may be called deictic as it refers to the present

moment which separates past from future (see diagram 1.2.). (Comrie 1985: 9-16)

Bernard Comrie defines the English tense as a „grammaticalized expression of location in time“ (1985: 9). This means that the verbs *go* and *went* express time. *Go* expresses present time and *went* past time. According to his definition the phrases *Sam goes* and *Sam went* can therefore be differentiated in terms of tense. On the basis of these phrases two conclusions can be drawn. Firstly, it is possible to say that a certain location in time is expressed, *goes* expresses present time and *went* past time. Secondly, that there are grammatical categories summarized under the term tense, this follows the fact that the verb carries the information of time. However, the phrases *Sam goes* and *Sam is going* can not be analyzed in this way as the difference is not one of time location but of aspect. (Comrie 1985: 9)

As Palmer said, tense marks temporal relations (1974:43). Normally situations are located in time by the form of the verb, and therefore by verbal inflection. (Aarts 2011: 243), (Greenbaum 1991: 49) In terms of morphological inflection the English tense system is very simple as only present and past tense exist. (Aarts 2011: 243), (Biber, Johansson, Conrad & Finegan 1999: 453) In English there is no future tense as such, as there is no future inflection or a grammatical form that can explicitly be called future tense, but there are other ways of talking about the future which will be discussed in the following chapters. (Aarts 2011: 243-244)

All in all it can be agreed that tense is a term which localizes a situation in time. If one considers tense to be strictly tied to morphology the future is not included, as verbs can not be marked for future time. The following chapters concentrate on absolute tense. Under this term those tenses which have the present moment as their deictic centre are summarized. (Comrie 1985: 36) It is also analyzed if there is a future tense or not.

1.3.1.1 Present Tense

The present tense is used to talk about events which happen around the reference point of the present moment. This chapter includes the grammatical formation of this tense, as well as its functions and special features.

There are only two tenses which carry time referential meaning, one of them is the present tense. This means that the verb is enough to function as the carrier of time and there is no need for an auxiliary to do that. The present tense verb is simply the base form of the verb. Only in the third person singular the present tense inflection is apparent on the lexical verb and on non-modal auxiliary verbs (primary verbs used as auxiliaries), for example *sees, feels, runs, has*. All the other present tense forms are in plain base form, except the verb *be*, which has three present tense forms: *am, is* and *are*. Psychological verbs like *think, believe* and *know* are often used when talking about cognitive states. Furthermore the present tense is often used in combination with so called 'performative verbs'.(Aarts 2011: 244-246)

Performative verbs, like *ap* *omise* are used by speakers who perform the action:

27. *I name this blog „An Editor's Blog“.* (Aarts 2011: 244-246)

The present tense is used for actions which are considered to be happening at the present moment and not before that or after that time. But the rules or theories on the usage of the present tense prove that a sentence in present tense does not necessarily have to be true only for the present moment. This might be due to the fact that it is highly unlikely that a certain situation is happening exactly at the moment of a speech act. All universal truths and habits, which are not only true for the present moment but which are generally considered to be true or habitual are grammaticalized in present tense sentences. (Comrie 1985: 36-37)

It has to be noted that the time of utterance and time of performance do not always coincide. Still there are some exceptions in which a situation or a performed utterance only consumes the present moment. Firstly, when the utterance describes that something will happen:

28. *I promise to give it back to you.*

In the above sample sentence a will-future construction is not used as the focus is not on when the thing referred to will be given back but on the promise itself. The act

of uttering the promise does not need more time than the present moment, equally as the act of naming the ship in the following example does not take longer than the present moment. (Comrie 1985: 37)

29. *I name this ship 'Prinz Eugen'.*

The English present tense is used to state that a certain situation is true for or at the time of speaking. (Aarts 2011: 244) The present tense may be said to have „two major meanings: to describe a state existing at the present time, and to describe present habitual behavior“ (Biber, Johansson, Leech, Conrad & Finegan 1999: 453).

30. *He wants a new pair of jeans.*

31. *She watches the ducks in the lake a lot.*

The first sample sentence expresses that a state exists at the present moment, the person referred to be in want of a pair of jeans now. The second, however, implies that the event is a habit. This is indicated by the adjunct *a lot*. A repeated situation, a habit, can be considered to be „making up a state“(Aarts 2011: 245).

The present tense can also state that an action is going on at the moment, for example *Here we go*. (Biber, Johansson, Leech, Conrad & Finegan 1999: 454) Furthermore it is used „to mean any period of time that includes the present moment“, and consequently includes 'all time' (Palmer 1965: 69). Examples for such a case may be scientific truths, like mathematical formulae, and laws of nature. For example:

32. *The Earth circles around the Sun.* (Aarts 2011: 245)

Comrie calls these universalities. (1985: 40) According to Palmer the present tense can refer to every period of time, either short or long and even eternal, but only if it „includes the present moment“. (1965: 69) Aarts also talks about 'instantaneous' or 'event' use of the present tense. There the tense is used to

describe an occurrence (event) that is simultaneous with the moment of speaking for example in spontaneous commentaries on [...] events (2011:245)

as in:

33. *Ronaldo gets the ball and shoots.*

Aarts's definition can be compared to Comrie who says that the present tense is used for situations and events which last no longer than the present moment. Ronaldo shooting the ball in the example above is such an event which happens momentarily at the moment of speaking (1985: 37).

The present tense is also used in sentences which do not refer to a certain time. This use is called 'timeless' present tense by Aarts (2011: 247). He claims that present tense is used to talk about „general descriptions, descriptions of journeys or museum exhibits in travel guides, instructions in manuals, stage directions in plays, [and] photo captions“(2011: 247). For example:

34. *And then you see this beautiful marble statue of a naked Celt.*
(Aarts 2011: 247)

In conclusion the present tense refers to the present moment. It is used for facts which are generally true, for habitual behavior, for references to emotions and psychological states as well as for performances. The main factor which has to be fulfilled is that situations or events are true for the present moment, and that the utterance is true at this time.

1.3.1.2 Past Tense

The past tense is the second grammatical construction in which the verb carries location in time. The reference point of this tense is in the present but one. Every event that happened before that is referred to by using the past. This chapter gives an account of grammatical constructions of the past tense which implies regular

and irregular verbs as well as modal verbs. Additionally, an explanation of its various functions will be given.

In the past tense the difference between regular and irregular verbs is of great importance as these verbs follow different rules of past tense formation. Regular verbs form their past tense form by adding the verb inflection *-ed* to the lexical word, for example *watch - watched*. Irregular words, however, form their past tense form quite differently mostly by a vowel change like in *drink – drank*. (Aarts 2011: 249)

It has already been stated that the past tense form of a verb is formed either with the *-ed* from of the verb or does follow other rules like a vowel shift. Modal auxiliary verbs however follow a different pattern. The progressive and passive *be* and the perfect *have* which do have a past tense form: *be- was/were, have- had*. *Must* does not have a morphological past tense form. If one wants to use the past tense of *must* one has to switch to the past tense form of *have*. All the remaining modal verbs have a morphological form of the past tense for example, *would, should, could, and might*. There are, however, a few exceptions which have to be taken into consideration. The semantic meaning of the modal verb *can* is twofold. (Aarts 2011:252)

35. *Can you hand me the plate please?*
36. *Could you hand me the plate please? (possibility)*
37. *Can /Are you able to drive your bike alone?*
38. *Were you able to drive your bike alone? (ability)*
39. *Were you allowed to drive your bike alone? (permission)*

Can is used to express 'ability', 'permission' or 'possibility'. The past tense form *could* is used to refer to 'possibility' but to express 'ability' mostly *to be able to* is used and to express 'permission' a form of *be allowed to* is used.

Also the semantic meaning of *shall* and its past tense form *should* differ While *shall* expresses 'intention', *should* expresses 'obligation'. (Aarts 2011: 252) This can be concluded from the following examples found in Aarts (2011: 252):

40. *We shall make up our mind when the IMF has reported. (intention)*
41. *We should make up our mind when the IMF has reported. (obligation)*

While the present tense is located at the present moment the past tense is located in time before that. The past tense can last for longer periods of time in the past but can also be used to talk about a certain situation which happened at a precise moment in the past. (Comrie 1985: 41)

Location of the Past Tense on a Timeline

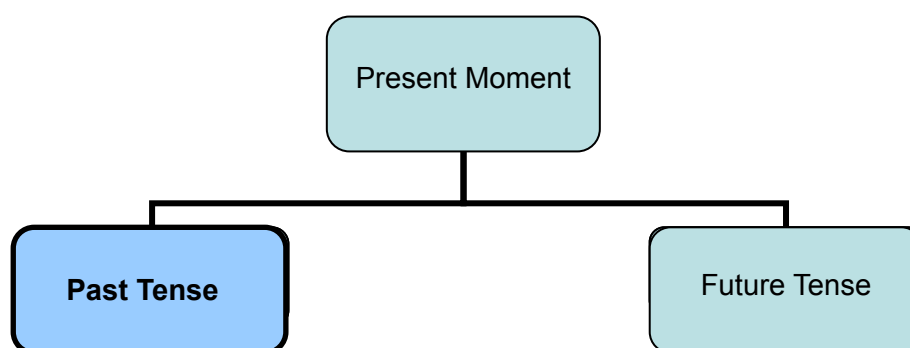


Table 7. Representation of the Past Tense on a Timeline

Comrie even claims that past tense can also be used for occurrences which last from the past until the present moment. However, it has to be argued that this might also be a matter of aspect and the present perfect tense should be used as it is used to talk about occurrences which started at a point in the past and last until the present moment or even longer and are relevant at the time of speaking. (Aarts 2011: 13, 71)
Comrie uses the example:

42. *up to this moment this disease was incurable (1985: 41).*

He claims,

[t]hat [the] use of the past tense only locates the situation in the past, without saying anything about whether that situation continues to the present or into the future (1985: 41)

According to Baas Aarts (2011: 249 -252) there are five typical uses for the past tense, obviously the past tense which refers to past situations, then the past tense used for reasons of politeness, modal past tense, past tense in reported speech and the past futurate.

Firstly, the past tense is used to talk about situations which happened in the past. Adjuncts like *last week*, *in (1983)* and *ago* may locate the situation at a certain time in the past. Secondly, the past tense may be used for polite requests as it “distances the speaker from their request, which makes it easier for the addressee to turn it down” (Aarts 2011: 250). Thirdly, the past tense suggests a distance between the speaker and the request. The past tense would be used if the speaker wants to give the addressee the chance to turn the request down more easily. Fourthly, the past tense can have modal implications and then is means to talk about situations which are non-factual or hypothetical. This is called *modal past tense* and is used to express *remote conditions* in clauses. Remote conditions are conditions which are unlikely to be fulfilled. They are introduced by *imagine if*, *if (only)* or *suppose*. *Remote conditions* can be transformed into open conditions by changing the past tense into the present tense. The modal past tense is also used after *it's time* or verbs like *wish*

All in all it may be said that the past tense is used to refer to past events which happened before the present moment. Although some modal verbs form different past tense forms they do not convey the same semantic meaning as present and past tense forms have differing connotations. The past tense of modal verbs is mostly used as a modal and not as a tensed lexical verb.

1.3.1.3 Future Tense

Even though the future tense does not have its own form Comrie counts the future tense to the absolute tenses (1985:43). The future tense is marked by either a modal or semi-modal verb, for example *will*, *shall* and *be going to*. While *will* and *shall* are not tensed *to be going to* can also be tensed. Using the semi-modal's past tense *was/were going to* is a means to refer to a future time but from a past view. The time to which such a sentence refers to may be before the present time but can also refer to an occasion that never took place. (Biber, Johansson, Leech, Conrad & Finegan

1999: 456)

Palmer (1974:36-38) does not include the future tense into his canon of tenses. He claims that only present tense and past tense are considered to be tenses. The future has no formal category and therefore does not belong to his primary paradigm, but to the modal paradigm. Reasons for that are that firstly, some forms like *I shall go* belong to the modal paradigm because of the modal verb *shall*. Secondly, some primary paradigms are used to refer to the future, for example the present tense progressive. However this has to be considered to be a matter of aspect. The present tense simple may refer to the future if a precise time, in the form of an adjunct, in the future is given. This can be illustrated as follows:

43. *The train leaves at 4 o'clock this afternoon.*

Palmer claims that the use of *will* or *shall* as future markers are not justified as there are other constructions which can be used to express the future, the already mentioned present tenses as well as forms of *to be going to* and *shall*. However, another reason for Palmer not to include the future tense in the primary paradigm is that the modal verb *will* does not always refer to a future incident but indicates probability or habitual behavior. He also claims that *will* often does not only refer to the future but also to willingness. (1974: 38) According to the evidence he has gathered modal verbs may refer to the future but „with additional reference to ability, probability, etc“(Palmer 1974:38 Nevertheless it has to be agreed that the future tense is not a member of the absolute tenses as verbs cannot carry a futurate time location.

Aarts seems to take an intermediate position as he mentions the different forms of future forms in his chapter on tense but he discusses them in more detail in his chapters on aspect and mood. He as well does not seem to count the future tenses to the absolute tenses but he in contrast to Palmer is of the opinion that there is some grammatical structure referring to the future which can be called tense nevertheless. As future time references he talks about *will*, *shall*, forms of *be going to* and the *progressive futurate* which corresponds to the present progressive use as a future form, which Palmer has already mentioned, but he also includes the past

progressive as Biber and his co-writers do. (2011: 254)

Bernard Comrie in contrast to Palmer, who is of the opinion that there is no such thing as a future tense and does not even consider its existence, is aware of the problematic nature of suggesting that the future tense belongs to the category of absolute tenses. He argues that although the modal verbs *will* and *shall* do express modality they also express a future notion. It is only natural that the future tense is something speculative as the future as such is something unsure and one can only make assumptions of what is going to happen. While the past can not be changed and is definite, the future is strongly connected to any kind of intervention and therefore not definite. From this point of view it can be understood that the future is a matter which should be treated in the grammatical field of mood and not tense. (1985: 43-44) However, Comrie goes on to explain what the future tense is and for what it is used, namely to

make a clear prediction about some future state of affairs, and is in this way clearly distinct from modal constructions that make reference to alternative worlds. (1985: 44)

In contrast to Palmer Comrie states that the usage of the modal *will* does not imply a reference to modality. He proves this on the following examples:

44. *It will rain tomorrow.*

45. *It may rain tomorrow.*

While the first sentences' truth can be tested by waiting to see if it rains tomorrow the second sentence is only a suggestion which cannot be tested. Still he agrees with Palmer in stating that *will* also functions as a modal verb expressing modality. Only by analyzing conditional subordinate clauses he comes to the conclusion that the English language in fact has a grammatical category of future tense. (Comrie 1985: 44-46) He found out that „future time reference uses of *will* are grammatically distinct from modal uses of *will* in subordinate clauses“ (Comrie 1985: 48).

All in all the future tense is not an absolute tense as verbs do not form future forms. However, because of didactic reasons and based on Comrie's research which shows

the existence of a future tense, it can be claimed to exist. The future tense refers to situations which will happen in the future and have their reference point there.

All theorists mentioned so far agree that *will*, *shall*, *to be going to* and present progressive express future time. In some circumstances even past progressive may be used. Yet the progressive aspect forms need an adjunct which locates the action in a certain time.

1.3.2 Aspect

Aspect and Tense are very closely connected. This is also made apparent by the structure of books on it. Aarts discusses both in one chapter while Palmer and Biber need two separate chapters and Comrie two books. In one point they all agree: every tense can have the notion of aspect, progressive and perfect respectively. While tense refers to a point in time, past, present or future aspect refers to the state of completion of an event which is described by a verb phrase. (Biber, Johansson, Leech, Conrad & Finegan 1999: 460) Aarts says that „aspectuality is a notion which concerns how a situation is perceived to be unfolding in time in the real world“ (2011:254). Aspect then is its grammatical coding. In the English language two kinds of aspect exist: progressive and perfect aspect. While the former is used to refer to actions which started in the past and are in progress, the latter refers to actions which lasted from some moment in the past to a certain place in time. The progressive aspect is formed with the auxiliary *be* and the *ing- participle*. The perfect aspect is marked by the auxiliary *have* and the *ed- participle*. Their difference in meaning is that the progressive aspect refers to an event or state which lasts longer or is still in progress or not yet finished at the time of the utterance. The perfect aspect, however, refers to events or states which started to take place earlier and last for a longer period up to a specific moment in time. (Biber, Johansson, Leech, Conrad & Finegan 1999: 460)

The main possible forms are according to Biber (1999:460):

- Perfect aspect present tense*
- Perfect aspect past tense*
- Progressive aspect present tense*
- Progressive aspect past tense*

But also combinations exist like the *perfect aspect progressive aspect present tense* and the *perfect aspect progressive aspect past tense* and even *perfect aspect progressive aspect future tense*. In the following chapters the progressive and perfect aspect will be discussed as well as combinations of those, but only with respect of the forms found in the analyzed data.

1.3.2.1 Progressive Aspect

The progressive aspect's grammatical construction is a combination of a tense marked form of *to be* and the *ing- participle*. It is used to talk about events or activities which are in progress, mainly with a time limit. One can distinguish three constructions, the present progressive, past progressive and non-finite progressive constructions. (Biber, Johansson, Leech, Conrad & Finegan 1999: 470), (Aarts 2011: 265). The present and past progressive are discussed in detail. Non-finite constructions are left out as they have no relevance for the data analysis.

Biber et al and Aarts do not share the same opinion on the use of the present progressive. Biber et al only state that it is used to talk about events which are happening at the moment of speech or which will happen in the near future. There mostly is an indication of what is going to be happening, which would be present progressive with future meaning. (1999:470) Examples would be:

- 46. *I am looking for a job at the moment. (is happening now)*
- 47. *What is Susan doing? (is about to happen)*

See the following timeline for illustration:

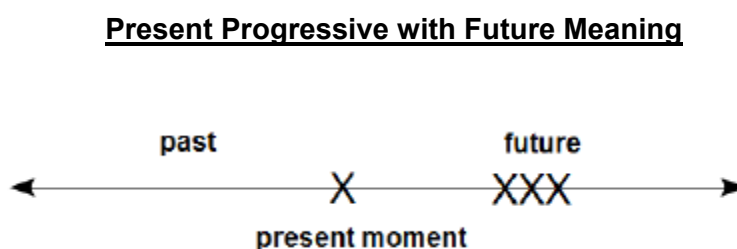


Table 8. Representation of the Present Progressive with Future Meaning

The single mark 'x' refers to the first sentence which is happening at the moment and the mark 'xxx' depicts the second sentence which is about to happen.

Aarts in contrast to Biber claims that the progressive aspect for present time has three different usages. The present progressive is used for situations that last from the past into the future, he agrees with Biber (1999: 470-471) in the fact that the duration is restricted. Examples which Aarts uses in his book on tense are:(2011:254)

- 48. *And he will now present her to those who are waiting inside the West Door of this great old church.*
- 49. *I'm rambling.*

Both examples are happening at the moment of speaking, and thus conform to Biber's claim that the present progressive is used for an event that is happening at the present moment. Additionally they do not only happen now but also start at a point in the past as the first of Aarts's examples does, the audience was waiting and is still at the present moment waiting. *I am rambling* is true for the present moment but also for the near future as *rambling* is a process which goes on. It can be concluded that the first usage given by Aarts goes hand in hand with Biber's usage. While Biber's definition is exclusive to the present moment and near future, Aarts claims that there is an implication of the event starting in the near past. Still Aarts' timeline has the present moment as reference point: (Aarts 2011: 265), (Biber, Johansson, Leech, Conrad & Finegan 1999: 460)

Present Progressive with Future Meaning and Past Tense Implication

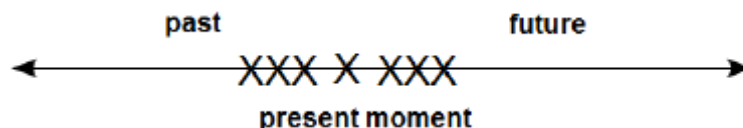


Table 9. Representation of Present Progressive with Future Meaning and Past Tense Implication

In the given examples the reference point is the present moment, however it has to be noted that the event does not have to take place at the moment of speaking but can take longer and simply has to happen in the broad sense of 'present time'. (Aarts 2011: 266) All in all „the progressive is used to denote temporary situations but „the

situation does not have to be transient“(Aarts 2011: 266). There is, however, a semantic conflict of meaning of progressive constructions, with its meaning that a situation is unfolding, and the meaning of the verb. This is the reason why a sentence as *I am believing in God* is impossible. Therefore the present progressive can not be used in sentences where the following verbs occur: *belong, contain, remain, believe, know, understand* and *want*. (Aarts 2011: 266)

While Aarts devotes a chapter to the past progressive Biber only discusses it marginally. While the present progressive presents an event which unfolds over a time in the present tense the past progressive does the same over a period in the past. Past progressive sentences may include an indication of when the event happened but some do not, however from the context such an indication may be implied. Only when an adjunct is given then a reference point can be specified. (2011: 266-267)

A sentence without a reference point would be illustrated like in the diagram below.

50. *John was playing the piano. (no reference point)*

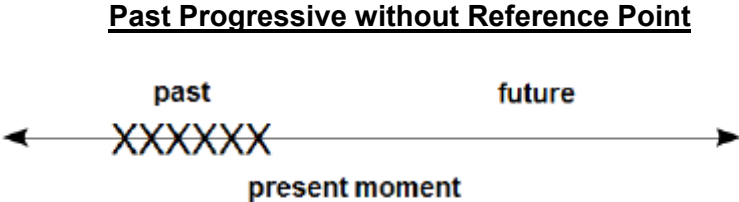


Table 10. Representation of Past Progressive with no Reference Point

A sentence with a given reference point however would be depicted like in the table, The symbol // is used for the reference point and the xxx show the duration of the action.

51. *John was playing the piano when the phone rang. (The underlined part of the sentence is the adjunct specifying the reference point.)*

Past Progressive with Reference Point



Table 11. Representation of the Past Tense including a Reference Point

Not only can a reference be given in a past progressive sentence but also a certain timeframe which is represented by the symbol II framing the beginning and the end of the time span. (Aarts 2011: 268) For example:

62. *My parents were on holiday in July.*

Past Progressive with a Timeframe

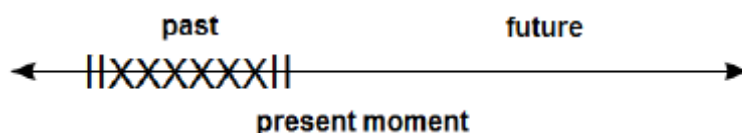


Table 12. Representation of the Past Progressive use in a Timeframe

Both progressive aspect constructions have in common that they express the duration of an event, which cannot be done by using tensed verbs without words expressing aspect.

1.3.2.2 Perfect Aspect

Further to past constructions which refer to past events there is also the present perfect construction based on the perfect aspect. The present perfect simple is used for events which happened in the past and have relevance for the present or the future moment. The present perfect construction is formed by the auxiliary *have* or *has*, for third person singular, and the *-ed participle*. Equally to the present perfect which has relevance in the present tense the past perfect construction has relevance for a moment in the past. As mentioned before the past tense is used when talking about events which happened in the past and are over having no direct relevance for the present moment. The past tense may refer to situations which have current relevance (1.2.1.2.), but they do not encode it, the present perfect however does.

This construction is formed by the past tense of the auxiliary *have*, therefore *had*, and the *-ed participle*. The focus of this chapter is the present perfect construction as past perfect constructions are very rare in conversation, especially in classroom discourse. (Aarts 2011: 255-256)

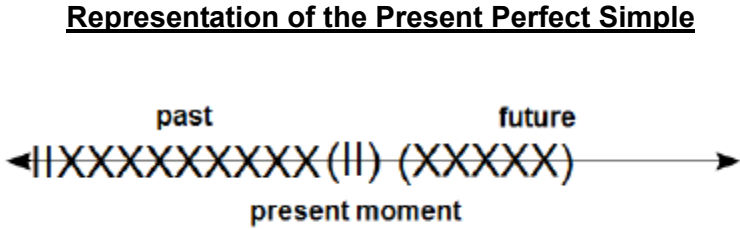


Table 13. Representation of the Present Perfect Simple

This graph shows that in sentences, like

63. *She has washed her car.*

and

64. *She has wasted her time for the last few weeks.*

the event started in the past 'II' and is relevant 'xxxx' at the present moment '(II)', which means that they have *current relevance*. But the graph also depicts that an event may have relevance for the near future which is represented by '(xxx)' on the timeline. (Aarts 2011: 256)

Present perfect constructions can be used to refer to several meanings. They are used to refer to the recent past, to results, and experiences, additionally the continuative present perfect refers to the duration of an event. All the examples are taken from Aarts (2011: 257-259) if not differently indicated. If present perfect constructions refer to the recent past adjuncts like *recently*, *just*, and *of late* are used. An example is given below:

65. *Of late, there has been research to indicate that procedural learning is retained in a great many contexts.*

When the present tense is used to indicate a result, this result has to have *current relevance*, as illustrated by the next example. The current relevance is that a lot of music has been downloaded and that this is relevant for the present moment.

66. *Both of you have recorded quite a good deal of music.*

The *experiential present perfect* expresses that a situation has already been mastered during a period of time which is not defined but which started in the past and lasted until the present moment. The following example refers to a person's previous experiences:

67. *Have you seen it before, Caroline?*

The last usage can be problematic for students as duration is mostly expressed by present perfect progressive constructions but may also be indicated by present perfect simple constructions. The so called *continuative present perfect* is used to talk about situations which are habitual and for situations which began in the past and lead to the present or future. (Aarts 2011:257-259)

68. *And on occasion I have lent handbooks and not got them back.*

In the next example lending handbooks to other people is a habit and the present perfect construction indicates that the person is still lending handbooks to other people.

1.3.2.3 Combination of Progressive and Perfect Aspect

The English language knows a large variety of constructions which combine the progressive and perfect aspect. Examples are the past perfect progressive, future perfect progressive and present perfect progressive. Some grammarians use the term modal perfect progressive tense to refer to such constructions. However, they are not tenses such as the present and past tense as they do not denote grammatical systems which locate situations within a timeframe. Therefore, as Aarts does, this phenomenon will be referred to as modal perfect progressive construction.

The focus for this thesis is the present perfect progressive construction as it is the only grammatical construction of interest for the following data analysis. (Aarts 2011: 270-271)

The distinction between the present perfect simple and present perfect progressive is not clear and has to be analyzed. As stated above the present perfect simple construction is used to refer to situations which have relevance for the present moment. This means that the result is of importance. Yet duration is also a criterion for the present perfect simple as the events last from a moment in the past until the present moment or even future. This is the problematic factor as duration is also of special interest for the present perfect progressive. The difference between the usage of these two grammatical constructions is that the present perfect progressive denotes that a situation is still in progress or ongoing. (Aarts 2011:272, Palmer 1985:68, Collins 2000: 74) Collins and Aarts agree that the function of the present perfect progressive construction is to convey the meaning that something is in progress. For example the sentence below expresses that relevance of meaning for the present, for which the perfect auxiliary *have* is combined with progressive element the *-ing participle*.

69. *Susan has been buying the tickets.*

Further to the intentional and 'ongoingness' expressing functions Palmer adds the factor of incompleteness. As in the following sample sentence:

70. *Someone has been using my pen.*

The difference between the continuative aspect of the present perfect and the present perfect progressive is that the former has an inception point, indicated by adjuncts, but the latter does not have such a reference point, adjuncts in such clauses only imply duration. All in all combined forms of progressive and perfect aspect are used to express progress, ongoingness, intention and incompleteness. In contrast to the simple perfective aspect not the result or relevance for the present or future moment are significant but the duration of the action or situation.

1.3.3 Voice

Grammatical voice is an often occurring syntactic feature, this however is only true for active voice and not passive voice. This feature distinguishes between active and passive voice. The terms 'active' and 'passive' stem from the differing semantic roles of the subject. The difference between active and passive clauses is described on the basis of the following sample sentences:

71. *Jenny opened the door. (active)*
72. *The door was opened by Jenny. (passive)*

In order to transfer active sentences into passive three changes had to be made. (a) The object of the active sentence was converted into the subject of the passive sentence. (b) The subject of the active sentence became the object of the passive sentence introduced by a *by-construction*. The subject of active clauses is called 'actor', and in passive clauses 'patient'. The noun phrase in the by-phrase is called 'agent'. (c) The verb form has also changed and marks voice. The passive form is constructed by a form of the auxiliary *be* combined with the *-ed participle*. Grammatical passive can be formed in all tenses but only with transitive¹ verbs. An alternative construction of passive voice is to use the verb *get* which functions as an auxiliary:

73. *I got chased by dogs.*

There are two forms of passive voice: long and short passive. Passive sentences with a by-phrase belong to *long passive constructions* and those which omit the by-phrase, as it is not obligatory, are *short passive constructions*. The short passive may also be called agentless passive. Aarts names a number of reasons why the agent can be left out. Either the agent is already known and can be deduced from the context or the agent is not known. Another reason is that the speaker does not want the agent to be of special interest. Leaving out the agent draws attention to the action. (Biber, Johansson, Leech, Conrad, & Finegan 1999:475, Collins 2000: 134-

¹Transitive clauses are clauses which have an object. Consequently, verbs in such clauses are called transitive verbs. (Collins:200:260)

146, Aarts 2011: 323-324) All in all active and passive clauses have the same propositional meaning, but may convey differing non-propositional meanings. The reason is the change of subject. (Aarts 2011: 323) This becomes obvious from the examples found in Aarts (2011:323):

74. *A man called Conte devised the process in 1795.*

75. *The process was devised in 1795 by a man called Conte.*

While the first (active) sentence is about 'a man called Conte', the second (passive sentence) is about 'the process'. (Aarts 2011: 324)

From this introduction of the passive voice it has become obvious which functions it has. Firstly, the agent can be omitted if irrelevant or unknown. Secondly, thematization, which means that transforming the active object into the active subject, has the purpose of making it more prominent. (Palmer 1985: 83-84)

2 Corpus Analysis

According to Leech (2001 ix) corpus analysis of spoken English as carried out this thesis, aims at giving an account of the frequency of certain verb usages based on Biber (1999). Before using a corpus, one has to consider what corpus linguistics is concerned with. McEnery and Wilson claim that it is “the study of language based on examples of real life language use” (1996: 1). In chapter 4 data from a real life situation is analyzed which consequently will be compared to findings of the corpus. The data analysis is language based as its aim is to find out what verb forms are most common in CLIL classroom discourse. It has to be mentioned that only a limited number of spoken English corpora exists. Although spoken communication is primary to written communication. However analyzing it and transcribing speech acts into texts can be read by a computer is very time consuming. The British National Corpus is one of the few which has a large enough number of items to be regarded as representative. (Leech 2001: 1) Leech (2011: ix) claims that

[f]or the teaching of languages, whether as mother tongue or as foreign or second language, information about the frequencies of words is important to vocabulary grading and selection

This however is not the point of the conducted research. The goal is to find out what verbs are commonly used and, therefore, have to be taught in second language acquisition. Embedding a chapter on corpus analysis of spoken English is necessary as the findings concerning the frequency of certain verbs as well as tense, aspect and voice are compared to the frequency in the analyzed data of CLIL classroom discourse transcripts. (Leech 2001: ix-x)

2.1 *What is a Corpus?*

The term corpus stems from the Latin term 'corpus' which means 'body'. A corpus, therefore, can be defined as a body of texts which functions as a base for the studies of empirical linguistics. But as such it has to be distinguished from other ways of examining texts such as semantics which aims at defining the meaning of expressions. In fact a corpus has four connotations. A corpus seeks to sample and represent data, it has a finite size. It is also important that its form has to be readable

for a machine and it has to be a standard reference. (McEnery 1996: 21) A corpus can be „any collection of more than one text“(McEnery 1996:21). The issue of sampling and representing data means that samples of a variety of languages are taken and analyzed. This is done as single utterances are easier analyzed and an analysis of all speech as such is impossible to undertake. A sample has to be representative which means that it should draw a picture as accurate as possible. To achieve this it is necessary to decide beforehand size of the sample is needed to be representative. Even though it is said that a corpus has a finite size this is not necessarily true. The John Sinclair's COBUILD team has launched a monitor corpus to which texts can be added at any time. The advantage of this kind of corpus is that it is not static and it has a larger scope. But since it is not sampled as rigorously (SYN) as a finite corpus it is not as reliable. When writing a finite corpus it has to be decided on a number of running words which shall be described in advance. One of the first corpora the Brown Corpus was based on 1,000,000 running words. The BNC is based on 100,000,000 running words which is one of the reasons why it is considered to be highly representative. In addition to its finite size a corpus has to be a machine-readable. Formerly a corpus was only considered to be a printed text, nowadays this has changed. The advantage of a machine-readable corpus is that one can navigate through it in a way which was impossible before. This means that accessing and finding information has become easier by this development. New texts or information can easily be added to such a corpus. A corpus is assumed to be of standard reference. This is, however, not an absolutely necessary part of a corpus. A standard reference refers to the language variety which is represented and presupposes (SYN) its availability for research. If a corpus is available for a variety of other researchers it is open for further research. Differing results can be compared without recomputerizing former researched data. (McEnery 1996:21-24) In general McEnery claims that a corpus nowadays is not a collection of texts but a

“finite-sized body of machine-readable text, sampled in order to be maximally representative of the language variety under consideration.”
(1996:24)

2.2 Corpus Linguistics

The research area of corpus linguistics is 'real life' language use and the study of 'real life' language. Corpus linguistics is a fairly modern term. This branch of linguistics was first introduced in the 1950s and was highly criticized in the 1960s and 1970s by generative grammar linguists since it was considered to be inadequate and unnecessary as it aims at descriptive adequacy and not explanatory adequacy, as generative grammarians do. (McEnery 1996:1-5) For example, corpus linguistics is not interested in the grammatical correctness of a sentence but in the number of times single words or grammatical forms occur. Aiming at descriptive adequacy was a major problem for generative grammarians since, as corpus linguists claim, complexity and variation are natural in language. Taking such a point was not accepted by generative grammarians and, therefore, criticism arose. While generative grammarians have a limited idea of language, corpus linguists are considered to be open-minded and open to change and variation in every day language use. Although other linguistic fields, such as semantics, can use corpus findings a major problem for many linguists is that corpora are rather performance influenced than competence influenced. However, performance is based on competence which makes this argument overstated. (Meyer 2002:1-2)

Corpus linguistics is rather a methodology which can be used in every other branch of linguistics. It is used to distinguish approaches and to define areas of linguistics. Today corpus linguistics has become an accepted kind of linguistics and the highly criticized Brown Corpus written by W. Nelson Francis and Henry Kučera have become famous and is considered to be a visionary work. (Meyer 2002:1-5) (McEnery & Wilson 1996: 1-17)

. According to Douglas Biber and Susan Conrad the importance of corpus linguistics for teaching English as a second language is twofold. They state in their TESOL article Quantitative Corpus- Based Research: Much More Than Bean Counting, that corpus linguistics is crucial because of „the centrality of register for studies of language use“ and „the unreliability of intuitions absolute use“. (Biber 2001:332) The former refers to the fact that patterns in different registers can either be strong or weak. In order to do an absolute analysis of all grammatical patterns, it

has to be borne in mind that patterns differ between registers. The latter claims that, for example, authors testing professionals and teachers often rely on intuition when it comes to the choice of the most important words and grammatical functions. However, corpus linguistics proves that intuition is very often inaccurate. (Biber 2011: 332)

2.3 *Relevant Corpus Uses*

A corpus is a source of empirical data used for a variety of different fields in applied linguistics. Empirical data is used to make objective statements about the language which is analyzed. For this thesis, however, only corpora in speech research, in lexical studies, in grammar and also in the teaching of languages are of interest. (McEney 1996: 87-88)

Corpora which aim at researching speech events provide a relevant sample size with many variables such as the speaker's sex, age and class across different activity types, like task-orientated speech, conversations and political speeches. This high number of data and the fact that spoken corpora include data from naturalistic speech allow making generalization about spoken language. Analyzing naturalistic data has the benefit of giving information about 'real' and every day life language. (McEney 1996: 88-89)

As this thesis is concerned with the analysis of different verb forms and seeks to get information about the distribution of lexical verb forms, corpus studies in lexicography is also an issue. The advantage for lexicographers is that computerized, machine-readable data can easily be found and listed. This thesis will compare corpus findings on the distribution of lexical verbs with CLIL classroom discourse data of history and marketing lessons. (McEney 1996:90) In addition to verb distribution also the distribution of lexical verb forms is of interest. According to McEney grammatical studies have used corpora for their research from a very early stage onwards. Due to quantitative research outcome it is representative for grammar and as its empirical data is useful for testing hypotheses' corpora are so popular in grammatical studies. (1996: 93)

Finally, the importance of corpora in teaching languages has already been mentioned. As many teachers but also authors rely on their intuition when it comes to the importance of verb forms to be learned and grammatical forms, corpus linguistics is used to give empirical and reliable information which should help to design course books and lesson plans. (McEnery 1996: 103) At the end of the data analysis it may be deduced that the findings on grammatical form distribution and verb distribution correlate with the findings in the data. A positive outcome would be a hint that teachers have achieved to teach the students the most important grammatical forms and verbs to communicate successfully.

2.4 British National Corpus

The British National Corpus, short BNC, incorporates information of over 100 million words of spoken and written British English. It is a finite, balanced, sample corpus. The data is taken from no earlier than the 1960s. Leech uses the BNC as basis for his book *Word frequencies in written and spoken English*. For this thesis only spoken English is of importance. (Leech 2001: 1-2) 90% of the corpus consists of written British English data and 10% of various types of spoken British English. Although 10% does not sound much, the BNC is the largest collection of spoken data available. (Meyer 2001: 30) It is demographically sampled. (Meyer 2001: 34) The BNC uses spoken data from 1991 onwards, which is 10 percent of the data. This data is divided into two parts. The first part of data stems from conversations and the second part from task-oriented speech. This corpus can be said to be representative as its participants are from different age groups as well as social groups. Additionally, the analyzed speech acts consist of different types ranging from educational and informative to business and public and leisure speech types. (Leech 2001: 1-3)

When talking about words in a corpus, it is important to be aware of the difference of types and tokens. Type is the term used for verb forms as they occur in dictionaries. If verbs are represented as such in the corpus list they are listed by lemma. Tokens are all forms of verbs in a text for example past tense and participles. The BNC consists of three different kinds of list formats. Words are either listed alphabetically, descending according to their frequency or distinctiveness. Speech is far more difficult to analyze than written language, especially as it is so time

consuming and as there are a lot of differences in spelling. In addition to that English words often have numerous semantic meanings and one form of a verb may refer to different grammatical constructions, which is a problem not only of analyzing spoken but also written English. Therefore some distinctions have to be made. The BNC, for instance, does not state differences in the meaning of abbreviations and acronyms. Also words which are spelled in the same way, such as *like* which can either be a verb, a preposition or an adverb, are only differentiated on the level of major word patterns. Moreover grammatical functions are not distinguishable like past forms and the past participle, as in *visited*. If one verb can either be the base form of one verb (*found*) and the past tense (*found*) form of another verb the distinction is made in the corpus by lemmatization. (Leech 2001: 3-6)

The BNC only lists words which occur more than 10 times per one million words. English language speakers use an enormous variety of words and phrases most of them appear very rarely. For example from 757,087 words 397,041 occur only once which is 52 per cent of all words. Furthermore 570,344 words appear less than 5 times and only 186,043 words appear more often than 5 times. From these 186,043 words only 124,002 are used more than 10 times per one million words, which is a relatively small number. (Leech 2001:8-9)

In order to analyze spoken data it is necessary to transcribe spoken material into written data. The BNC has done that in accordance with standard orthographic practices. Word fragments, which are fragments of words, are not analyzed. Word fragmentation is a phenomenon which only occurs in spoken language when speakers start uttering a word and stop in midsentence. Words as *gonna* and *dunno* are transcribed as *going to* and *do not know*. If certain tokens or fragments are left out, the question of accuracy arises. As linguistic programs used for computer analysis are not able to distinguish between grammatical labels, errors can occur. The BNC does not attempt to illustrate the frequency of grammatical forms. All in all the accuracy of finding major parts of speech is slightly more than 98 per cent. (Leech 2001: 10-15)

The following chapter will illustrate the frequency of verbs and which grammatical forms are most commonly used in the BNC and the Longman Spoken

and Written English Corpus. The major contrastive features will be described as well and in chapter 4 the outcome will be compared to findings of the LSWE and BNC.

2.4.1 Frequency of verbs in general

The table below contains a list of the 20 most frequently used words in spoken English. They have been structured by lemma.

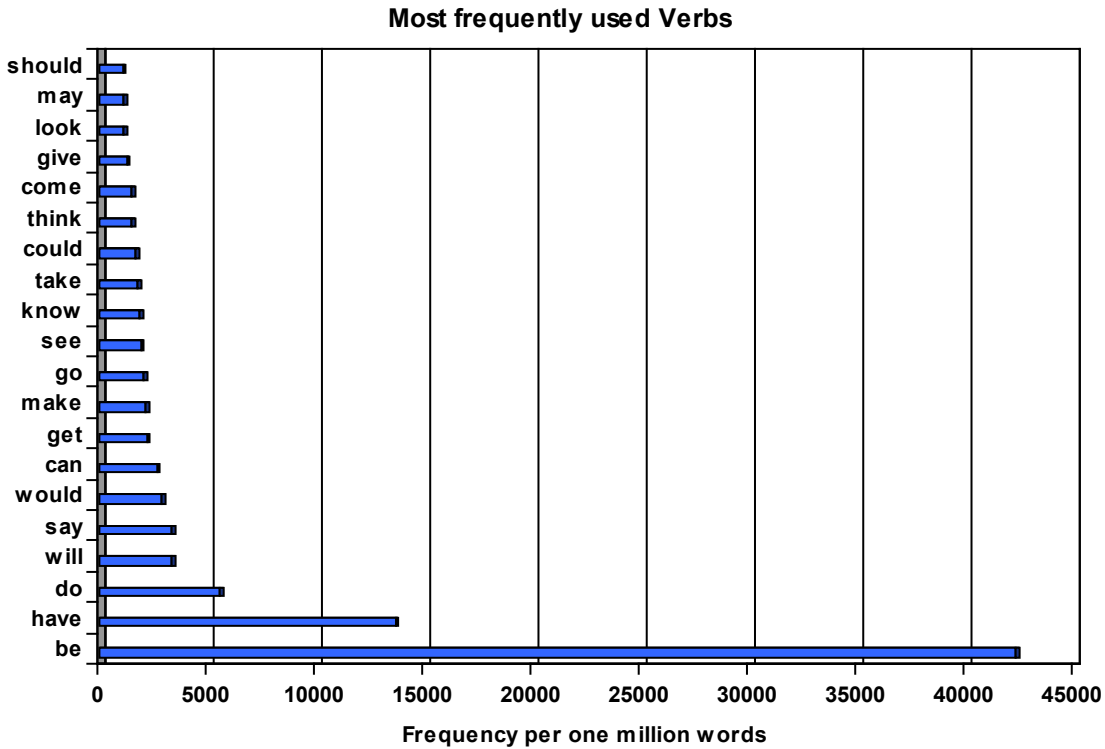


Diagram 1. List of most frequently used Verbs (BNC)

The three most frequently used verbs are *be*, *have* and *do*. All of them are primary verbs and function as primary auxiliary verbs. They are used to form questions and negations, *have* is furthermore used to express perfect or progressive aspect. The verb *be* is by far the most commonly used verb. This number includes all forms of *to be* like *is*, *am*, *are*, *was*, *were*, *been* and *being*. Reasons for this outcome, which had to be expected, are that *be* is not only a lexical but also a primary verb which has numerous uses, for example it is necessary to form questions and negations in present tense. All in all 9 of these 20 verbs are auxiliary verbs and the rest lexical verbs with differing semantic meanings, like mental (*know*, *think*) or activity verbs (*go*, *come*). (Leech 2001: 282)

2.4.2 Frequency of grammatical classes

The above table has shown that *be*, *have* and *do* are the most frequently used verbs in spoken English. The following tables illustrate the frequency of their verb forms per one million verbs. Also the distribution of the grammatical forms of a lexical verb are given, which is illustrated in the last table.

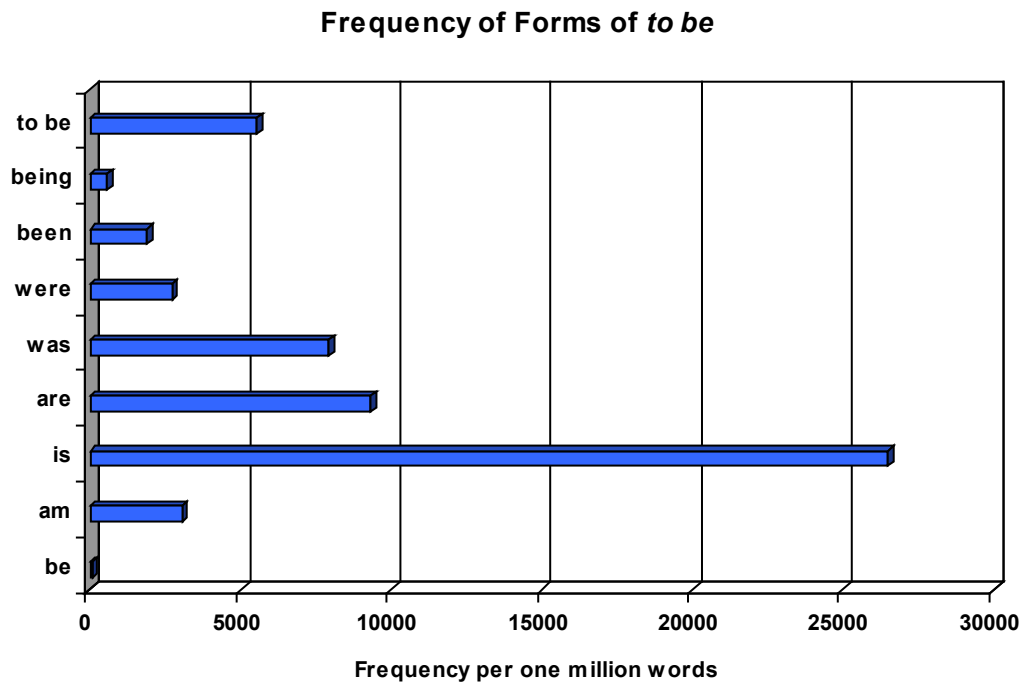


Diagram 2. Frequency of the Forms of BE

From this data which was taken from the BNC it becomes obvious that the present tense is the predominant tense used in spoken English. In total over 30.000 words per million are in present tense in contrast only 15.000 forms of *to be* are past tense. The considerably high number of *is* usages suggests that conversations are used to talk about single persons, events or situations as those are referred to by *is*. The past tense is naturally used to talk about past events, which is also very common in speech acts. An interesting factor is that the infinitive *to be* is used over 5000 times. This is due to the fact that grammatical constructions, like modal phrase constructions, are formed by a modal verb and a base form. (Leech 2001: 297)

Frequency of forms of do and have in comparison

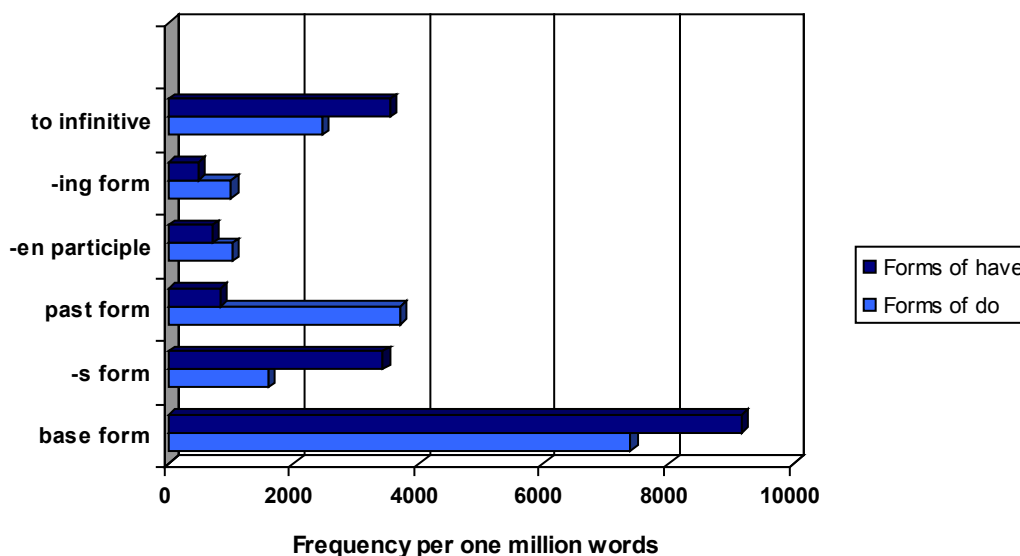


Diagram 3. Frequency of the Forms of DO and HAVE

The number of uses of the form of *do* and *have* are very similar, as can be seen in the graph below. Here, also the use of the present tense is predominant to other tenses and also the base form without the inflectional *-s* is used far more frequently than all the other verb forms. While *have* appears over 9000 times do only appears over 7000 times per one million words. Further similarities are that the to infinitive is quite often used; however, the *-ing* participle as well as the *-en* participle are the least used forms. All this is also true for forms of *to be* discussed above. One major difference is that while the past tense forms of *be* and *do* are used with the second highest frequency, the past tense of *have* in comparison is very seldom used. This might occur due to the fact that people rarely speak about what they have or had experienced and rather discuss differences in life or what they are or were able to do. However what they have and had done or experienced is another issue and the rare usage may be due to psychological reasons. (Leech 2001: 297-298)

When taking a look at the frequency of lexical verb forms in contrast to the three primary verbs, which as already said have different usages, the most obvious observation is that the infinitive is the most used form of lexical verbs followed by the base form. A reason for this is that lexical verbs very often occur with modal auxiliary verbs. The lexical verb often only needs to convey semantic meaning and not tense,

aspect, voice or mood. Therefore the infinitive is a commonly used form. As the present tense is the mostly used tense in conversations the base form is the second in the lead followed by modal auxiliaries and the past participle. The -en or -ed form is very frequently used with lexical verbs, but not with primary verbs, this due to the fact that perfective aspect and voice are very rarely occurring grammatical functions.

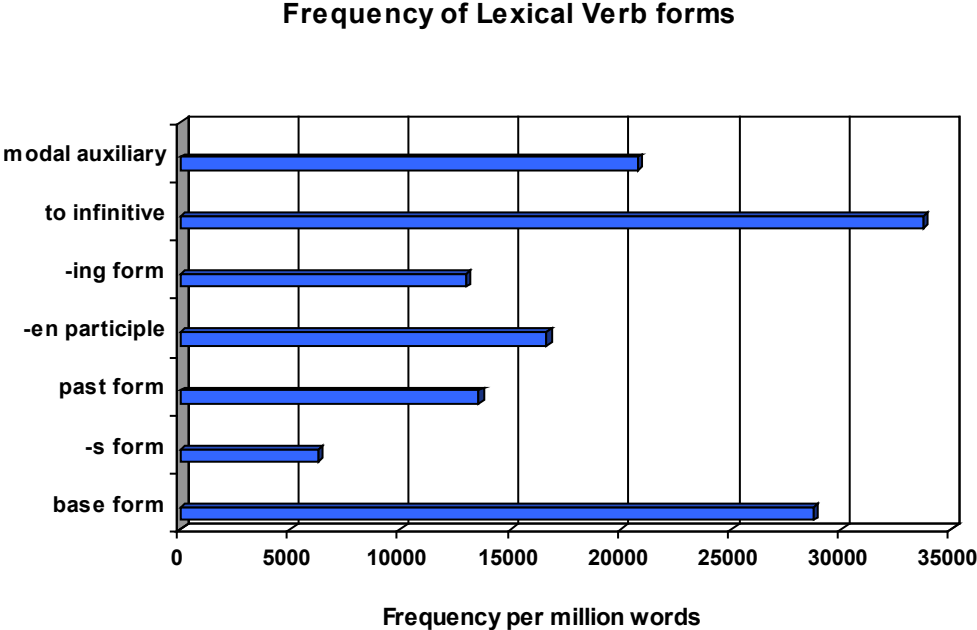


Diagram 4. Frequency of Lexical Verb Forms (BNC)

The most striking difference between the occurrence of certain verb forms of lexical verb forms in contrast to the primaries *have* and *do* is that the to infinitive is the most often used verb form in lexical verbs with nearly 34.000 occurrences per million words. While the base form is the mostly used verb form of *have* and *do*, the base form is only the second most used verb form in lexical words. It is also interesting that the –s form is not so common in lexical verbs. This is due to the fact that *have* and *do* are very often used when talking about people and their actions. Additionally *has* and *does* are required when negating sentences or posing questions. Another important difference is the fairly often used past participle, over 16.600 occurrences, in lexical verbs as they are necessary in perfective aspect constructions. Neither the past participle *done* nor *had* are very common, they are used less than 1.000 times per million words. (Leech 2001: 296-297)

2.4.3 Differences between conversation and task-oriented speech

The British National Corpus does not only include numbers for verb occurrences in conversations as the LSWE does, but it contrasts conversational to task-oriented speech. This is interesting as the conducted research does not analyze every day conversations, but task-oriented speech. Speech being task-oriented means that the use of the English language has an educational aim, therefore, it can be said to occur in CLIL classroom discourse. The following diagrams illustrate the difference in conversations and task-oriented speech in the participant's usage of *be*, *do*, *have* and lexical verbs. The dark blue columns show the numbers for task-oriented speech and the light blue columns bear evidence for the conversational numbers. (Leech 2001: 301-302)

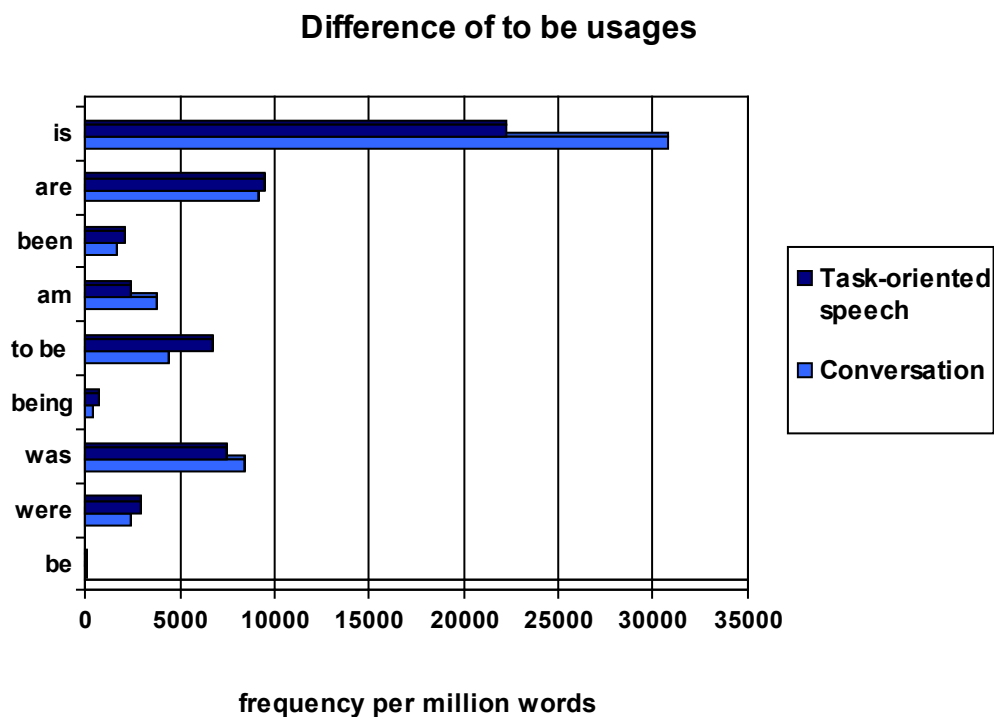


Diagram 5. Difference of the Uses of BE in Task-oriented speech and conversations (BNC)

Both in task-oriented speech and in conversations *is* is the most frequently used form of *to be*; however, it is used 1.000 times more often in conversations, which might be caused by the fact that in conversations one normally talks more about people and things than one does in task-oriented speech which circles around

abstract topics and requires a more eloquent language. Also in task-oriented speech people are not so often addressed and the employed language is thus impersonalized. This is shown by the fact that the infinitive is a few hundred times more often used in task-oriented speech than in conversational speech. The other verb forms do not show such striking differences. (Leech 2001: 301)

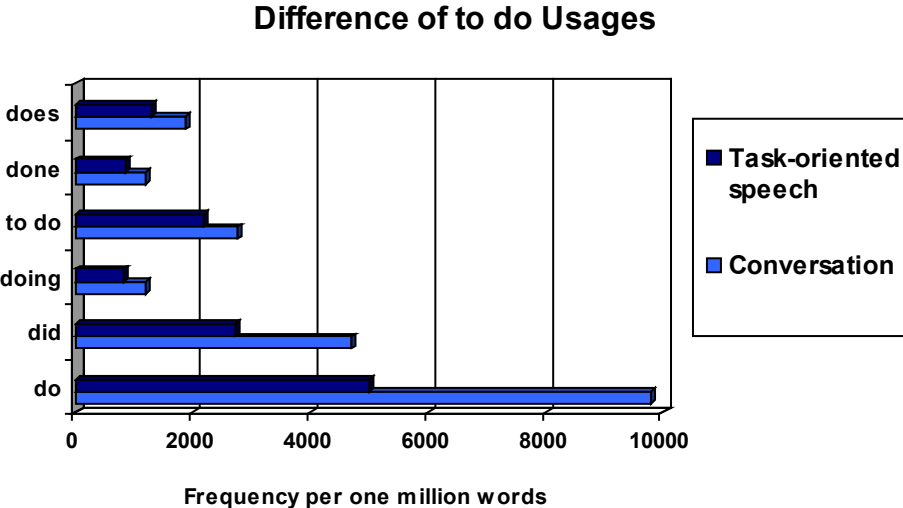


Diagram 6. Difference of the Uses of DO in Task-oriented speech and conversations (BNC)

As already mentioned, the forms of *to do* are 4 times more common than forms of *to be*. Furthermore the base form *do* is the most frequently used verb form in contrast to the 3rd person s- form *is*. From the illustration above it becomes obvious that *do* is approximately half as common in task-oriented speech. Task-oriented speech is not as concerned with what people are doing therefore its occurrence is limited. Still expressing activities is a bigger issue than referring to states of being in both task-oriented speeches and in conversations. (Leech 2001: 301)

Difference of to have usages

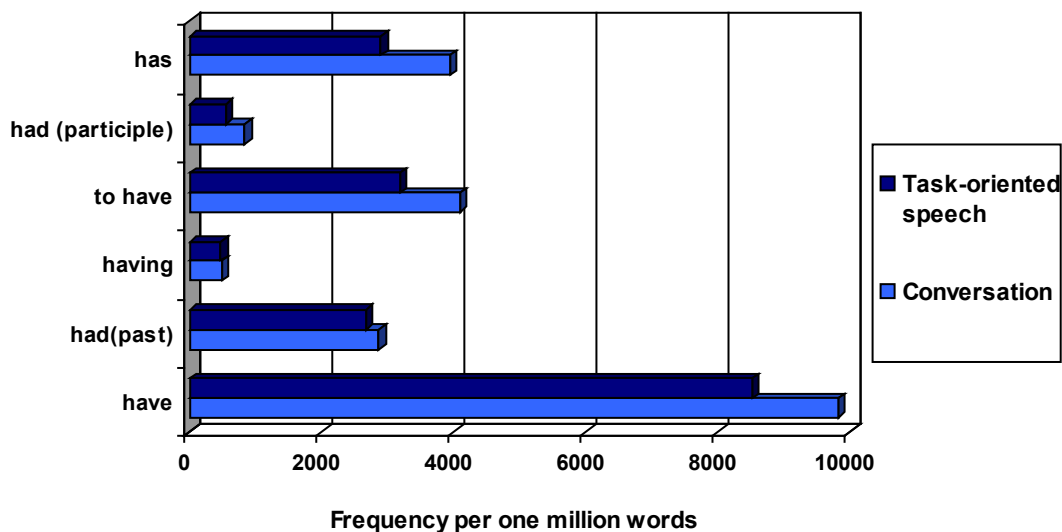


Diagram 7. Difference in the Use of HAVE forms in task-oriented speech and conversations (BNC)

Forms of *have* are as frequent as forms of *to do*. Also here *have* is less used in task-oriented speech, however the difference is not as prominent. The present tense, as in the usage of *do* and *be* is the tense mostly used. All in all the primary verbs are very common as they do not only have primary modal, meaning but they are also used as modal verbs and main verbs. (Leech 2001: 301)

Difference in the Use of Lexical Forms

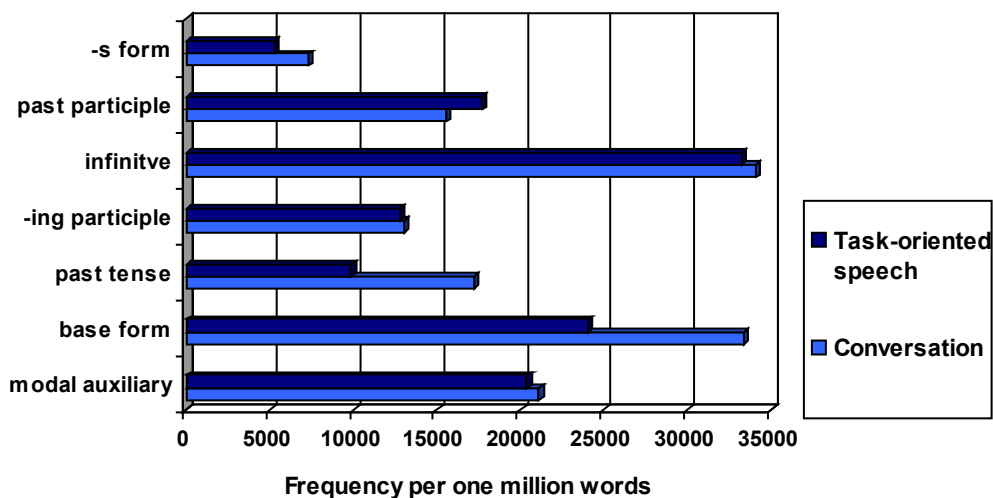


Diagram 8. Difference in the Use of Lexical Verb forms in Task-oriented speech and conversations (BNC)

Lexical verbs are very common in both conversations and task-oriented speech. Except for the –en form all lexical verbs are less frequent in task-oriented speech. It, however, more often used in task-oriented speech. The difference between the occurrences is most dominant in the appearance of the base and past tense form. The base form is used 1.000 times less in task-oriented speech and over 500 times less in the past tense. Overall lexical verbs are 3 times more common than the individual primary verbs as they express action, behavior and feelings. All the primary verbs put together are slightly less used than all lexical verbs, due to the fact that they are only used when expressing modality, tense or aspect. However, they are not used for present simple statements, which is the reason why lexical verbs are more frequently used. (Leech 2001: 302)

2.5 Longman Spoken and Written English Corpus

Another corpus on spoken English worth mentioning is the *Longman Spoken and Written English Corpus*, short LSWE corpus, by Biber, Johansson, Leech, Conrad and Finegan. The LSWE is based on 37,244 texts and 40,026,000 words. It provides a solid basis for the analysis of grammatical patterns. Furthermore it systematically represents four major registers, which are conversation, academic prose, fiction and news. For this thesis only the register of conversation is of interest as well as the additional register of non-conversation speech which happens in lectures or public meetings. The register of importance can be subdivided into American English, which is based on 329 texts and 2,480,800 words, and British English with 3,436 texts and 3,929,500 words, as well as non-conversational speech also in British English with 751 texts and 5,719,500 words. On average 7,000 words can be registered in an hour which is an approximate number of 120 words per minute. Considering the speech rate one-million words equal 140- 150 hours of spoken data. The distinction between American and British English is only of importance if an item is differently used in either one dialect. Just as the BNC the LSWE also does not give an account of all variations in register. It is stated that presenting readers with too much information on every variation in register is overwhelming and does not provide an overall picture. While the BNC's data dates back to the 1960's the LSWE is more modern using data produced after 1980. (Biber, Johansson, Leech, Conrad & Finegan 1999: 24-27)

The data of the LSWE was collected by choosing a range of English speakers, with differing age, sex, and social groups. Those people had to record their conversations on a tape recorder, just as the informants of the BNC had to. Then the data was transcribed in order to make an orthographical, lexicographic and grammatical research possible. The non-conversational speech acts, gathered by Longman, consist mostly of monologues and speeches. Only about one third of this data comprises answer and question interactions. (Biber, Johansson, Leech, Conrad & Finegan 1999:29-33)

2.5.1 Distribution of lexical, modal and primary auxiliary verbs

The LSWE confirms that lexical verbs are the most common verb form which occurs more than 100,000 times per one million words. This means that more than one out of ten words is a lexical verb. Lexical verbs are followed by primary verbs in terms of the hierarchical order and modal verbs are the least commonly used verb forms. It is worth noting that the primary verb *be* is the most common verb more or less as common as all modal auxiliary verbs together. Primary verbs in conversations often denote a progressive form and are, therefore, very often used when talking about events which are happening at the moment of speech. From the table below it can be concluded that as lexical verbs are the most commonly used verb they do not always occur together with an auxiliary verb. The reason why primary verbs are more often used than modal verbs is that they express aspect and modality and, therefore, have an important position in conversation. Moreover the single auxiliary verb *do* is often used as it is necessary to form negations and questions. The graph furthermore shows that the appearance of lexical words in conversations is extremely common in conversations. A reason for this high frequency is that in conversations people talk about actions and events which are formulated in short clauses. Short clauses mainly express one idea and are governed by one lexical verb. Also talking about personal information is common, which are built around lexical words. (Biber, Johansson, Leech, Conrad & Finegan 1999: 359-360)

Frequency of Verb Classes according to the LSWE Corpus

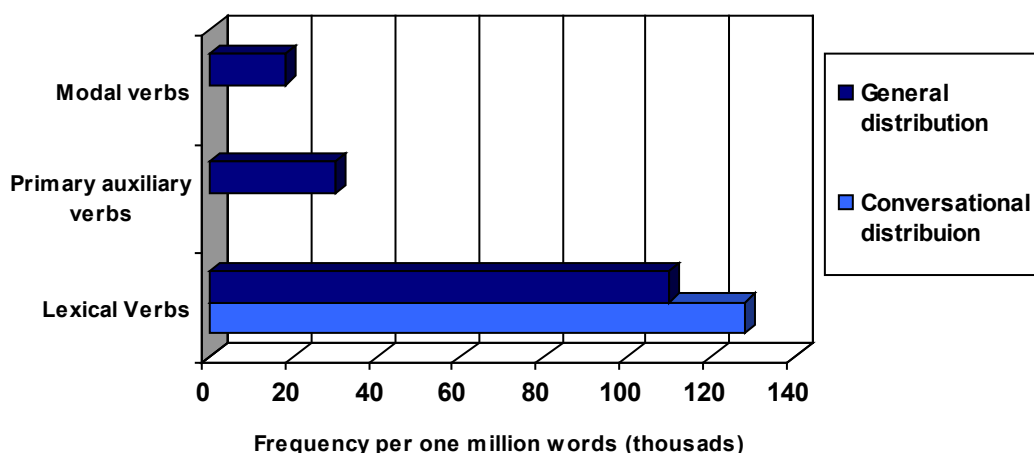


Diagram 9. Frequency of Verb Classes in general and in conversations (LSWE)

2.5.2 LSWE Findings on Semantic Domain Distribution

In the previous chapter different semantic domains have been discussed. According to the semantic domain categorization verbs can be sorted into activity, mental, communication, existence, occurrence, causative and aspectual verbs. Now their distribution and frequency patterns will be illustrated. The LSWE corpus found out that approximately 50 per cent of all verbs are activity verbs. While mental and communication verbs are also often found the other domains are not so frequent. The distribution of verbs which hold a special semantic meaning gives evidence of the topics which are usually discussed. As already mentioned most common are activity verbs, such as *bring, buy, get, go, give*, closely followed by mental verbs, as *believe, feel, find, know, and like*. Existence and communication verbs are also common, however, aspectual, occurrence and causative verbs are very rare. (Biber, Johansson, Leech, Conrad & Finegan 1999: 365-366)

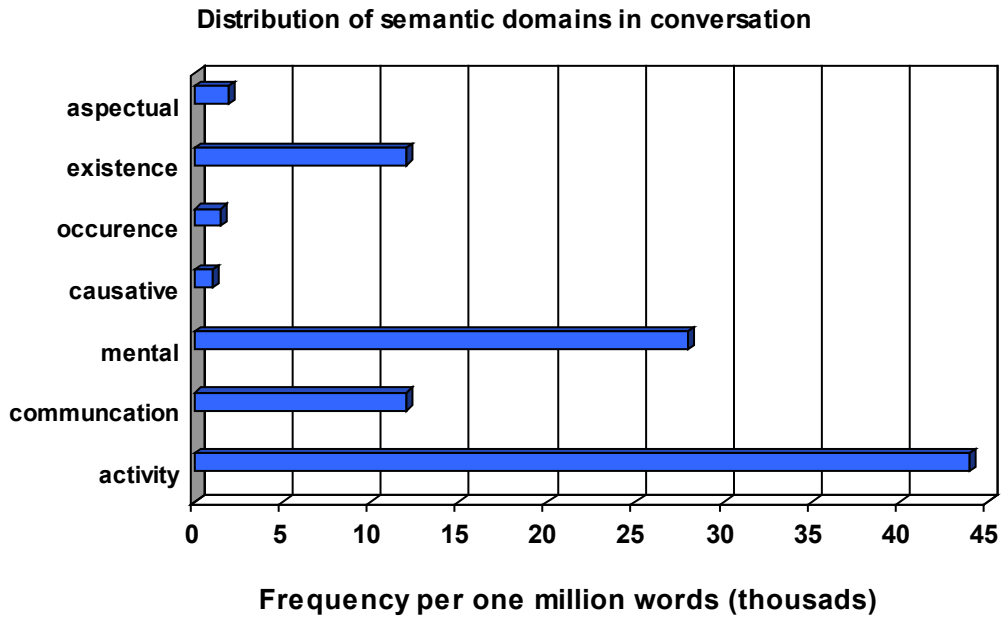


Diagram 10. Semantic Domain Distribution (LSWE)

2.5.3 Most Common Lexical Verbs

A list of the most common lexical words in conversation is provided in order to compare it a list of the most used lexical verbs in classroom discourse.

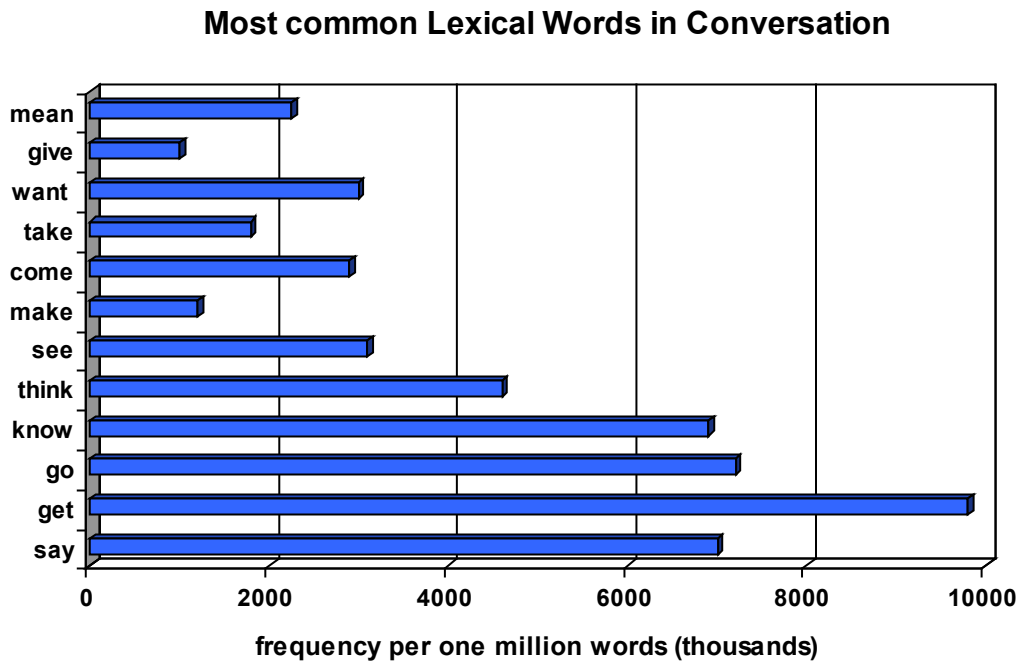


Diagram 11. List of the most Common Lexical Verbs (LSWE)

The verb *get* is the most common word in conversations, due to the fact that it can be used in numerous grammatical patterns. It also has various semantic meanings. It is an activity verb used to express that something is obtained or it refers to something being moved away from something else. It also is causative as it is used to express the cause of an event. When it refers to changing states, it belongs to the semantic field of occurrence and it is also a mental word to confirm that something has been understood. Additionally it is used to substitute *have* with static meaning in the form of *have got*. Furthermore *get* is used as an auxiliary verb in passive voice, in multi-word phrases and elisions. Also the verb *say* is very frequently used. Mostly the past tense form of this verb is used. The present and past forms are especially common in reported speech constructions. It is also used to express habitual behavior. All the other words are only interesting because of the distribution in other registers, which is not of relevance for this thesis. (Biber, Johansson, Leech, Conrad & Finegan 1999: 373-377)

2.5.4 Frequency of Tense and Modality

Verbs can either be marked for tense or modality, but never both. The LSWE states, that in both written and spoken language, tensed verb phrases are more common than modal verb phrases. In conversations the present tense is most frequently used, three times more than the past tense and modal verbs. This is illustrated by the table below.

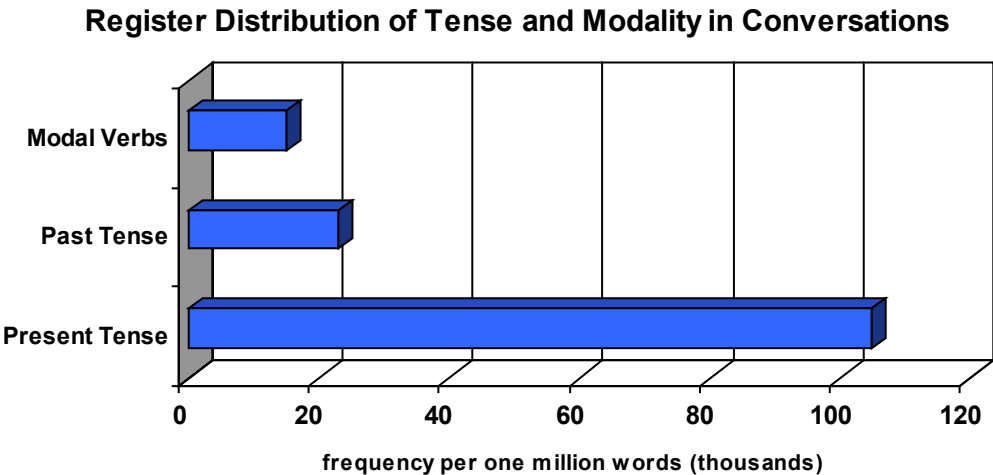


Diagram 12. Distribution of Tense and Mood (LSWE)

The present tense, as said in the previous chapter, is used to refer to some past events, habitual behavior, states true for the presence and also for future events. This is however not the reason why it is so common in conversations. The present tense is used to mark the speaker's focus on the context. In conversation it mostly has the function of expressing habitual behavior and current states. The progressive aspect in the present tense is common to display intention and the perfect aspect refers to events which happened in the past and its result is of interest in the present. (Biber, Johansson, Leech, Conrad & Finegan 1999: 456-458)

2.5.5 Corpus Findings on Aspect

The English aspect is a very rare grammatical form in the English language as it refers to states of ongoingness. The number of verbs unmarked for aspect is more than twice as high as the number for aspectual verbs. The progressive aspect is slightly more often used than perfect aspect. Aspect mostly occurs in the present tense. A combination of perfect and progressive is very rare and is not illustrated in the table as it comprises only 0.5 per cent.

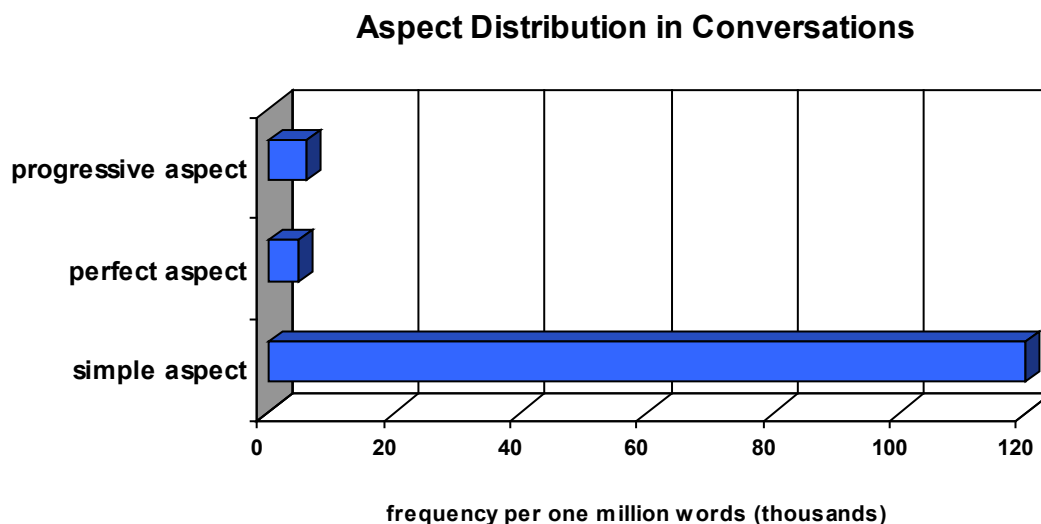


Diagram 13. Distribution of Aspect (LSWE)

In conversations the following verbs are most frequently used with the present perfect aspect: *been, had, got, gone, and done*. Also quite often *made, seen, come, said, taken, though, called and put* occur. The present perfect aspect is in general used to indicate physical or communicative activities, naturally implying that the

consequences exist over a period of time. Verbs containing logical or mental semantic meaning are very rarely, used with the perfect aspect.

Have had and *have got* have the most prevalent occurrence. Both mainly mark possession in a very general sense. *Has got* has more or less the same meaning as *have*.

76. *He has got the keys. = He has the keys.*

Often *have* is omitted leaving out the past participle when referring to a current possession. The uses of *have had* and *have got* differ in British and American English, but this would burst the limit of the present thesis. (Biber, Johansson, Leech, Conrad & Finegan 1999: 461-469)

Just as the perfect aspect also the progressive aspect is mostly used with the present tense. Predominantly the semantic fields of physical and communicative activities, physical situation and mental, attitudinal and perceptual states go hand in hand with the progressive aspect. Both dynamic and stative verbs can occur, even though it is said that stative verbs are seldom used with the progressive aspect. Some verbs are nearly exclusively used to refer to the progressive aspect. Those verbs which often occur with the progressive aspect are also verbs which are in general very common. If the human subject functions as an agent that controls the action, then the progressive aspect is used. Also if a situation, state or an action is described. However, when talking about perceptions then rarely occur with the progressive aspect. In conclusion the progressive aspect is used when referring to situations which are in progress, typical words used are activity verbs like *bring*, *drive* and *walk*. (Biber, Johansson, Leech, Conrad & Finegan 1999: 471-475)

2.5.6 Frequency of active and passive voice

In conversation passive voice is very rarely used. Only 2 per cent of finite verbs are passives. The *get* passive is a grammatical form, which is even more seldom used, only 0.1 per cent of all passives belong to this passive form. The

passive is used to put back human-centered concerns. In spoken English, however, this is hardly the case as human concerns in speech refer to the speakers. (Biber, Johansson, Leech, Conrad & Finegan 1999: 475-477)

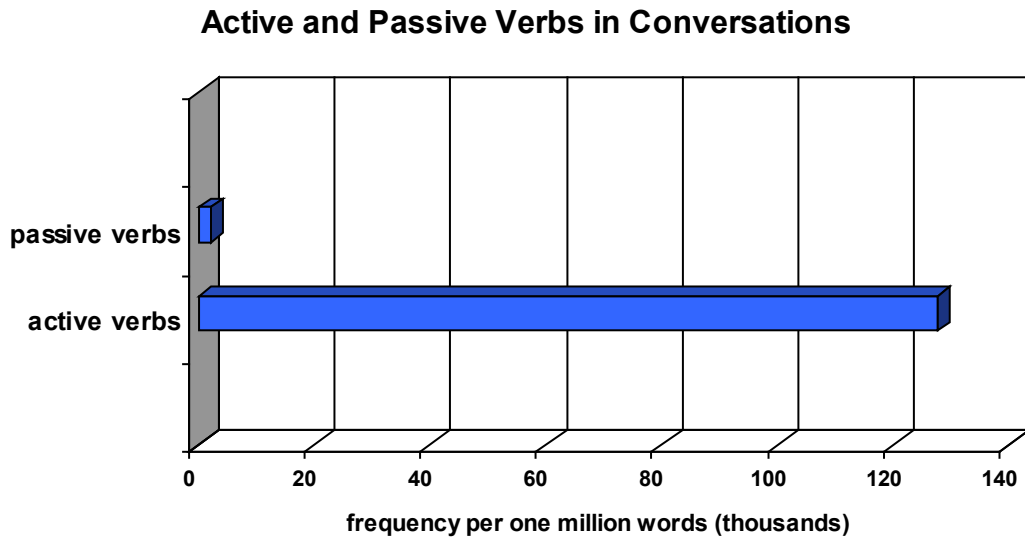


Diagram 14. Voice Distribution (LSWE)

The LSWE says that the scarce use of the *get* passive is caused by its recent emergence. Only a small number of words are commonly used with the *get* passive, for instance *get married*, *get hit*, *get involved*, *get left*, and *get stuck*. Generally *get* passives have a negative connotation. While be-passives refer to stative meanings, the *get* passive has a dynamic implication. (Biber, Johansson, Leech, Conrad & Finegan 1999: 477-481)

3 Research Methodology in Applied Linguistics

The research of data requires a hypothesis, thorough planning and the use of research methods in order to verify the conclusion. The choice of methods depends on their suitability for the research. All research which is conducted has to be reliable and valid. In this chapter it will be discussed how linguistic research is planned in classroom discourse, which questions one has to ask oneself before starting with the analysis and what outcome is expected. Furthermore, the research methods used for the data analysis are described in detail. First of all it has to be noted that classroom research is every kind of empirical examination carried out in a classroom. A classroom is every room in which teaching in a planned way takes place. In applied linguistics the classroom is a primary research site investigating how teaching and learning take place in context. However, the research in this thesis does not aim at investigating how teaching or learning take place but to examine speech events which are compared to LSWE and BNC corpus findings. Therefore, not typical classroom research is conducted but a corpus analysis comparison. Corpus linguistics studies the language of real life situations. Corpus linguistics uses for example “frequency information about occurrences of particular linguistic phenomena” (Baker: 2006:1) In order to do so quantitative research methods are used. Computerized data makes quantitative methods necessary which represent association patterns in numbers. (Baker 2006:1) (Biber 1998: 4) Quantitative research is applied to describe the numbers of verb forms and grammatical functions which shall be compared.

3.1 *An Introduction to Research Methodology*

Research is a means of reaching answers for a question, in other words finding an answer to ones hypothesis. There are two ways to achieve this: 'conceptual' and 'empirical' research. The former aims at answering the research question by secondary or also called library research. When conducting conceptual research one analyses what experts in the field have already found out. This corresponds to the former two chapters as well as this chapter as already existing experiences and theoretical findings are reproduced. The latter, however, is based on collected data.

Conclusions are drawn from results of the data analysis. This kind of research is also called primary research. (Dörnyei 2007:15-16)

The research conducted in this thesis is both conceptual and empirical, in addition to that it is corpus based. The statistics of corpora and statistics of the research are compared and either verify or do not verify the hypothesis. The method used in this research is a quantitative method. It was chosen as the results have to be illustrated by using statistical tools.

3.2 *The Roots of Quantitative research*

Quantitative research can be traced back to the nineteenth century's natural scientists who used the 'scientific method' for their exploration. This method represents data in numbers. The scientific method is the forerunner of quantitative research. Research done according to the scientific method had to follow three stages. The research process involved

“(a) observing a phenomenon or identifying a problem; (b) generating an initial hypothesis; and (c) testing the hypothesis by collecting and analysing empirical data using standardized procedures.” (Dörnyei 2007: 31)

Only testing the hypothesis did not make it generally true. First it needed to be validated by replication. Only after having established its validity it could become a scientifically satisfactory theory or a law. Scientific method researchers used tools which helped to answer and analyze questions in an objective way. These tools had to assure that the outcome was not influenced by the researcher. Only an objective analysis can give a reliable description. The scientific method is illustrated by numbers. Formerly, every research which was not described numerically was considered to be unscientific. In the 20th century not only the scientific method but also statistics were further developed and quantitative research methodology was used increasingly all over the world. In the 1970s and 1980s quantitative research became gradually more used in applied linguistics. (Dörnyei 2007:30-31)

3.3 Main features of quantitative data research

Using quantitative methods to investigate data of social sciences is problematic because when for example analyzing language patterns the data outcome depends on the moment it has been conducted, as different results will be achieved when doing the same research of language in the 19th and the 21st century. (Dörnyei 2007: 32) In applied linguistics and more precisely in corpus studies variation is a major point which has to be considered. Language is a scientific field of research which is very dynamic and changes constantly. A phrase like “*this rocks*” is today understood perfectly well but 50 years ago no one would have referred to something as “*rocking*”. Still quantitative research corresponds to the principles of the quantitative method. Nevertheless, it has some features which are not regular. The research at hand only measures verb form occurrences and grammatical functions. These normally do not change through the ages. But there are differences between conversational speech acts and task-oriented speech acts in CLIL classroom discourse. Results obtained by quantitative research are depicted by numerical data which is analyzed by statistical methods. This means that data taken from the social world are conveyed by numbers. The variables are defined before conducting the research. Quantitative research is defined by its use of numbers, its beforehand done categorization, its preference of variables rather than cases, statistics, standardized procedures used to assess objective reality and its claim to be universally true. (Dörnyei 2007: 16-34)

The usage of numbers in quantitative research is critical, as they are only relevant if the value of a number is exactly defined and the categories of the data are specified. Both the boundaries of a category and the content of the category need to be defined. For this research this means that one category is called primary modal verbs, forms of *do*, *have* and *be* are included, modal auxiliary verbs and lexical verbs, however, are excluded. The numbers given illustrate their occurrence in either a single lesson or history or marketing lessons in general. Most importantly numbers are only valid and reliable if the categories are defined and described precisely. (Dörnyei 2007: 32-33) An important aspect of categorization is that each category is absolutely unambiguous. Preparing these categories and embedding them into a

software, which counts them, is a task which might last for weeks or even month. However, after this has been achieved, the research analysis does not take as long as its preparation. (Dörnyei 2007: 33)

Another characteristic of quantitative research is that it is more concerned with common features of a group of people than with individual features. It is interested in the overall picture of the analyzed data, as it studies variables and not cases. Due to the fact that quantitative research aims to identify common features, for example tense usage in history lessons in contrast to marketing lessons, its study is based on variables. These variables are quantifiable and can be counted as well as scaled. (Dörnyei 2007: 33)

Counting makes the need of using statistics obvious. Averages may be calculated or absolute numbers of occurring verb forms are illustrated by graphs and tables. Also occurrences per 1000 words or relative numbers are acceptable. In order to be objective and to assess data in such a way it is necessary to agree on a standardized procedure. This should guarantee that other researchers get the same results, which then lead to the fact that qualitative research outcomes may be said to generally true. (Dörnyei 2007: 33-34)

On the basis of these characteristics quantitative research produces reliable and replicable data as it is very focused and controlled. However, this is a cause of criticism by qualitative researchers, as only distinct groups are observed this research fails „to do justice to the subjective variety of an individuals life“(Dörnyei 2007: 34-35) Branner even says that quantitative research is too simplistic and reductionist in order to claim general truth (2005:7). Nevertheless, the aim quantitative methodology is to ban subjectivity and to ensure that the research is objective and the research remains stable also when investigations are made by other people. (Dörnyei 2007: 34)

All these factors like the interest in variables and standardized procedures and their expression by numbers and statistics which convey the objective reality of the worlds show that quantitative methodology aims at the discovery of generalizable truth's, wide- ranging accuracy and universal laws. A question which can not be

answered is if there are universal laws when it comes to people's behavior and their language usage. (Dörnyei 2007: 34)

3.4 *The advantages and disadvantages of quantitative research*

The benefits of quantitative research are overwhelming. Some of its strengths are that it is

“systematic, rigorous, focused, and rightly controlled, involving precise measurement and producing reliable and replicable data that is generalizable to other contexts.” (Dörnyei 2007: 24)

Additionally, the statistical tools are very sophisticated and extensive. Some of them even include quality checks which are helpful when deciding if the figures are valid. Although collecting and analyzing data can be time consuming it nevertheless is, in comparison to other methods, quickly done and, what is more important, valuable. A benefit is also that the analysis' accuracy is achieved by using statistic computer software.

Quantitative methodology does also have some drawbacks, which are often stated by qualitative researchers. They claim that this kind of methodology is too simplistic and decontextualized to capture the meaning of the participant's actions. As quantitative methodology often works with averages of responses it is impossible to give reliable data on the whole set of participants. Accordingly, subjectivity and difference in an individual's behavior can not be stated and embedded into the results. (Dörnyei 2007: 35) Furthermore, it is of no importance for quantitative researchers to find out “the reasons for particular observations, or the dynamics underlying the examined situation or phenomenon” (Dörnyei 2007: 35). Nevertheless, quantitative research does not aim to arrive at conclusions but it illustrates statistically the numerical outcome of countable information.

3.5 How to do Empirical Investigation using Quantitative Research Methods

As quantitative research bases its findings on already existing theoretical findings it follows a deductive approach. The hypothesis is developed on known theory and the data analysis aims at proving it. From the hypothesis the necessary methodology is developed by which data is generated. The data is then analyzed which leads to a result from which it can be deduced that the hypothesis is right or wrong. (Rasinger 2008: 10) The illustration below shows the stages of quantitative research.

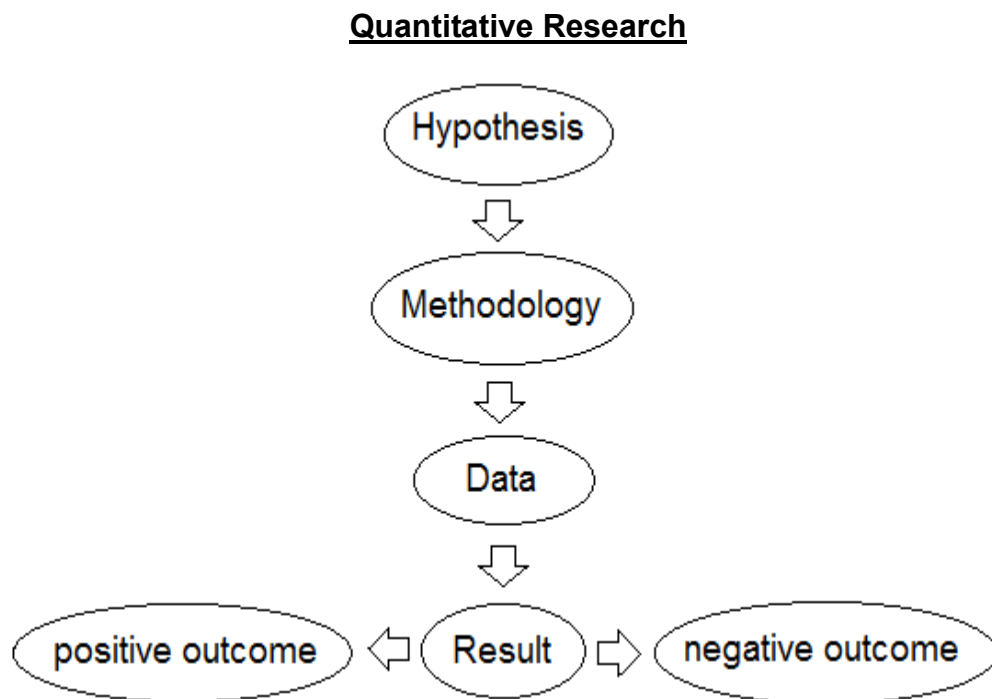


Diagram 15:. Illustration of Quantitative Research

After having formulated ones hypothesis or research questions one develops the methodology which consists of a number of instruments capable of measuring the numerical outcome of the data in such a way that the end result has either a positive or negative outcome. The development of relevant methodology demands analytical tools which analyses the data. The collected data has to suit the methodological framework and has to be reliable and valid. Normally such frameworks have to be tested in order to assure their reliability. If the data is considered to be valid it is analyzed and illustrated in numbers. At this point one compares the results with ones

hypothesis which shows if the research outcome is positive and the hypothesis has been proved to be true or not. (Rasinger 2008: 10-15)

3.5.1 Data Collection

Now it will be described how data can be collected, how it can be analyzed and reported as this needs a lot of consideration and guidelines. All the information is based on the findings of Zoltán Dörnyei in *Research Methods in Applied Linguistics*. (2007: 96-100)

First of all the researcher has to agree on the sample group of an acceptable sample size. The population and the representativeness of both have to be taken into consideration. (Dörnyei 2007:96) This means that “the population is the group of people whom the study is about” and “the sample is the group of participants whom the researcher [...] examines” (2007: 96). A sample is claimed to be good if it is alike the population in its characteristics but also if it is similar to all the other presupposed features. In fact the sample represents the whole population. This is highly important as the more accurate the sample is to the population the more representative it is. (2007: 96)

Quantitative research can be done on the basis of a variety of sample procedures. The main categories are *probability* and *non-probability sampling*. Probability sampling involves various scientific procedures such as random sampling. Non-probability sampling is favored by applied linguistics. In quantitative research this is problematic as non-probability samples are not as precise and representative. Kinds of non-probability sampling are quota sampling and dimensional sampling, snowball sampling and convenience or opportunity sampling. The samples for this study have been taken from a pool of CLIL classroom discourse transcripts. As history and marketing lessons shall be compared, appropriate samples have been chosen from the set of transcripts. This corresponds to quota sampling as according to a sampling frame a subgroup is defined. Samples which match this frame are then chosen. Also convenience or the so called opportunity sampling is used here. As the opportunity to do such a research arose and it was convenient to take samples from

an already existing set of transcripts. The members of the target population met the criteria and the availability of the data was given. As the participants met the characteristics necessary opportunity sampling is another way of sampling used for the data analysis in chapter 4. (2007: 97-98)

3.5.2 Data analysis

For the purpose of stating the number of occurrences of verb forms and grammatical functions the following guidelines have been followed.

Additionally to priming, coding and computerizing data it is necessary to screen and clean the data first. Initial data may have mistakes which have to be found and corrected. For the analysis in chapter 4 this had also to be done as sometimes verbs were misspelled or inflections were added. These factors would only be important for an analysis of pronunciation. (Dörnyei 2007: 202) For example, *happenah* had to be changed into *happen* and added into the category of lexical verbs and added to the other forms of *happen* like *happened* or *happening*.

Analyzing data is a very time consuming process, quantitative data uses statistics for its evaluation. Before analyzing the data it has to be primed for analysis. First of all the recordings of classroom discourse have to be coded and then computerized in order to run it through the software. Quantitative data software analyses data numerically. So first the information which shall be analyzed has to be transformed into numbers. (Dörnyei 2007: 195-199) For example, if one wants to find out the difference in lexical and primary verb distribution all lexical and primary verbs have to be found and numbered first. Typing in the numbers needs a lot of concentration as mistakes distort the outcome.

The outcome of the analysis has to be illustrated. Descriptive statistics is “used to summarize sets of numerical data in order to conserve time and space.” (Dörnyei 2007, 209) All the scores are listed to show the overall picture of the data outcome. As in this thesis two groups are compared, history and marketing classes and their verb usage. The results were illustrated in a table.

3.5.3 Reliability and Validity

In order to be of importance research has to be reliable, valid and verifiable. Dörnyei claims that reliability and also measurement and research validity have to be given for a research being of good quality. (2007:49-50)

According to Rasinger the concept of reliability is used to refer to a method which always, meaning at all times, measures what it was designed to measure. The method is supposed to lead to the exact same outcome when applying it to the same variables at any other point of time. In linguistics it is nearly impossible to do a research with the exact same variables at differing moments of time. The reason is that the subjects researched change their character over time as well as they change their behavior due to additional experiences they made. This means that the variables can never be 100 percent the same. (2008: 28-29) Still it has to be ensured that the tool used to analyze the data stays reliable over time because if the variation of the variables is the cause of inconsistency then the results are considered to be unreliable (Dörnyei 2007: 50). In addition to that the method should also be able to be used by other researchers and the same results should be delivered. For this an exact definition of the variables is essential. (Rasinger 2008:28-29) Linguistic research and especially quantitative research are always based on variables. A variable is considered to be a “measurable features of a particular case” (Rasinger 2008: 18). A case can be a person, groups of people or even linguistic units. Per definition a case is a particular entity which depicts the feature which is measured. (Rasinger 2008:18) In linguistics and every other kind of research tests are necessary to prove the validity of a result. Reliability can be tested in two different ways. Firstly, by using the split-half method and secondly, by applying the test-re-test method. For the former the data set, samples are split and the method used for the whole sample is applied for the test sample. The result should be the same or at least nearly the same. In test-re-test method the same set of samples is tested again after a period of time. (Rasinger 2008: 28-30) Summing up, not the test or the measurement tools can be regarded as being unreliable it is the results. The reliability of each score has to be given. To ensure the reliability of each score such tools have to be chosen which, no matter who uses them, lead to the same results.

Researchers strive for the highest level of accuracy when doing a research therefore another factor, apart from reliability, has to be regarded when doing a quantitative research. This factor is called validity. Dörnyei claims that one has to differentiate between *research validity* and *measurement validity*.

Research validity is concerned with the value of the whole examination and in more detail it aims at the distinction between *internal validity* meaning to make sure that the result is gained by a function of the variables and *external validity* which seeks for generalizability. A study is internally invalid if other factors than the ones considered have affected the findings, it is externally invalid if the results can not be generalized and are only relevant for the one study. (Dörnyei 2007: 50-52) Measurement validity or construct validity refer to the test results' interpretation being meaningful. (Dörnyei 2007: 50) In linguistics measurement validity is of special importance. As this term is difficult to explain Rasinger's definition is quoted here:

“[it] refers to the issue of whether our method actually measures what it is supposed to measure allowing us to draw appropriate conclusions.”
(2008:30)

Of course this can only be achieved by defining ones variables in advance. The main problem is if the data is influenced by something and we as researchers do not know by what and cannot even measure by what. (Rasinger 2008: 30-31) The key points of measuring validity are that validity measures the quality of the interpretation made and not the research tools or results. Validity can not be proven to 100 per cent, but plausible arguments can be given. Another fact is that validity is specific to certain situations and interpretations can not be applied form other situations. Finally, validity can be sustained by various kinds of evidence. (Rasinger 2008:18-52)

4 DATA ANALYSIS

The final part of the thesis illustrates the findings based on CLIL classroom discourse transcripts of history and marketing lessons. Each transcript is divided into teacher and student talk. The verb forms discussed in the theoretical discussion of the English verb functions as a basis, as well as the corpus findings which will be compared to the outcome of the analysis. It will be shown which lexical verbs are most frequently used. Also the use of tense, voice, aspect and modality are of interest and numbers of their occurrence will be given.

4.1 Objectives

In this thesis 4 CLIL classroom discourse transcripts are analyzed and compared. The basis are two history and two marketing lessons each. CLIL classroom discourse is analyzed in order to examine if grammatical forms and functions as well as the occurrence of semantic domains can be compared to the LSWE and BNC. It is assumed that the language use by the participants of CLIL discourse is similar to spoken English of native speakers in terms of tense, verb form and semantic domain distribution. The hypothesis is that the most commonly used tenses are the present and past tense simple, both in teacher and student talk. Furthermore, it is presumed that the past tense is more often used in history lessons than in marketing lessons where present tense and possibly future tense are more common, due to the fact that in history lessons past events are discussed and marketing lessons are concerned with facts and what the outcome of a marketing strategy will be. It therefore has to be noted that the verb choice strongly depends on the subject, as certain topics demand a special choice of tense, semantic domain as well as verb form usage. Consequently the list of most frequently used lexical verbs might not equal the one found in the BNC and LSWE. Student talk and teacher talk are individually analyzed and then compared. It also has to be analyzed if teachers use a higher variety of grammatical verb forms than students. It is supposed that students mostly use simple past and present tense forms, as well as simple passive constructions. Both progressive as well as perfect aspect are supposed to mostly be used by teachers, while modal verbs will be used by both target groups. The variables of the study, therefore, are on the one side the teacher of the history and

the marketing class, as well as the students in both classes. Further features are that the grammatical structure and tense frequency of each of the groups is counted. All the individual features analyzed are variables, for example primary, modal and lexical verb distribution.

In this research the population is the CLIL teachers and learners and the samples would be history and marketing classes using CLIL. A drop-out rate did not have to be taken into account as the participants have already agreed to be recorded and the recordings have already been transcribed. (Döyneyi 2007: 99-100) The findings are based on tokens which are the single verb forms and these are summarized into types, for example all the verb forms of *do* are then summarized under the type primary verb-

4.2 HIST 1

The transcript HIST 1 was collected on November 19, 2001. It is a history lesson conducted in eleventh grade (HFB). Its topic is social hierarchy in ancient Greece. The class consists of 35 students. An overall number of 6,207 words are analyzed 3,577 words of teacher talk and 2,630 of student talk. In teacher talk an overall number of 601 verbs were found, which is 17.1 per cent of all words. The data of HIST 1 student talk consists of 2,630 words. 252 of these words are verb forms. This is 9.6 per cent which is 8 per cent less than in the teacher talk. A large variety of verb forms and grammatical forms could be found which will be described in detail in the following subchapters.

4.2.1 Word Class Distribution

According to the LSWE corpus, lexical words are the most common verbs with over a hundred thousand occurrences per one million words. Primary and modal verbs are far less often used with 30 and 18 thousand occurrences. All the numbers of the LSWE are found in Biber. (Biber, Johansson, Leech, Conrad & Finegan 1999: 359-360) The table below illustrates the absolute and relative distribution of these three verb classes in teacher and student talk, respectively. Furthermore the overall

distribution in the first analyzed history lesson, HIST 1, is shown in comparison to the relative outcome of LSWE findings. The numbers and percentages show that in HIST 1 the distribution differs greatly from the LSWE findings. First the results of teacher talk and then of student talk are stated before looking at the overall picture.

Distribution of Verb Classes in HIST 1

TEACHER			
	Primary Verb	Modal Verbs	Lexical Verb
Teacher absolute frequency	301	47	253
Teacher relative frequency	50.1	7.8	42.1
STUDENT			
Student absolute frequency	94	29	129
Student relative frequency	37.3	11.5	51.2
Classroom overall %	46.3	8.9	44.8
LSWE			
LSWE %	19	11.4	69.6

Table 14. Distribution of Verb classes in HIST 1

In teacher talk primary verbs are the principal verb class with 50.1 per cent. Lexical verbs follow closely with 42.1 per cent. This differs greatly from the LSWE corpus as primary verbs are far less common there with a relative number of 19 per cent. The findings of student talk are more similar to the ones of the LSWE corpus. 37.3 per cent of primary verbs oppose 51.2 per cent of lexical verbs. Overall the class of primary verbs shows that primary verbs are slightly more common than lexical verbs. In contrast, according to the LSWE corpus, the class of lexical verbs is by far most common. The only parallel between HIST 1 and the LSWE is that the modal verb class is the least used verb class in both sets.

4.2.2 Primary Verb Distribution

As the high number of primary verbs is surprising a special focus needs to be laid on their distribution in their function as modal auxiliary verbs and main verbs. Neither the LSWE nor the BNC give account of this division. Nevertheless, due to the unusual verb class distribution the need to investigate primary verbs in more detail

arose. Table 15 shows the share of auxiliary and main verb usage in absolute as well as in relative numbers also the overall number of both sets of speech is stated.

Primary Verb Distribution in HIST 1

TEACHER		
	Auxiliary Verb	Primary Verb
Teacher absolute frequency	98	203
Teacher relative frequency	32.6	67.4
STUDENT		
Student absolute frequency	31	63
Student relative frequency	33	67
Classroom overall %	32.7	67.3

Table 15. Primary Verb Distribution in HIST 1

The distribution of primary verbs in their function as auxiliary and main verbs surprisingly is the same in both student and teacher talk and therefore also in the overall distribution of the data set HIST 1. Their main verb usage of 67 per cent is higher than the auxiliary verb distribution with 33 per cent. The share of *be*, *do* and *have* is illustrated below. The numerical occurrence and the relative number of occurrence in the primary verb class are given. In addition to the findings in the single speech sets also the overall circulation in HIST 1 and the results of the BNC (Leech 2001: 282) in regular conversations and in task-oriented speech are given. Classroom discourse is a kind of task oriented speech, the analysis will show if HIST 1 discourse is more similar to conversational speech or task-oriented speech. It of course should rather correspond to the latter.

Table 16 shows how often each primary verb is used in HIST 1 also the findings of the BNC (Leech 2001: 282) are given in order to compare them to the classroom data. Both in teacher as well as in student talk *be* is the most common primary verb followed by *have* and *do*. The share of *be* in the overall distribution is 60.9 percent, the one of *do* 17 per cent and of *have* 22.1 per cent. These results correspond to the BNC figures of task-oriented speech. For example, nearly 61 per cent of *be* usage in HIST 1 are opposed to 63 per cent in task-oriented speech. The primary verb *do* shows the highest divergence of 1.9 per cent, while *have* shows nearly the same result with 22.1 per cent in HIST 1 and 21.6 per cent in

BNC task-oriented speech.

Distribution of BE, DO ad HAVE

TEACHER			
	BE	DO	HAVE
Teacher absolute frequency	181	55	65
Teacher relative frequency	60.1	18.3	21.6
STUDENT			
Student absolute frequency	59	12	22
Student relative frequency	62.7	13.8	23.4
Classroom overall %	60.9	17.0	22.1
BNC			
BNC % in conversations	58.4	20.5	21.1
BNC % in task-oriented speech	63.3	15.1	21.6

Table 16. Distribution of BE, DO and HAVE in HIST 1

BE

After having established the overall distribution, a closer look is taken at the single primary verbs. First the auxiliary and primary verb distribution is given, then their verb form occurrence and finally for which functions they are used as auxiliary verbs. This structure is used firstly for *be*, secondly for *do* and thirdly, for *have*. The auxiliary and main verb distribution is illustrated in table 17.

Distribution of BE as Auxiliary and Main Verb

TEACHER		
	Auxiliary Verb	Main Verb
Teacher absolute frequency	35	138
Teacher relative frequency	19.3	76.2
STUDENT		
Student absolute frequency	11	48
Student relative frequency	18.6	83.1

Table 17. Distribution of BE as Auxiliary and Main verb in HIST 1

Although the primary verb *be* can be used for the formation of a variety of different grammatical formations its usage as a auxiliary is very low. In teacher talk the auxiliary has a share of 19.3 per cent. The difference to student talk is not as

distinct, with 18.6 per cent, as one would suspect due to the students' lower proficiency level. Summing up, *be* is more often used as a main verb and its frequency in the two speech samples does not differ greatly.

The next issue is to show which forms of the verb are most common in both teacher and student talk. Their relative distribution is stated as well as the figures of the BNC (Leech 2001: 297). Only those verb forms which occur in the data are listed.

Verb form Distribution of BE

TEACHER						
	to be	be	is	are	was	were
Absolute frequency	8	0	61	26	36	50
Relative frequency	4.4	0	33.7	14.4	19.9	27.6
STUDENT						
Absolute frequency	0	2	17	16	12	12
Relative frequency	0	3.4	28.8	27.1	20.3	20.3
Classroom overall %	3.3	0.8	32.5	17.5	20	25.8
BNC						
BNC % in conversation	7.2	1.1	50.5	15	13.7	7.8
BNC % in task-oriented speech	12.4	0.1	41.1	17.5	13.8	5.5

Table 18. Verb form Distribution of BE in HIST 1

Of all of the seven forms the verb *be* can take only 5 are used by the teacher and the students, respectively. The former uses the infinitive and the two present tense forms *is* and *are* as well as both past tense forms. The latter also *is* and *are*, and *was* and *were* can be found but instead of the infinitive the base form occurs. Comparing the BNC to the data analysis results one sees an enormous difference in the distribution of verb forms. According to the BNC, in both task-oriented and conversational speech, the –s form is by far the most common verb form of the primary verb *be*. The following verb forms are the present plural form, the –s form of the past tense, the infinitive, the plural past tense form and the base form. In teacher and student talk, however, the past tense forms follow the present tense –s form. All in all, the findings show a quite different distribution than the BNC only when it comes to the present tense form *is* and the base form the findings are similar.

The following table illustrates for which grammatical functions the primary verb *be* is used. Progressive aspect, passive voice, *going to* future, the substitution of a modal verb as in *was able to* instead of *can* and negations are the only functions found in HIST 1. The table shows their absolute frequency of present and past tense as well as absolute and relative overall frequency. From this one can also deduce which tense is predominant, namely the past tense

Grammatical Functions of BE

TEACHER							
	Prog. Aspect	P. Voice	going to	Substitution Verb	Neg.	Sum	% of verb form use
is	3	3			1	5	14.3
are	1	3		1		4	11.4
Present	4	6		1	1	12	34.3
was	3	5	1	1		10	28.6
were		6		7		13	37.1
Past	3	11	1	8		23	65.7
Overall	7	17	1	9	1	35	100
% of overall usage	20	48.6	2.9	25.7	2.9		
STUDENT							
is		2				2	18.2
are		1		1		2	18.2
Present		3		1		4	36.4
was		2				2	18.2
were		2		3		5	45.4
Past		4		3		7	63.6
Overall		7		4		11	
% of overall usage		63.6		36.4			

Table 19. Grammatical Functions of BE in HIST 1

The modal auxiliary occurrences of teacher talk are discussed before the ones in student talk. In HIST 1 teacher talk progressive aspect, passive voice, *going to* future, negations and the use as a substitution verb occur. The most applied grammatical function is passive voice with nearly 50 per cent. As mentioned above the most used form of *be* is the tensed verb *were* followed by *was*. Therefore, the past tense is predominant with 65.7 per cent. Contrary to the various uses of the modal auxiliary in teacher talk it is uncommon in student talk. The students in HIST 1 only use *be* to express passive voice and as a substitution verb, voice leads with 63.6 per cent. All

in all passive voice is the most frequently used grammatical function followed by the primary verbs use as a substitution verb. Moreover past tense is more regularly used than the present tense.

DO

The next primary verb discussed is *do*. The table below illustrates the absolute and relative distribution of the modal and main verb function in the teacher and student data set.

Distribution of DO as Auxiliary and Main Verb

TEACHER		
	Auxiliary Verb	Main Verb
Absolute frequency	41	14
Relative frequency	74.5	25.5
STUDENT		
Absolute frequency	4	9
Relative frequency	30.8	69.2
Classroom overall %	65.7	34.3

Table 20. Distribution of DO as Auxiliary and Main Verb in HIST 1

In HIST 1 teacher talk the auxiliary verb use of the primary *do* is more frequent than its main verb usage. 74.5 per cent opposing 25.5 per cent speak for themselves. Student talk shows a different distribution, main verb usage is more common than modal verb usage. The main verb has a share 69.2 per cent and auxiliary verb 30.8 per cent. When adding these figures the result is that the auxiliary verb usage leads with a total of 65.7 per cent. The fact that *dummy do* is so predominant leads to the conclusion that the most frequent forms are *do*, *does* or *did*. The following table verifies this and it furthermore gives information on the exact verb form distribution as well as equivalent figures from the BNC.

Verb Form Distribution of DO

TEACHER						
	do	does	did	done	to do	base
Absolute frequency	12	3	33	1	1	5
Relative frequency	32.7	5.5	60	1.8	1.8	9.1
STUDENT						
Absolute frequency	4	1	4		3	1
Relative frequency	30.8	7.7	30.8		23.1	7.7

BNC Findings					
BNC % in conversations	45.6	8.7	21.8	5.51	12.8
BNC % in task-oriented speech	38.8	10.1	21.1	6.6	17

Table 21. Verb Form Distribution of DO in HIST 1

The verb form distribution in teacher and student talk differs greatly, also in comparison to the BNC. According to the BNC, The present tense form of the primary verb *do* is the most common one. Followed by the past tense form, the infinitive and the –s form. However, in the teacher’s talk the past tense form is by far the most common verb form followed by the present tense form. In student talk both present and past tense forms are equally frequent with 30.8 per cent. These forms are followed by the infinitive. In conclusion it may be said that the results of the teacher’s talk differs more to the BNC findings than the results of student talk. (Leech 2001: 297-298)

The high share of auxiliary use and the interesting verb form distribution make it necessary to discover for which purposes *do* is used in the formation of grammatical functions. The dummy *do* can only be used to express questions and negations in both the present and the past tense, apportionment of both sample sets of HIST 1 is depicted.

Grammatical Functions of DO

TEACHER				
	Question	Negation	Sum	% of verb from use
do	4	5	9	22
does	3		3	7.3
Present:	7	5	12	29.3
did	14	15	29	70.7
Past:	14	15	29	70.7
Overall:	21	20	41	
% of overall distribution of to do	51.2	48.8		

STUDENT				
	Question	Negation	Sum	% of verb from use
do	1		1	25
does				
Present	1		1	25
did	1	2	3	75
Past	1	2	3	75
Overall	2	2	4	
% of overall distribution of to do	50	50		

Table 22. Grammatical Functions of DO in HIST 1

The findings in teacher talk and student talk are fairly similar. In the former questions are slightly more common with 51.2 per cent than negations with 48.8 percent. Past tense as already assumed is more often employed with 70.7 per cent than the present tense with 29.3 per cent. In the latter the distribution of question and negation is identical with 50 per cent each. This is surprising as the word form list hints at a different result.

In conclusion the primary verb *do* is more frequent as an auxiliary. Furthermore the past tense is the leading tense and the verb form distribution in teacher talk does not equal the one represented in the BNC while the one of student talk is more similar to the BNC.

HAVE

The third and last primary verb which has to be discussed is *have*. This word is necessary in the formation of aspect, passive voice and is also often applied in its function as substitution verb for the modal auxiliary *must*. Since it forms so many grammatical functions it is interesting to investigate its distribution as main and auxiliary verb. As in the previous primary verbs its frequency as auxiliary and main verb is given as well as the hierarchy of verb forms. Table 10 shows the primary's share as auxiliary and main verb in both teacher and student talk.

Distribution of HAVE as Auxiliary and Main Verb

TEACHER		
	Auxiliary Verb	Main Verb
Absolute frequency	22	43
Relative frequency	33.8	66.2
STUDENT		
Absolute frequency	12	10
Relative frequency	54.5	45.5
Classroom overall %	39.1	60.9

Table 23. Distribution of HAVE as Auxiliary and Main Verb in HIST 1

The teacher in HIST 1 uses the primary *have* more often as a main verb than as an auxiliary verb. Two thirds of all forms belong to the main verb class and only one third to the auxiliary verb class. In student talk the discrepancy in the share is not so

distinctive. In contrast to teacher talk, the students use the primary more often as an auxiliary verb with a share of 54.5 per cent. In order to get a first idea which tense is more frequent and also for which grammatical functions *have* is used. A closer look is taken on the hierarchy list of primary verb forms.

Distribution of HAVE's Verb Forms

TEACHER					
	have	has	had	to infinitive	base form
Absolute frequency	26	5	22	1	11
Relative frequency	40	7.7	33.9	1.5	16.9
STUDENT					
Absolute frequency	3	3	16		
Relative frequency	13.6	13.6	72.8		
BNC findings					
BNC % in conversations	44.5	17.9	13	18.6	
BNC % in task-oriented speech	46.4	15.9	14.6	17.4	

Table 24. Distribution of HAVE's Verb Forms in HIST 1

As can be deduced from the table the results of the BNC are more or less equal to the findings in teacher talk but they hardly correspond to HIST 1's student talk. According to the BNC the present tense form *have*, including the base form, are most often used with over 40 per cent in both conversations and task-oriented speech. This is also true for teacher talk, if adding the number of the base form to the present tense form the number is even higher than in the BNC, with 56.9 per cent. A major difference is the high usage of the past form in the teacher's talk. In student talk the past tense form is by far the most frequently used form with over 70 per cent. In contrast, the past shows a distribution of 13 or 14.6 per cent in the BNC corpus. (Leech 2001: 298)

After having mentioned the most important facts I will go on to analyze for which grammatical functions *have* is used and also which tense is more common.

Grammatical Functions of HAVE

TEACHER					
	Perfect Aspect	Substitution Verb	Neg.	Sum	% of verb form use
have	6	6	1	7	31.8
has	2	1		3	13.6
Present:	8	7	2	17	77.3
had	1	4		5	22.7
Past	1	4		5	22.7
Overall:	9	11	2		
% of overall distribution	40.9	50	9.1	100	

STUDENT					
	Perfect Aspect	Substitution Verb	Neg.	Sum	% of verb form use
have	1			2	15.4
has	2	1		3	23.1
Present:	3	1		4	30.8
had		9		9	69.2
Past		9		9	69.2
Overall:	3	10		13	
% of overall distribution	23.1	76.9		100	

Table 25. Grammatical Functions of HAVE in HIST 1

Teachers use the primary verb in order to express perfective aspect and in negations, moreover, it is applied to substitute the modal auxiliary verb *must*. The perfect aspect has a share of 40.9 per cent, and negations 9.1 per cent. Most often it is utilized as a substitution with 50 per cent. Present tense use is most recurrent with 77.3 per cent. Students only use *have* to convey perfect aspect and as a substitution. In contrast to teacher talk in student talk the past tense is predominant. All in all, the primary verb is most often used to substitute the modal auxiliary *must* and the past tense is the prime tense.

4.2.3 Distribution of Modal Verbs

In this section the modal verb class is examined. The focus is to list which modal verbs occur in the samples as well as to which tense is the most commonly used in HIST 1. As the modal auxiliary verb class is seldom used in comparison to primary and modal verbs the outcome is only briefly described. In table 13 one finds information on the distribution of the single modal auxiliaries in absolute and relative numbers, and the percentage of tense distribution in the modal verb class, both in the two subsets and also in the whole data set of HIST 1.

Distribution of Modal Verbs

TEACHER					
Present		Past		Number of distribution	% verb of distribution
can	13	could	16	29	61.7
need	1	needed	2	3	6.4
will	1	would		1	2.1
may	1			1	2.1
		should	1	1	2.1
Sum	16		31	47	
% of tense distribution (teacher)	34		66		
% of overall tense distribution	2.6		5.2		
STUDENT					
can	6	could	20	26	89.7
will		would	1	1	3.5
may	1			1	3.5
		should	1	1	3.5
Sum	7		22	29	
% of tense distribution (student)	24.1		75.9		
% of overall tense distribution	2.8		8.7		
% of overall HIST 1 distribution	30.3		69.7		

Table 26. Distribution of Modal Verbs in HIST 1

Both students and teachers use almost the same modal auxiliaries. These are *can*, *could*, *will* or *would*, *may* and *should*. In both sets of data *can* and *could* are by far most frequently used. *Can* has a share of 61.7 per cent found in teacher talk and an even higher number, 89.7 per cent in student talk. All the other modals occur very rarely. Just as in the class of primary verbs the past tense is predominant. In comparison to the overall number of verbs the frequency of modal verbs is very low which also corresponds to the LSWE corpus. (Biber, Johansson, Leech, Conrad & Finegan 1999:259-260)

4.2.4 Lexical Verb Distribution

The lexical verb group is the second largest word group in HIST 1. The results of the analysis are compared to the findings in the BNC (Leech 2001:96-297). Table 27 shows the distribution of the individual verb forms in which lexical verbs can occur.

Distribution of Main Verb Forms

		TEACHER						
		base form	-s form	present form	past form	-ing participle	-en participle	to infinitive
Absolute frequency		61	17	52	43	21	26	34
Relative frequency		24.1	6.7	20.6	17	8.3	10	13.4
		STUDENT						
Absolute frequency		18	9	25	28	5	10	35
Relative frequency		14	7	10.1	21.7	3.9	7.8	27.1
		BNC						
BNC % in conversation distribution		27.7	6.0	<i>see base form</i>	14.3	10.8	12.9	28.2
BNC % in task-oriented speech		23.3	5.1	<i>see base form</i>	9.6	12.4	17.2	32.3

Table 27. Distribution of Main Verb Forms in HIST 1

The outcome of the verb form analysis corresponds mostly to the findings of the BNC, especially when it comes to student talk. The BNC's category called base form consists of the base form of the verb as well as its present tense use. For a better representation of true present tense use in comparison to past tense use this does not apply for the data representation. The predominance of the as suggested by the corpus is only verified in student talk but not in teacher talk. The base form group is the second most frequent verb form in the BNC. Also in HIST 1 it shows a high frequency. The major difference is that the past tense is more often used in both data sets than in conversations or task-oriented speech. Especially the dissimilarity between the data and task-oriented speech is enormous.

4.2.5 Summary of Findings

In conclusion it may be said that there are similarities between the data of HIST 1 and the findings of the BNC and LSWE still there are also some divergences. The most prominent one is that the past tense is by far more often used in this history lesson than in conversations or task-oriented speech. This fact can be found in nearly all of the points discussed. Table 28 gives an overview of the most important numbers concerning grammatical functions and present and past tense use.

Distribution of Grammatical Functions

TEACHER									
	Tense		Aspect		Mood		Voice		
Number of occurrence	112		15		67		18		
% of occurrence	52.8		7.1		31.6		8.5		
	<i>Present</i>	<i>Past</i>	<i>Present</i>	<i>Past</i>	<i>Present</i>	<i>Past</i>	<i>Present</i>	<i>Past</i>	
Times of occurrence	69	43	12	3	24	43	6	12	
% of occurrence	32.5	20.3	5.7	1.4	11.3	20.3	2.8	5.7	
Overall number	Present		236		Past		215		
% of overall occurrence			52.3				47.7		
STUDENT									
	Tense		Aspect		Mood		Voice		
Number of occurrence	53		29		10				
% of occurrence	57.6		31.5		10.9				
	<i>Present</i>	<i>Past</i>	<i>Present</i>	<i>Past</i>	<i>Present</i>	<i>Past</i>	<i>Present</i>	<i>Past</i>	
Times of occurrence	25	28	7	22	6	4			
% of occurrence	39.7	60.3	24.1	75.9	60	40			
Overall number	Present		38		Past		54		
% of overall occurrence			41.3				58.7		
% of overall HIST 1 distribution			50.5				49.5		

Table 28. Distribution of Grammatical Functions in HIST 1

In the LSWE it is claimed that tense is the most used grammatical form followed by mood and then aspect. Passive voice, conversely, is the least common grammatical functions in spoken English. (Biber, Johansson, Leech, Conrad & Finegan 1999:359-360) When taking a look at the above table one sees that once again the results of student talk are more similar to both corpus findings because of the small number of verbs at hand passive voice is not used at all. The only difference is that mood is less common than aspect in student talk. In terms of tense distribution the past tense is

predominant with 58.7 per cent. This is extraordinary and does not correspond with the corpora referred to in which present tense is by far the most common tense. Turning now to teacher talk one can see that the hierarchy of grammatical functions equals the one stated in the LSWE. In contrast, to student talk all four grammatical functions appear with relatively high frequency. Another difference is that in teacher talk the present tense is more common than the past tense. Still the past tense is more frequent than claimed by the LSWE. The unusually high share of past tense forms in HIST 1, due to the fact that past events are discussed in history lessons, suggests that in CLIL history classroom discourse past tense is more common than in conversations and task-oriented speech.

4.3 HIST 2

The second history lesson transcript is named HIST 2. The lesson was recorded on November 30, 2001 also in a third grade class but in a HTA class. The teacher is the same but the group of students is not. Similar results as in HIST 1 can be expected for teacher talk. The topic of the lesson is how Athens was governed. The data consists of an overall number of 5,313 words of which 978 words belong to the class of verbs. This is a percentage of 18.4. The data is also subdivided into teacher and student talk. Teacher talk consists of 3,766 words of which 708 are verbs, which are 18.8 percent. Student talk has an overall number of 1,547 words of these only 270 are verbs which are 17.5 per cent. In HIST 2 nearly twice as many verbs occur in comparison to HIST 1. The relative numbers show that the share of verbs is very similar in teacher and student talk. HIST 2 data is analyzed in the same way as HIST 1. Also the same findings are listed in order to compare them afterwards. Not only are the numbers of word classes given but also their verb form, tense and grammatical structure distribution.

4.3.1 Word Class Distribution

Biber claims in the LSWE that the lexical verb class is the most common one with nearly 70 per cent. Primary and modal auxiliaries are by far less frequently used in spoken English with a share of 19 and 11.4 per cent respectively. Table 29 gives account of the absolute and relative distribution of both HIST 2 teacher and

student talk. Furthermore, also the numbers found in the LSWE are given in order to be able to compare them. (Biber, Johansson, Leech, Conrad & Finegan 1999: 359-360)

<u>Verb Class Distribution</u>			
TEACHER			
	Primary Verb	Modal Auxiliary	Lexical Verb
Absolute frequency	308	80	320
Relative frequency	43.5	11.3	45.2
STUDENT			
Absolute frequency	114	25	131
Relative frequency	42.2	9.3	48.5
Classroom overall %	43.1	10.7	46.1
LSWE findings			
LSWE %	19	11.4	69.6

Table 29. Verb Class Distribution in HIST 2

The analysis shows that lexical verbs are the most common verb forms in both student and teacher talk. The class of lexical verbs is followed closely by the class of primary verbs. This stands in contrast with the LSWE which shows that primary verbs are by far less used than lexical verbs. Modal auxiliary verbs are the least common verbs in both the data and the LSWE corpus. Interestingly, the relative numbers of the LSWE and the data are nearly the same. All in all, lexical verbs are more common than primary verbs but the contrast is not as distinctive as in the LSWE corpus.

4.3.2 Primary Verb Distribution

The very high number of primary verbs in HIST 2 is of interest in this section. Factors which are described are the overall share of the single primaries and individual primary verb share of auxiliary and main verb occurrence. The general number of the single primary verbs is given as well as their verb form distribution. Due to the fact that primary verbs are used to form different grammatical functions also their individual field of usage is part of the analysis.

Auxiliary and Main Verb Distribution of Primary Verbs

TEACHER		
	Auxiliary Verb	Main Verb
Absolute frequency	104	204
Relative frequency	33.8	66.2
STUDENT		
Absolute frequency	53	61
Relative frequency	46.5	53.5
Classroom overall %	37.2	62.8

Table 30. Auxiliary and Main Verb Distribution of Primary Verbs in HIST 2

As can be seen, primary verbs are in general more frequently used as a main verb than as an auxiliary verb. The overall share of the main verb function is 62.8 per cent and that of the auxiliary verb 37.2 per cent. For both teachers and students the main verb shows a higher share. The difference in teacher talk is more obvious than in student talk. Let me now turn to the distribution of the individual primary verbs. The following table shows the absolute and relative distribution of *be*, *do* and *have*. Furthermore, their overall distribution in HIST 2 data is given as well as comparable numbers from the BNC. (Leech 2001:282)

Distribution of BE, DO and HAVE

TEACHER			
	BE	DO	HAVE
Absolute frequency	199	43	66
Relative frequency	64.6	14	21.4
STUDENT			
Absolute frequency	68	13	33
Relative frequency	59.6	11.4	29
% of overall distribution in HIST 2	63.3	13.3	23.5
BNC Findings			
BNC % in conversations	58.4	20.5	21.1
BNC % in task-oriented speech	63.3	15.1	21.6

Table 31. Distribution of BE, DO and HAVE in HIST 2

In teacher talk the primary verb *be* is by far the most often used one with 64.6 per cent. *Have* and *do* are less common with a share of 21.4 and 14 per cent, respectively. Also in student talk the highest score can be found for the primary *be* with 59.6. The other two verbs are less frequent just as in teacher talk. This

correlates with the findings from the BNC corpus both in conversations and in task-oriented speech.

Next the individual primary verb functions and distributions are illustrated and described. As the primary verb *be* is the most common primary verb it is described first, followed by *do* and *have*.

BE

The first point is to find out if *be* is more common in its auxiliary or main verb usage. The table below gives the number of occurrence and the relative distribution.

Share of Auxiliary and Main Verb Use of BE in HIST 2

TEACHER		
	Auxiliary Verb	Main Verb
Absolute frequency	49	150
Relative frequency	24.6	75.4
STUDENT		
Absolute frequency	10	58
Relative frequency	14.7	85.3

Table 32. Share of Auxiliary and Main Verb Use of BE in HIST 2

The primary verb *be* is three times more often used in teacher talk than in student talk. Both show a higher frequency of *be* as a main verb than as an auxiliary verb. The difference between these functions is more pronounced in student talk. As the main verb function is so high it is interesting to investigate the verb form distribution in HIST 2. Table 33 shows which verb forms can be found in the data and what their distribution is. Furthermore, the results of the BNC are given to make comparison easier. (Leech 2001:297)

Verb Form Distribution of BE in HIST 2

TEACHER							
	be	is	are	was	were	being	been
Absolute frequency	14	80	25	41	36	1	2
Relative frequency	7.0	40.2	12.6	20.6	18.1	0,5	1,0

STUDENT							
Absolute frequency	5	27	10	15	11		
Relative frequency	7.3	39.7	14.7	22.1	16.2		
Classroom overall %	7.1	40.1	13.1	21	17.6	0.4	1.7
BNC Findings							
BNC % in conversation	1.1	50.5	15	13.7	7.8	0.6	2.8
BNC % in task-oriented speech	0.1	41.1	17.5	13.8	5.5	1.3	3.8

Table 33. Verb Form Distribution of BE in HIST 2

Teacher talk consists of nearly three times as many words as student talk, this might be a reason why in the former seven, out of eight, forms of *be* can be found and in the latter only five. The teacher uses the base form, *is* and *are*, the past tense forms, and both participles. The most common form is the –s form with a share of 40.2 per cent. The past tense’s –s form is used half as often and *were* even less. Very rare are the –ing and –en participles. The distribution in HIST 2 does, however, not correspond to the one of the BNC. The present tense’s –s form is most common in both subsets as well as in the BNC. According to the BNC the second most frequent form is the primary *are*. But in HIST 2 it is the past tense form *was*. Another difference is that the frequency of *be* is relatively high in HIST 2. In the analyzed data it has a distribution of 7.1 percent in BNC conversations only 1.1 and in BNC task-oriented speech 0.1. However, the rare use of the –ing and –en participle is found in both the BNC corpus and the analyzed. (Leech 2001:297)

The next point worth noting are the purposes for which the primary *be* is used in its auxiliary verb function. *Be* can be used to express progressive aspect, voice, future, questions and negations and also as a substitution for the modal auxiliary *can*. The share of these purposes is illustrated in the table below. Additionally, the distribution of present and past tense is given. The verb form list shows that the present tense is predominant, if this is true can be seen in the list.

Grammatical Functions of BE

TEACHER								
	Prog. Aspect	Prog. Perf. Aspect	P. Voice	going to	Substitution Verb	Neg.	Sum	% of verb form use
is	4		3	1		2	10	20.4
are	3	1	5	1	3		13	26.5
Present	7	1	8	2	3	2	23	46.9
was			7			2	9	18.4
were			10		5	2	17	34.7
Past			17		5	4	26	53.1
Overall	7	1	25	2	8	6	49	
% of overall usage	14.3	2	51	4.1	16.3	12.2		

STUDENT								
	Prog. Aspect	Prog. Perf. Aspect	P. Voice	going to	Substitution Verb	Neg.	Sum	% of verb form use
is	7						7	28
are			1				1	4
Present	7		1				8	32
was			6		3		9	36
were			7		1		8	32
Past			13		4		17	68
Overall	7	0	14	0	4	0	25	
% of overall usage	28	0	56	0	16	0		

Table 34. Grammatical Functions of BE in HIST 2

In teacher talk the primary verb *be* is used to express aspect, passive voice, *going to* future and negations, and it is also used as a substitution verb. The most common structure is passive voice with a frequency of 51 per cent. Aspectual combinations, negations and the substitution usage are less often used. All these grammatical forms are more often found in past tense than in present tense. The difference is, however, not very distinctive. Students do not use such a variety of grammatical functions. In their talk only progressive aspect, passive voice and the substitution are applied. These mostly occur in their past tense form. The similarities between the sets of data are that passive voice is most common with more than 50 per cent and the past tense is the dominating tense.

DO

The next primary verb discussed is *do*. This verb is called *dummy do* when

used as an auxiliary verb as it does not convey any meaning and is only used to form questions and negations in present and past tense. The first point is its main and auxiliary verb distribution which is shown in the following table. Its frequency in both data sets is given as well as the overall share.

Share of Auxiliary and Main Verb Use of DO

TEACHER		
	Auxiliary Verb	Main Verb
Absolute frequency	33	10
Relative frequency	76.7	23.3
STUDENT		
Absolute frequency	3	10
Relative frequency	23.1	76.9
Classroom overall %	64.3	35.7

Table 35. Share of DO as an Auxiliary and Main Verb in HIST 2

The absolute number of occurrences of *do* is considerably higher in teacher talk than in student talk. It has to be noted that conclusions based on absolute figures are problematic as teacher talk consists of more than twice as many words as student talk. Therefore comparisons are made on the basis of relative figures. In teacher talk the auxiliary verb usage is most common with a frequency of 76.7 per cent. A theory of why this is so, might be that the teacher is bound to ask a lot of questions which makes the use of the *dummy do* necessary. In student talk, on the other hand, the main verb use is more common. Consequently, overall main verb use is less frequent with 35.7 percent than auxiliary verb use with 64.3 per cent. Before describing the share of questions and negations the verb form distribution is illustrated by the next table. Also the BNC's findings are given in order to compare the numbers (Leech 2001:297-298).

Verb Form Distribution of DO

TEACHER							
	do	does	did	done	doing	to do	base
Number of occurrences	21	3	16	1	2		5
% of overall <i>do</i> forms	48.8	7	17.2	2.33	4.65		11.6

		STUDENT					
Number of occurrences	of	7	1	5		1	
% of overall forms	<i>do</i>	53.9	7.7	38.5		7.7	
		BNC findings					
% in conversations		45.6	8.7	21.8	5.51	5.6	12.8
% in task-oriented speech		38.8	10.1	21.1	6.6	6.4	17

Table 36. Verb Form Distribution of DO in HIST 2

The primary *do* can occur in seven different forms. In teacher talk six of them occur in student talk only four. All of the other forms except of *do* and *did* are relatively rare. According to the BNC *do*, including the present tense form and base form, dominates followed by the past tense form and infinitive. The other verb forms are rather rare. Just as in the BNC both in teacher and in student talk *do*, when including as well as when excluding the base form occurrences, is the most common verb form followed by the past tense form. An interesting fact worth noting is that the infinitive which has a 12.8 and 17 per cent distribution in conversations and task-oriented speech is not at all found in teacher talk and only once in student talk. This is even more extraordinary if one considers that in teacher talk more verbs occur than in student talk. Therefore the infinitive should be more likely to be found in the former than in the latter.

As already mentioned, due to the high number of auxiliary verb uses in teacher talk it may be that in the question/ negation distribution questions are more common as the teacher is bound to ask many questions in classroom discourse. In student talk *do*-questions are expected to be rather uncommon. The next table shows their distribution and if present or past tense is used more frequently.

Grammatical Functions of DO

TEACHER				
	Question	Negation	Sum	% of verb from use
do	9	6	15	45.5
does	1	1	2	6.1
Present:	10	7	17	51.5
did	14	2	16	48.5
Past:	14	2	16	48.5
Overall:	24	9	33	
% of overall distribution of to do	72.7	27.3		

STUDENT				
	Question	Negation	Sum	% of verb from use
do	2	3	5	50
does	1	1	1	10
Present:	3	3	6	60
did	1	3	4	40
Past:	1	3	4	40
Overall:	4	6	10	
% of overall distribution of to do	40	60		

Table 37. Grammatical Functions of DO in HIST 2

As suggested, in teacher talk questions occur more often than negations. The difference is very high with a 72.7 per cent share of question use and 27.3 per cent share of negation use. Also the suggestion made for the distribution in student talk is verified. Questions (40 %) are less frequent than negations (60 %). There is a clear divergence in the count of questions and negations. While in teacher talk present and past tense occur equally often in student talk the past tense is predominant.

HAVE

The last primary to be analyzed is *have*. This verb is used to express a variety of grammatical functions. These are perfect aspect, combination of aspect and passive voice, questions and negations in present perfect or present tense and its use as a substitution for the modal auxiliary *must*. Table 38 shows the main and auxiliary verb distribution, table 39 its verb form distribution and table 40 illustrates for which grammatical functions it is used.

Auxiliary and Main Verb Share of HAVE

TEACHER		
	Auxiliary Verb	Main Verb
Absolute frequency	22	44
Relative frequency	33.3	66.7
STUDENT		
Absolute frequency	18	15
Relative frequency	54.5	45.5
Classroom overall %	40.4	59.6

Table 38. Auxiliary and Main Verb Share of HAVE in HIST2

Firstly, in teacher talk more uses of the primary verb were counted, this is unsurprising, as the general number of verbs is also higher than in student talk. The relative numbers give more reliable information on their occurrence in this respect one can see that the share of auxiliary verbs is higher in teacher talk with 66.7 percent. Surprisingly, in student talk the main verb usage is more common than the auxiliary verb use with 54.5 per cent. In general *have* is more frequent as a main verb which is shown by the overall main verb distribution of 59.6 per cent. As the main verb use is so high, the next point worth discussing is the verb form share. Table 39 illustrates the distribution of verb forms as well as the BNC's findings. The primary verb *have* can occur in seven different forms: base form, -s form, present tense, past tense, -en and -ing participle and infinitive. The BNC summarizes the base form and present tense forms in the verb form *have*. This is not applied in my verb form distribution list for the classroom data.

Verb Form Distribution of HAVE

TEACHER				
	have	has	had	base form
Absolute frequency	23	10	22	11
Relative frequency	34.8	15.2	33.3	16.7
STUDENT				
Absolute frequency	10	4	18	1
Relative frequency	30.3	12.1	54.5	3.0
BNC findings				
BNC % in conversations	44.5	17.9	13	
BNC % in task-oriented speech	46.4	15.9	14.6	

Table 39. Verb Form Distribution of HAVE in HIST 2

According to the BNC *have* is the most common verb form with nearly 50 per cent. The –s form has a share of 17.9 per cent in conversations and 15.9 per cent in task-oriented speech. The past tense form is slightly less frequent. Adding up the numbers of the present tense use of *have* and the base form in teacher talk *have* is by far the most common verb form. (Leech 2001:97-298) Conversely to the corpus, in teacher talk not the –s form but the past tense form is more frequent. The outcome in student talk is even more unusual as the past tense form is the most frequent verb form with over 50 per cent. Only the result of *has* corresponds to the findings of the corpus as it is the least common verb and also has a similar distribution.

The last aspect discussed is the share of grammatical functions in need of the primary verb. In the table below one can see which functions are applied and also which tense dominates in the data samples. Taking all the information into consideration it can be said that teachers use a higher variety of grammatical functions which need a primary verb for their formation.

Grammatical Functions of HAVE

TEACHER							
	Perfect Aspect	Voice	Substitution Verb	Quest.	Neg.	Sum	% of verb form use
have	1	2			4	7	31.8
has				2	1	3	13.6
Present	1	2		2	5	10	45.5
had			12			12	54.5
Past			12			12	54.5
Overall	1	2	12	2	5	22	
% of overall have distribution	4.5	9.1	54.5	9.1	22.7		
STUDENT							
have			2			2	11.1
has			3			3	16.7
Present			5			5	27.8
had		2			11	13	72.7
Past		2			11	13	72.2
Overall		2	5		11	18	
% of overall have distribution		11.1	27.8		61.1		

Table 40. Grammatical Functions of HAVE in HIST 2

First of all in teacher talk five different grammatical functions utilizing *have* can be found. The most common use is using the primary as a substitution verb in the past tense. As such it has a distribution of 54.5 per cent. Even though present tense is less frequently used it is applied for a greater variety of grammatical functions. The most common are negations with present reference. Students only use it to express passive voice, as a substitution or in negations. Negative sentences are most frequent with 61.1 per cent. Also for students the past tense is used more often than present tense.

4.3.3 Distribution of Modal Verbs

The least common verb group, in both the LSWE corpus and HIST 2, is the modal verb class (Biber, Johansson, Leech, Conrad & Finegan 1999:359-360). Still its usage is described as it is necessary to gain information about the overall picture of verb usages in history lessons. Table 41 shows the present and past tense distribution and the list of modal verbs used including their number of occurrence.

List of Modal Verbs and their Distribution

TEACHER					
Present	Past		Number of distribution	% verb of distribution	
can	16	could	40	45	56.3
need	3	needed	1	4	5
will	4	would	9	13	16.3
may	1			1	1.25
must	6			6	7.5
		should	1	1	1.25
Sum	30		50		
% of tense distribution (teacher)	37.5		62.5		
STUDENT					
can	5	could	14	19	76
need	1	needed		1	4
will	1	would	3	4	16
must	1			1	4
Sum	8		17	25	
% of overall tense distribution (student)	32		68		
% of overall HIST 2 distribution	59.0		40.9		

Table 41. List of Modal Verbs and their Distribution in HIST 2

The variety of different modal verbs is higher in teacher talk than in student talk. Teachers use six different modal verbs, students only four. The most repeatedly occurring modal verb in teacher and student talk is *can* in both present and past tense. It has a share of 56.3 per cent in the former and 76 in the latter. Another similarity is that in both data samples the past tense is more frequent than the present with more than 60 per cent.

4.3.4 Lexical Verb Distribution

The most frequent verb class according to the BNC is the lexical verb class. (Leech 2001: 282) This is also true for the data of HIST 2. Lexical verbs can occur in seven different forms, their distribution is illustrated in table 42. The author of the BNC summarizes the base form and present tense form occurrences into the base form group (Leech 2001.282). This is, however, not done in the description of the classroom data.

		<u>Lexical Verb Form Distribution</u>						
		TEACHER						
		base form	-s form	present form	past form	-ing participle	-en participle	to infinitive
Number of occurrence		73	23	82	52	18	41	38
% of occurrence		22.8	7.2	25.6	16	5.6	13	11.9
		STUDENT						
Number of occurrence		20	4	29	26	13	18	16
% of occurrence		15.3	6.9	22.1	20	9.9	6.7	12.2
		BNC Findings						
% of conversation distribution		27.7	6.0		14.3	10.8	12.9	28.2
% of task-oriented speech		23.3	5.1		9.6	12.4	17.2	32.3

Table 42. Lexical Verb Form Distribution in HIST 2

In teacher talk the present tense form is the most common verb form with 25.6 per cent, when adding the base form one gets a share of 48.4 per cent which is two times

as much as the BNC's share. Another difference to the BNC is that not the infinitive but the past tense form is next in line. Also the results of student talk differ to the BNC corpus. The base and present tense forms are the most common with a general number of 37.4 per cent. After that follows the past tense form with 20 per cent. The main reason for these dissimilarities is probably the fact that history lessons make a higher use of past tense forms necessary. Because of this the order of verb forms is influenced and the results differ from the ones of the BNC.

4.3.5 Semantic Domains and Lexical Verb Distribution in HIST1 and HIST2

The English language knows seven semantic domains which have been explained in detail in Chapter 1. Lexical verbs can be categorized into activity, mental, communication, occurrence, existence, causative and aspectual verbs. Two more issues remain to be discussed. Firstly, the semantic domain share of lexical words of both history lessons. Secondly, the most frequently used verbs and their correspondence to the LSWE (Biber, Johansson, Leech, Conrad & Finegan 1999 365-366) and BNC (Leech 2001: 282). Table 43 shows the distribution of semantic domains in both transcripts and their total number. The total number's percentage is then compared to the results of the LSWE.

Semantic Domain Distribution in History Lessons

	HIST 1	HIST 2	Total number	% of distribution	% of LSWE
Activity	49	45	94	48	43.8
Mental	18	16	34	17.3	27.9
Communication	17	11	28	14.3	11.9
Occurrence	12	4	16	8.2	1.5
Existence	9	5	14	7.1	11.9
Causative	3	4	7	3.6	1
Aspectual	1	2	3	1.5	1

Table 43. Semantic Domain Distribution in History Lessons Compared to the LSWE

Biber and his co-authors state in their Longman corpus that the most frequent category is the one of activity verbs with 43.8 per cent. Also very common are mental, communication and existence verbs. The data corresponds to this order. (Biber, Johansson, Leech, Conrad & Finegan 1999: 365-366) Activity verbs, like *come*, *get*, *got*, *give* and *make* are the most common ones with a share of 48 per cent. Mental verbs have a share of 17.3, some examples found in the data were *find*, *know*,

see and *think*. Communication verbs, such as *ask*, *say*, *tell* and *talk*, have a share of 14.3 per cent in the data of history lessons. Occurrence verbs have, in contrast to the LSWE results, a quite high share, the following verbs were found in the data: *happen*, *become*, *die* and *grow*. Causative verbs, as *let* and *allow*, and aspectual verbs as *finish* and *bring* are used very seldom in both the data and the LSWE. All in all the results of the data analysis equal the ones of the corpus.

The last point aims at an analysis of the most frequent verb forms. The most common verbs of the HIST data are given as well as the lists of the LSWE (Biber, Johansson, Leech, Conrad & Finegan 1999: 373-377) and BNC (Leech 2001: 296-297). The bold verbs in the data list also occur in the LSWE and the underlined ones in the BNC.

Most frequently used Lexical Verbs

	HIST		LSWE	BNC
1.	get	28	get	say
2.	go	17	go	get
3.	<u>know</u>	15	say	make
4.	meet	15	know	go
5.	mean	14	think	see
6.	choose	13	see	know
7.	happen	12	want	take
8.	write	11	come	think
9.	work	11	mean	come
10.	<u>say</u>	10	take	give
11.	leave	9	make	look
12.	elect	9	give	use

Table 44. Most frequently used Verbs in History Lessons, LSWE and BNC

Only four verbs of the LSWE list and five verbs of the BNC can be found in the top twelve verb occurrences, these are *get*, *go*, *know* and *mean*. However, not all verbs can be spotted in every data set, which is the reason for not being in the most frequent verb list. The verbs which are included are verbs which were important for the lessons and were used to explain happenings in the past. As the lesson was about life and social hierarchy and the legal systems in antiquity verbs such as *choose*, *elect* and *work* are not surprising.

4.3.6 Summary of Findings

The general insights are summarized in this section. This means that the major differences but also similarities between the two analyzed history lessons are recapitulated. Additionally, the overall relative distribution of tenses is given as well as the grammatical functions of aspect, mood and passive voice. The corpus data is taken from Biber (Biber, Johansson, Leech, Conrad & Finegan 1999: 359-481) and Leech (2001: 282-297)

The share of primary, modal and lexical verbs is compared as a first point. In HIST 1 teacher talk primary verbs are more common than lexical verbs but in all the other sets of data the suggested order of the LSWE can be found. Still the enormous number of primary verbs is extraordinary as, according to the Longman corpus, primary verbs are just slightly more often used than modal verbs and their numbers put together do not reach the total number of lexical verb occurrence. In the primary verb distribution of auxiliary and main verb function, however, all the history data sets result in a higher use of the main verb than the auxiliary verb function. A further similarity is that *be* is always the most common verb form followed by *have* and then *do*. These findings equal the ones of the BNC. While all of the data sets show a predominant use of the main verb function in the single primary verbs in teacher talk *do* is more often used as an auxiliary verb than as a main verb. In terms of tense the past tense is predominant in all the primaries. Quite interesting is also that in terms of verb form distribution the infinitive is very uncommon, although the BNC claims that the infinitive is the, or at least, one of the most frequently used verb forms. Also the past tense forms are far more often used than they statistically should be according to the BNC finding on conversations and task-oriented speech. A further fact is that teachers use a higher variety of grammatical functions than students. Now I will turn to the modal verb usage in history lessons. Overall the most frequent modal in both sets of data is *can* and its past tense form. Overall, also *would* occurs relatively often. While in HIST 1 the past tense distribution is higher in HIST 2 the present tense is more common. Finally, the lexical verb form distribution is quite different from the one found in the LSWE. Also here the infinitive is not as common as it should be. All in all the past tense is predominant which is also quite unusual, when taking a look at the corpus. The table below shows a detailed description of

grammatical structure usage and present and past tense distribution once again.

Distribution of Tense, Aspect, Mood and Passive Voice

TEACHER									
	Tense		Aspect		Mood		Voice		
Number of occurrence	429		9		92		27		
% of occurrence	77.1		1.6		16.5		4.8		
	<i>Present</i>	<i>Past</i>	<i>Present</i>	<i>Past</i>	<i>Present</i>	<i>Past</i>	<i>Present</i>	<i>Past</i>	
Times of occurrence	262	167	9	0	24	68	10	17	
% of occurrence	61.1	38.9	1.6	0	26.1	73.9	37.0	63.0	
Overall number	Present				Past				
% of overall occurrence	54.8				45.2				
STUDENT									
	Tense		Aspect		Mood		Voice		
Number of occurrence	168		7		28		16		
% of occurrence	79.2		3.3		13.2		7.5		
	<i>Present</i>	<i>Past</i>	<i>Present</i>	<i>Past</i>	<i>Present</i>	<i>Past</i>	<i>Present</i>	<i>Past</i>	
Times of occurrence	93	75	7	0	11	17	1	15	
% of occurrence	66.4	44.6	100	0	39.3	60.7	6.3	93.7	
Overall number	Present				Past				
% of overall occurrence	51.1				48.9				

Table 45. Distribution of Tense, Aspect, Mood and Passive Voice in HIST 2

According to Biber's findings in spoken English tense is the most common grammatical category followed by mood, aspect and passive voice in this order. Both teacher and student talk have different distributions. Passive voice is not the least common grammatical structure but aspect. Tense, however, has the highest share with over 70 per cent. Teacher talk shows a higher share of present tense in tense and aspect, which leads to an overall dominance of this tense, however, the difference is not clear as present tense has a sharing of 54.8 per cent and past tense 45.2 per cent. In student talk the difference is even smaller with 51.1 per cent for

present tense and 48.9 per cent for past tense. Overall this is quite extraordinary, as past tense is not as common in conversations and even less in task-oriented speech. It was assumed that in history lessons the past tense would be used more often than in regular spoken English, due to the fact that information on past events is conveyed. Although the past tense is not the predominant form it still shows a considerably higher occurrence than in conversations and task-oriented speech. Speech in a classroom is a kind of task-oriented speech. The reason why the results in classroom discourse and task-oriented speech are not the same is that that in the latter a vast range of task-oriented speech acts are included.

4.4 MARK 1

MARK 1 is a transcript of a marketing lesson, recorded on January 27, 2003 in the third grade of a “Handelsakdamie”. The students there are in 11th grade and of an average age of 16 to 17. The lesson is about how to promote a product and which qualities a product has to have to be successful in different markets. The data set MARK 1 consists of an overall number of 7,288 words of which 1,591 words are German and 5,697 English. The German words of course are excluded from this analysis. Teacher talk consists then of 3,908 words of which 787 are verbs. This is a share of 20.1 per cent. Student talk consists of 1789 words 269 of which are verbs, which is 15.1 per cent.

4.4.1 Word Class Distribution

The Longman corpus on written and spoken English says that the lexical verb form is the most verb common class of all. (Biber, Johansson, Leech, Conrad & Finegan 1999: 359-260) Table 46 shows the distribution in teacher and student talk of MARK 1.

<u>Verb Class Distribution</u>			
TEACHER			
	Primary Verb	Modal Auxiliary	Lexical Verb
Absolute number of occurrences	415	86	286
% of distribution	52.7	10.1	36.3
STUDENT			
Absolute number of occurrences	106	25	138
% of distribution	39.4	9.3	51.3
% of overall distribution	49.3	10.5	40.2
LSWE Findings			
% of LSWE findings	19	11.4	69.6

Table 46. Verb Class Distribution in MARK 1

The marketing teacher in MARK 1 uses an astounding number of 415 primary verbs. This is an enormous share of 52.7 per cent. Lexical verbs have a share of 36.3 per cent, and modal verbs of 10.1 per cent. This equals the numbers suggested by

the LSWE corpus. Student talk, on the other hand, shows the same frequency ranking as the LSWE. Lexical verbs are with 51.3 per cent the most used verbs followed by a 39.4 per cent share of primary verbs and only 9.3 per cent of modal verbs. However, when adding the numbers of both data sets primary verbs dominate with 49.3 per cent. In the LSWE primary verbs only have a share of 19 per cent. This means that the overall results do not correspond to the findings of the LSWE corpus because primary verbs occur more often than the other verb classes. This may be due to the high usage of questions and negations in classroom discourse. Another reason might be a high occurrence of a variety of grammatical functions which are formed by using a primary verb.

4.4.2 Primary Verb Distribution

As the number of primary verbs is so extraordinarily high special interest is taken in its analysis. The first issue is to find out if the primary verbs occur more often in their main verb or in their auxiliary verb function. Furthermore, the distribution of the single primary verbs is of interest which is also compared to the findings of the LSWE (Biber, Johansson, Leech, Conrad & Finegan 1999: 359-360). After this every individual primary verb is analyzed in detail. Table 47 shows the main and auxiliary verb distribution and table 48 the distribution of the single primary verbs in comparison to the corpus.

Auxiliary and Main Verb Distribution of Primary Verbs

TEACHER		
	Auxiliary Verb	Main Verb
Absolute frequency	121	294
Relative frequency	29.2	70.8
STUDENT		
Absolute frequency	18	88
Relative frequency	17.0	83.0
Classroom overall %	27.0	73.0

Table 47. Auxiliary and Main Verb Distribution of Primary Verbs in MARK 1

As can be seen in the table the main verb usage dominates strongly in both teacher and student talk. The absolute numbers are difficult to compare as also the overall

verb count differs greatly. The relative numbers show that the main verb has a share of 70.8 per cent in teacher talk and 83.0 per cent in student talk. This leads to an overall frequency of 73 per cent of main verbs and only 27 per cent auxiliary verb usage. The distribution of the individual primary verbs is described below.

Distribution of BE, DO and HAVE

TEACHER			
	BE	DO	HAVE
Absolute Number of occurrences	296	46	73
% of Primary Verb Class	71.3	11.1	17.6
% of overall distribution (teacher)	37.6	5.8	9.3
STUDENT			
Absolute Number of occurrences	69	15	22
% of Primary Verb Class	65.1	14.2	20.7
% of overall distribution (student)	25.7	5.6	8.2
% of overall distribution in MARK 1	51.6	11.9	18.5
BNC Findings			
% of BNC findings in conversations	58.4	20.5	21.1
% of BNC findings in task-oriented speech	63.3	15.1	21.6

Table 48. Distribution of BE, Do and HAVE in MARK 1

The primary verb *be* has to be used in the formation of the greatest variety of grammatical functions. *Have* is used to form numerous grammatical structures and *do* is only necessary in its *dummy do* usage in questions and negations or to add further stress on a statement. According to the BNC the primary *be* is the most frequently used verb in both conversations and task-oriented speech. While in conversations *have* is more predominant than *do*, in task-oriented speech it is the other way round. Both sample sets show the same weighting as task-oriented speech in the BNC. Especially the relative distribution in student talk is very similar to the corpus findings. (Leech 2001: 282) In teacher talk the primary verb *be* shows an even higher share than in student talk. The table above also shows the frequency of the single primary verbs in the light of all verbs put together. All of them are used more often than any lexical verb. The primary *be* is the most used verb.

BE

Be is the most common primary verb, it occurs 296 times in teacher talk and 69 times in student talk. Table 49 shows its auxiliary and main verb distribution. Later on its verb form share is illustrated. Additionally, it is discussed for which grammatical functions it is used.

Auxiliary and Main Verb Share of BE

TEACHER		
	Auxiliary Verb	Main Verb
Absolute frequency	63	233
Relative frequency	21.3	78.7

STUDENT		
Absolute frequency	3	66
Relative frequency	4.3	95.7

Classroom overall %	18.1	81.9
----------------------------	-------------	-------------

Table 49. Auxiliary and Main Verb Share of BE in MARK 1

In both data samples the main verb use of the primary is predominant. In teacher talk the main verb form use of *be* has a share of 78.7 per cent and in student talk this number is even higher with 95.7 per cent. Table 50 shows the distribution of verb forms in contrast to the BNC findings of conversations and task-oriented speech (Leech 2001: 297).

Verb Form Distribution of BE

TEACHER							
	to be	be	is	are	was	were	been
Number of occurrence	36		168	68	20	2	2
Overall % of <i>to be</i> forms	12.2		56.8	23	6.8	0.7	0.7

STUDENT							
Number of occurrence		6	34	18	8	3	
Overall % of <i>to be</i> forms		6.7	49.3	26.1	11.6	4.3	

% of overall distribution	9.9	1.6	55.3	23.6	7.7	0.7	0.7
----------------------------------	------------	------------	-------------	-------------	------------	------------	------------

BNC Findings							
% of BNC conversation	7.2	1.1	50.5	15	13.7	7.8	2.7
% of task-oriented speech	12.4	0.1	41.1	17.5	13.8	5.5	3.8

Table 50. Verb Form Distribution of BE in MARK 1

It is extraordinary that the results of teacher talk and student talk show the nearly the same order as both the BNC corpus findings in conversations and task-oriented speech. The base forms –s form is the most common verb form with 56.9 per cent in teacher talk and 49.3 per cent in student talk. This number is followed by the second present tense form with 23 and 26.1 per cent, respectively. The only difference is that the past tense *was* is not as often used in the analyzed data as it is in the BNC (Leech 2001:297). Table 51 shows for which grammatical functions the primary verb is used.

Grammatical Functions of BE

TEACHER						
	Prog. Aspect	Voice	going to	Question	Sum	% of verb form use
is	2	6	4	26	38	60.3
are	8		16		24	38.1
Present	10	6	29	26	62	98.4
was		1			1	1.6
were					0	
Past:	0	1	0	0	1	1.6
Overall	10	7	29	26	63	
% of overall usage	15.9	11.1	46.0	41.3		

STUDENT						
is				1		33.3
are			1			33.3
Present			1	1	2	66.7
was						
were		1			1	33.3
Past		1			1	33.3
Overall		1	1	1	3	
% of overall usage	overall	33.3	33.3	33.3		

Table 51. Grammatical Functions of BE in MARK 1

The teacher uses the primary *be* to express aspect, passive voice, *going to* future and for questions. The *going to* future is the most common grammatical structure in MARK 1 with 46 per cent. Also a lot of questions are asked as the percentage of 41.3 shows. Progressive aspect and passive voice do not occur as frequently. All in all present tense is predominant with 98.4 per cent. Students use the modal function of the primary only three times. Two of these uses are present tense. This corresponds to the overall predominance of the present tense.

DO

The primary *do* is the least frequently used primary verb. In contrast to the primary *be* which is more common as a main verb, for *do* the auxiliary verb usage is more frequent as can be seen in table 52, which shows the share of auxiliary and main verb usage in the two sample sets and the overall distribution in MARK 1.

Auxiliary and Main Verb Share of DO

TEACHER		
	Auxiliary Verb	Main Verb
Absolute Number of occurrences	33	13
% of <i>do</i> verb forms	71.7	28.3

STUDENT		
Absolute Number of occurrences	11	4
% of <i>do</i> verb forms	73.3	26.7
% of overall distribution	72.1	27.9

Table 52. Auxiliary and Main Verb Share of DO in MARK 1

In teacher talk the share of auxiliary verb usage is 71.7 per cent and 73.3 per cent in student talk. The main verb use of the primary is not even one third of the overall number. Before turning to the grammatical functions the verb form distribution is illustrated in table 53. Also the findings of the BNC (Leech 2001: 297-298) are included in the table in order to compare them to the results of the two datasets.

Verb Form Distribution of DO

TEACHER				
	do	does	did	to do
Absolute Number of occurrences	44		2	1
% of overall <i>do</i> forms	95.7		4.3	2.2

STUDENT				
Absolute Number of occurrences	13	1		1
% of overall <i>do</i> forms	86.6	6.7		6.7
% of overall distribution	78.1	1.4	2.7	1.4

BNC Findings				
% in conversations	45.6	8.7	21.8	12.8
% in task-oriented speech	38.8	10.1	21.1	17

Table 53. Verb Form Distribution of DO in MARK 1

Only three different verb forms occur in both teacher and student talk. In the former these are the base and present tense form, the past tense form and the infinitive. The most common verb form is *do* with a share of 95.7 per cent. The past tense and the infinitive are rare with less than five percent each. The very low number of occurrence of the other verb forms is surprising as the numbers for these are comparatively higher in the BNC. In student talk the two present tense forms occur as well as the infinitive. Just as in teacher talk also here the base/present form is predominant with 86.6 per cent. All in all the very rare use of other verb forms, apart from *do*, is very interesting.

The final point of discussion is the auxiliary verb use of the dummy *do*. It can only be used to form questions and negations. Table 54 shows their share as well as which tense is used more frequently. From the verb form count it can already be said that the present tense will be the predominant tense, it may even be that the past tense does not occur at all.

Grammatical Functions of DO

TEACHER				
	Question	Negation	Sum	% of verb from use
do	23	8	31	93.9
Present:	23	8	31	93.9
did	1	1	2	6.1
Past:	1	1	2	6.1
Overall:	24	9	33	
% of overall distribution of <i>to do</i>	72.7	27.3		
STUDENT				
do		10	10	90.9
does	1		1	9.1
Present:	1	10	11	100
Overall:	1	10	11	
% of overall distribution of <i>to do</i>	9.1	90.9		

Table 54. Grammatical Functions of DO in MARK 1

For teacher talk the predominant use of questions is not surprising as teachers have to ask many questions in classroom discourse. The share of questions is 72.7 per cent. Only *do* questions are asked. As claimed above the main tense is the present tense with 90 per cent. In student talk only present tense questions and negations can be found. In contrast to teachers, students only use the primary *do* in

order to negate sentences as can be seen by the share of 90.9 per cent.

HAVE

The last primary verb is *have*, which occurs 73 times in teacher talk and 22 times in student talk. Table 55 shows if it is more common as an auxiliary or main verb. It not only illustrates the share of the individual sets of data but also overall share in the classroom data.

Auxiliary and Main Verb Share of HAVE

TEACHER		
	Auxiliary Verb	Main Verb
Absolute frequency	25	48
Relative frequency	34.3	65.7
STUDENT		
Absolute frequency	4	18
Relative frequency	18.2	81.8
Classroom overall %	30.5	69.4

Table 55. Auxiliary and Main Verb Share of HAVE in MARK 1

In both sets of data the primary *have* is more often used in its main verb function, with 65.7 per cent in teacher talk and 81.1 per cent in student talk, than in its auxiliary verb function. Although they have the same result the difference between auxiliary and main verb function is more distinct in student talk. These verbs are now sorted into verb form groups in table 56, which shows the verb form distribution in comparison to the results of the BNC (Leech 2001: 298).

Verb Form Distribution of HAVE

TEACHER				
	have	has	had	base form
Absolute frequency	55	3	4	11
Relative frequency	75.3	4.1	5.5	2.7
STUDENTS				
Absolute frequency	13	6		3
Relative frequency	59.1	27.3		13.6
Classroom Overall %	71.6	9.5	4.2	11.6

BNC Findings			
BNC % in conversations	44.5	17.9	13
BNC % in task-oriented speech	46.4	15.9	14.6

Table 56. Verb Form Distribution of HAVE in MARK 1

Just as in the primary verb *do* there is an enormous difference between the findings in MARK 1 and the BNC (Leech 2001: 298) when it comes to verb form distribution. In teacher talk four different verb forms appear, these are *have*, *has* and *had*. *Have* can be either the base form or a tensed verb. The BNC summarizes these in one. The base form, tensed or not, is the most frequently used form with 73.2 per cent, if including the base form as well. The –s form and the past tense form are in comparison very rare. In student talk only three verb forms can be found, *have* as a base form and as present tense, and the –s form. All in all, both the data and the BNC show a high use of the verb form *have*. The absence of the –en form and the low number of the –s form and the past tense is surprising.

Now a closer look is taken at the grammatical functions. *Have* is used to form perfect aspect, passive voice and as a substitution verb in MARK 1. Table 57 shows the share of grammatical functions in teacher and student talk, respectively. Furthermore, it gives account of the tenses used.

Grammatical Functions of HAVE

TEACHER					
	Perfect Aspect	Substitution Verb	Voice	Sum	% of verb form use
have	2	20	1	23	92
has			1	1	4
Present	2	20	2	24	96
had		1		1	4
Past		1		1	4
Overall	2	21	2	25	
% of overall have distribution	8	84	8		
STUDENT					
have		2		2	50
has	2			2	50
Present	2	2		4	
Overall	2	2		4	
% of overall have distribution	50	50			

Table 57. Grammatical Functions of HAVE in MARK 1

The teacher in MARK 1 uses *have* to form perfect aspect and passive voice. It is also used to substitute the modal verb *must*. The last of these purposes is the most common one with a share of 84 per cent. The reason might be that in marketing lessons rules are presented and the teacher therefore uses *have to* in order to express what has to be done in marketing. In terms of tense use, the present tense is predominant. Past tense is nearly never used.

Summing up, the primaries *be* and *have* are more often used in their main verb function. Only for *do* the auxiliary verb function is predominant. Also all the primary verbs are predominant in their present tense form. Past tense occurs only very rarely. Additionally, some *going to* future structures were found. This can be deduced from the taught subject, as marketing is a subject which teaches rules and facts and how a product has to be presented in order to sell it successfully in the future.

4.4.3 Distribution of Modal Verb

The least common verb class in this set of data and according to the LSWE is the modal verb class (Biber, Johansson, Leech, Conrad & Finegan 1999: 359-360). Which modal verbs are used and how often can be observed in the table 58. It also shows the share of present and past tense in the modal verb group itself and also in the overall number of present and past tense uses.

Modal Verb Distribution

TEACHER					
Present		Past		Word count	% of modal verbs
can	20	could	14	34	39.5
need	3	needed		3	3.5
will	7	would	22	29	33.7
may	6			6	7
		should	13	13	15.1
must	1			1	1.2
Sum	37		49	86	
% of tense distribution	43		57		

STUDENT					
can	12	could	1	13	52
need	1	needed		1	4
will	6	would		6	24
		should	3	3	12
must	2			2	8
Sum	21		4	25	
% of tense distribution	84		16		

Table 58. Modal Verb Distribution in MARK 1

One can find an overall number of six different modal verbs in MARK 1. In teacher talk *can*, *need*, *will*, in present and past tense, *may*, *must* and *should* are used. The most common modals are *can* with a share of 39.5 per cent and *will* with a 33.7 per cent. With 15.1 per cent also *should* is relatively frequent. Turning to tense past tense is more common with 57 per cent. In student talk only five modal verbs appear, the most common is *can* with a percentage of 52. Also *will* is relatively frequent with 24 per cent. In contrast to teacher talk, in student talk the present tense dominates with 84 per cent.

4.4.4 Lexical Verb Distribution

As we saw in table 46 in MARK 1 the lexical verb class is not the most common verb class as it should be according to the LSWE (Biber, Johansson, Leech, Conrad & Finegan 1999: 359-360). This section takes a closer look at the verb form distribution and aims to find out if there are many divergences. For reasons of comparison also the BNC's findings (Leech 2001: 282) are included in the table 59.

Verb Form Distribution of Lexical Verbs

TEACHER							
	base form	-s form	present form	past form	-ing participle	-en participle	to inf.
Absolute frequency	94	22	76	24	11	12	33
Relative frequency	32.9	7.7	26.6	8.4	3.8	4.2	11.5
STUDENT							
Absolute frequency	23	17	63	14	2	6	12
Relative frequency	16.7	12	45.7	10.1	1.4	4.4	8.7

BNC Findings						
BNC % in conversation distribution	27.7	6.0	14.3	10.8	12.9	28.2
BNC % in task-oriented speech	23.3	5.1	9.6	12.4	17.2	32.3

Table 59. Verb Form Distribution of Lexical Verbs in MARK 1

In teacher talk the most common verb form is the present tense and if adding the base form as it is done in the BNC the number is even higher with 59.5 per cent. This number opposes 23.3 per cent in BNC task-oriented speech. Just as in the BNC also in teacher talk the infinitive is the second frequent verb form with 11.5 per cent. In contrast to the BNC the -ing and -en participles are very uncommon in this set of data. (Leech 2001: 282) Student talk shows even more astonishing results. Just as in teacher talk the present and the base form have the highest frequency. More than half of all verbs occur in the base form. 62.4 per cent is an enormous number and, it is especially interesting, that it is three times higher than in the BNC. The -s form is the next verb form with a 12 per cent, this is surprising as according to the BNC the -s form is the least common verb form of all. Just as in teacher talk the -ing and -en participles are very rare. All in all the present and the base form are the most frequent verb forms in MARK 1, the -s form is more often used in the data, while all the other verb forms are far less common than the BNC suggests. Summing up the verb form distribution in MARK 1 does not correspond to the one of the BNC.

4.4.5 Summary of Findings

This section is devoted to a summary of the most important facts. The distribution of tense, aspect, mood and voice is illustrated as well as the present and past tense frequency.

Distribution of Tense, Aspect, Mood and Passive Voice

TEACHER									
	Tense		Aspect		Mood		Voice		
Number of occurrence	475		12		106		9		
% of occurrence	79.1		2		17.7		1.5		
	<i>Present</i>	<i>Past</i>	<i>Present</i>	<i>Past</i>	<i>Present</i>	<i>Past</i>	<i>Present</i>	<i>Past</i>	
Times of occurrence	423	52	12	0	57	49	8	1	
% of occurrence	89.1	10.9	100	0	53.7	46.3	88.9	11.1	
Overall number	Present		500		Past		102		
% of overall occurrence			83.1				16.9		
STUDENT									
	Tense		Aspect		Mood		Voice		
Number of occurrence	189		2		27		1		
% of occurrence	86.3		0.9		12.3		0.5		
	<i>Present</i>	<i>Past</i>	<i>Present</i>	<i>Past</i>	<i>Present</i>	<i>Past</i>	<i>Present</i>	<i>Past</i>	
Times of occurrence	164	25	2	0	23	4	0	1	
% of occurrence	86.1	13.9	100	0	85.2	14.8	0	100	
Overall number	Present		189		Past		30		
% of overall occurrence			86.3				13.7		

Table 60. Distribution of Tense, Aspect, Mood and Passive Voice in MARK 1

In teacher and student talk tense is the most recurrent grammatical form, with 79.1 and 86.3 per cent each. Mood has also a mentionable share of 17.7 per cent. In every single section the present tense is the most frequent tense. It should also be mentioned that the present tense always occurs over 50 per cent.

4.5 MARK 2

The transcript MARK 2 was recorded on February 3, 2003. The teacher is the same as in MARK 1. The students in this lesson present their findings on product sales. Also how to do a survey on the product's market share is discussed.

The lesson consists of 7,589 words. German was also relatively often spoken as 1,168 of all words were German and 6,421 English. Teacher talk consists of 4,306 words of which 865 are verbs. Student talk consists of only 2,115 verbs of which 313 are verbs. In the former 20.1 per cent of all words belong to the verb class in the latter 14.8 per cent.

4.5.1 Word Class Distribution

The share of word class is illustrated in table 61. Furthermore, results of the LSWE (Biber, Johansson, Leech, Conrad & Finegan 1999 359- 360) and BNC (Leech 2001: 282-298) are given, which are compared to the findings. Not only the absolute but also the relative distribution is given.

<u>Verb Class Distribution</u>			
TEACHER			
	Primary Verb	Modal Auxiliary	Lexical Verb
Absolute frequency	402	97	370
Relative frequency	46.5	10.7	42.8
STUDENT			
Absolute frequency	116	29	145
Relative frequency	37.1	16.6	46.3
Classroom overall %	49.3	10.5	40.2
LSWE Findings			
LSWE %	19	11.4	69.6

Table 61. Verb Class Distribution in MARK 2

Biber found out that in spoken English the lexical verb group is the most common verb class and that the primary and modal auxiliary verb group are less frequently used. This is not true for MARK 2, because of various facts explained in the following. Firstly, in teacher talk primary verbs are with 46.5 percent more common than lexical

verbs with a share of 42.8 per cent. Secondly, in student talk it is true that lexical verbs are more common with 46.3 per cent, but the primary verb class follows with 37.1 per cent which is far more than in the LSWE. In the overall distribution primary verbs are more often found than lexical verbs. The modal auxiliary class is the least frequent verb group in both the data and the LSWE.

4.5.2 Primary Verb Distribution

An overall share of 49.3 per cent of the primary verb class is extraordinary, especially in comparison to 19 per cent in the LSWE (Biber, Johansson, Leech, Conrad & Finegan 1999: 359-360). Now it will be illustrated if primary verbs are more common in their main or auxiliary verb usage. Also the share of the single primary verbs is given.

Auxiliary and Main Verb Distribution

TEACHER		
	Auxiliary Verb	Main Verb
Absolute frequency	97	305
Relative frequency	24.1	75.9
STUDENT		
Absolute frequency	29	87
Relative frequency	25	75
Classroom overall %	24.3	74.7

Table 62. Auxiliary and Main Verb Distribution in MARK 2

The numbers in student and teacher talk are very similar. The main verb use is with 75 per cent more frequent than the auxiliary verb function. This implies that the variety of grammatical function is rather low, especially in student talk where the primary only is used 29 times as an auxiliary verb. In teacher talk more grammatical functions can be expected as a total of 97 auxiliary verbs were counted.

Now the frequency of the single primaries is illustrated, as it is interesting to see which primary is most often used. Table 63 not only gives information on the individual share but also the frequency of the verbs in the total verb count. The three primaries are the most frequently used verbs. I will now give information on the

frequency ranking of primary verbs. The LSWE suggests that *be* is the most common primary verb followed by *have*. *Do* is the least common primary.

Distribution of BE, DO and HAVE

TEACHER			
	BE	DO	HAVE
Absolute frequency	247	59	69
Relative frequency	68.2	14.7	17.1
STUDENT			
Absolute frequency	81	20	15
Relative frequency	69.8	17.2	13
Classroom overall %	63.3	15.3	16.2
LSWE Findings			
BNC % in findings in conversations	58.4	20.5	21.1
BNC % in findings in task-oriented speech	63.3	15.1	21.6

Table 63. Distribution of BE, DO and HAVE in MARK 2

Just as in the LSWE the primary *be* is the most frequently used primary verb in the analyzed data. (Biber, Johansson, Leech, Conrad & Finegan 1999: 369-360) In both teacher and student talk it has a share of nearly 70 per cent. Differing results are found for *do* and *have*. While the order of the BNC is followed in teacher talk, students use the primary *do* more often than the primary verb *have*. In the individual representation of these verbs a detailed analysis can be found. Maybe a conclusion as to the different predominant usage can be drawn.

BE

In this section one finds information on the auxiliary and main verb use of the primary *be*. In addition to this its verb form distribution and occurrence of grammatical functions are given.

Auxiliary and Main Verb Distribution of BE

TEACHER		
	Auxiliary Verb	Main Verb
Absolute frequency	37	210
Relative frequency	15	85

STUDENT		
Absolute frequency	6	75
Relative frequency	7.4	92.6
Classroom overall %	13.1	86.9

Table 64. Auxiliary and Main Verb Distribution of BE in MARK 2

Be is more common as a primary verb than as an auxiliary verb. The main verb function has a share of 85 per cent in teacher talk and 92.7 per cent in student talk. Table 64 shows that the variety of grammatical functions should be higher in teacher talk than in student talk, as in teacher talk 37 different auxiliaries were found but only 6 in student talk. Table 65 illustrates which verb forms can be found and it also gives information about the tense use in MARK 2.

Verb Form Distribution of BE

TEACHER							
	to be	be	is	are	was	were	been
Absolute frequency		33	163	64	9	5	
Relative frequency		12.1	59.5	23.3	3.3	1.8	
STUDENT							
Absolute frequency	4	14	38	22	3	3	
Relative frequency	4.9	17.3	46.9	27.2	3.7	4.3	
Classroom overall %	1.3	14.5	63.2	27	5.5	2.5	
BNC Findings							
BNC % in conversation	1.1	0.2	50.5	15	13.7	7.8	2.8
BNC % in task-oriented speech	0.1	0.1	41.1	17.5	13.8	5.5	3.8

Table 65. Verb Form Distribution of BE in MARK 2

In teacher talk only five different forms of *be* appear. The most common ones are the present –s form with nearly sixty per cent and the plural present form with 23.3 per cent. The past tense forms are very rarely used. An interesting result is that the base form occurs very often, but according to the BNC the base form should be the rarest verb form. Students use six different verb forms. All those which occur in teachers talk plus the infinitive. The distribution is similar to the one in teacher talk. An interesting finding is that the infinitive which has a higher share than could be assumed from the low distribution in the BNC. Past tense forms are very seldom used. All in all the past tense *was* is in both sets very rare, but in the BNC it is one of the three most common verb forms. (Leech 2001: 297)

The next table shows the grammatical functions distribution:

<u>Grammatical functions of BE</u>					
TEACHER					
	Prog. Aspect	Passive Voice	going to	Sum	% of verb form use
is	1	5	5	11	29.7
are	4	9	6	19	51.3
Present	5	14	11	30	81.1
was	2	3		5	13.5
were		2		2	5.4
Past	2	5		7	18.9
Overall	7	19	11	37	
% of overall usage	18.9	51.3	29.7		
STUDENT					
			Quest.		
is			2	2	33.3
are	2	2		4	66.7
Present	2	2	2	6	
Overall	2	2	2	6	
% of overall usage	33.3	33.3	33.3		

Table 66. Grammatical Functions of BE in MARK 2

Progressive aspect, passive voice and *going to* future are the three grammatical functions found in teacher talk. The highest frequency is found in passive voice with 51.3 per cent. With 29.7 per cent the *going to* future is also relatively frequent. In this marketing lesson the present tense dominates with 81.1 per cent. In student talk only present tense is used to form grammatical functions. Students use *be* to form progressive aspect, passive voice and questions with the same share. Overall grammatical functions in student talk are not very significant as for each of the functions only two hits could be found. In the next section the primary *do* is described maybe the findings for student talk are more noteworthy.

DO

Do occurs 59 times in teacher talk and only 20 times in student talk. It is not only the least common primary verb, it also shows a completely different share in auxiliary and main verb use as can be seen in table 67.

Auxiliary and Main Verb Distribution of DO

TEACHER		
	Auxiliary Verb	Main Verb
Absolute frequency	48	11
Relative frequency	81.4	18.6
STUDENT		
Absolute frequency	17	3
Relative frequency	85	15
Classroom overall %	82.3	17.7

Table 67. Auxiliary and Main Verb Distribution of DO in MARK 2

While the primary *be* is more common in its main verb function, *do* is more frequently used in its auxiliary function in order to express questions and negations. In both data sets the auxiliary verb frequency is over eighty per cent. This means that both teachers and students ask questions and negate sentences. It can be assumed because of the role of both analyzed speech groups that in teacher talk more questions will be found than negations and that it may be the other way round in student talk.

From the following verb form description a first suggestion on present and past tense distribution can be made.

Verb Form Distribution of DO

TEACHER							
	do	does	did	done	doing	to do	base
Absolute frequency	32	7	13	3	1	2	2
Relative frequency	54.2	11.9	22.0	3.4	1.7	3.4	3.4
STUDENT							
Absolute frequency	16	2	1		1		
Relative frequency	80	10	5		5		
% of overall distribution	60.8	11.4	17.7	3.8	2.5	2.5	2.5
BNC Findings							
BNC% in conversations	45.6	8.7	21.8	5.51	5.6	12.8	
BNC % in task-oriented speech	38.8	10.1	21.1	6.6	6.4	17	

Table 68. Verb Form Distribution of DO in MARK 2

In teacher talk all possible forms of the primary verb can be found. In contrast to the

BNC which includes the base form and the present tense form in one category this is not done in the data representation in order to show the present tense use in more detail. But still this form, whether including the base form or not, is the most common verb form in teacher and student talk, just as stated in the BNC. According to the BNC corpus the next verb form in line is the past tense form, this is true for teacher talk not, however, for student talk. In student talk the next form is the –s form, which is uncommon according to the BNC, also in teacher talk it is quite often used with an 11.9 per cent distribution. All the other verb forms are rarely used. Especially interesting is that the infinitive which is one of the most common verb forms in the BNC it is extremely rare in teacher talk and not even once used in student talk. Why this form is so uncommon can not be answered. From the verb form count one can see that the present tense is the predominant tense, this is also expected in the grammatical functions distribution demonstrated in table 69. (Leech 2001: 296-298)

Grammatical Functions of DO

TEACHER				
	Question	Negation	Sum	% of verb from use
do	13	15	28	58.3
does	7		7	14.6
Present	20	15	35	72.9
did	2	11	13	27.1
Past	2	11	13	27.1
Overall	22	26	48	
% of overall distribution of to do	45.8	54.2		
STUDENT				
do	3	11	14	82.4
does		2	2	11.8
Present	3	13	16	94.1
did		1	1	5.9
Past	0	1	1	5.9
Overall	3	14	17	
% of overall distribution of to do	17.6	82.4		

Table 69. Grammatical Functions of DO in MARK 2

The primary *do* can only be used pose questions and negations. The suggestion that teachers foremost use *do* in order to ask questions can not be verified. Negations are with 54.2 per cent more often used than questions. But it can be confirmed that the present tense is predominant with a share of 72.9 per cent. Similar results were

found for student talk, however, with a higher difference in its frequency. Negations have a share of 82.4 per cent and the present tense of 94.1 per cent.

HAVE

The primary verb *have* shows a similar distribution to the primary *be*. Table 70 illustrates the absolute and relative share of its main and auxiliary verb use.

Auxiliary and Main Verb Distribution of HAVE

TEACHER		
	Auxiliary Verb	Main Verb
Absolute frequency	12	57
Relative frequency	17.4	82.6
STUDENT		
Absolute frequency	6	9
Relative frequency	40	60
Classroom overall %	21.4	78.6

Table 70. Auxiliary and Main Verb Distribution of HAVE in MARK 2

Have occurs 69 times in teacher talk and only 15 times in student talk. In contrast to the primary *do*, this primary is used more often as a main verb than as an auxiliary verb. The percentage, however, differs strongly. In the teacher data set the share for the main verb is 82.6 per cent and in the student data it is only 60 per cent. The next point described is the verb form distribution. The verb form count shows which tense is predominant.

Verb Form Distribution of HAVE

TEACHER				
	have	has	had	base form
Absolute frequency	60	4	2	1
Relative frequency	87	5.8	2.9	1.3
STUDENTS				
Absolute frequency	7	5		3
Relative frequency	46.7	33.3		20
BNC Findings				
BNC % in conversations	44.5	17.9	13	
BNC% in task-oriented speech	46.4	15.9	14.6	

Table 71. Verb Form Distribution of HAVE in MARK 2

According to the BNC the most common verb form is *have* (base and present form included) followed by the –s form and past tense form. (Leech 2001: 298) This distribution can also be found in teacher talk but not in student talk as they only use two different verb forms. The past tense does not occur once.

The last point worth discussing is for which grammatical functions the primary *have* is used.

Grammatical Functions of HAVE

TEACHER				
	Perfect Aspect	Substitution Verb	Sum	% of verb form use
have	1	9	10	83.3
has		1	1	8.3
Present			11	91.7
had		1	1	8.3
Past			1	8.3
Overall occurrence	1	11	12	
% of overall <i>have</i> distribution	8.3	91.7		
STUDENT				
have		5	5	83.3
has		1	1	16.7
Present			6	100
Overall occurrence		6		
% of overall <i>have</i> distribution		100		

Table 72: Grammatical Functions of HAVE in MARK 2

In MARK 2 *have* is only used to form perfect aspect constructions and as a substitution verb. The former occurs only once in teacher talk but not at all in student talk. The latter is used far more often with 91.7 per cent in teacher talk and 100 per cent in student talk. All in all *have* is more often used in present tense.

4.5.3 Distribution of Modal Verb

Just as in all other lessons analyzed also in MARK 2 the modal verb class is the least common one Table 73 shows the absolute and relative distribution of the single modal verbs. Also the overall tense distribution of the modal verbs is given.

List of Modal Verbs and their Distribution

TEACHER					
Present		Past		Number of distribution	% verb of distribution
can	22	could	17	39	41.9
need	4	needed		4	4.3
will		would	13	13	14
may	4			4	4.3
must	2			2	2.2
might	5			5	5.4
		should	26	26	28
Sum	37		56	93	
% of tense distribution	39.8		60.2		

STUDENT					
can	28	could		28	53.8
need	5	needed		5	9.6
will	5	would	7	12	23.1
may	2			2	3.8
might	4			4	7.7
		should	1	1	1.9
Sum	44		8		
% of tense distribution	84.6		15.4		

Table 73. List of Modal Verbs and their Distribution in MARK 2

In teacher talk seven different modal verbs were found. The mostly used modal verb is *can* and its past tense form, which has a distribution of 41.9 per cent. Also very common is *should* with 28 per cent and *would* with 15 per cent. All in all, in the modal verb class the past tense is more often applied than the present tense. In student talk six different modal verbs were found. The most common modals are once again *can* with a share of 53.8 per cent and *would* with 23.1 per cent. The remaining verbs are used very seldom. In contrast to teacher talk here the present tense is more often used than the past tense.

4.5.4 Lexical Verb Distribution

Lexical verbs are those verbs which convey what is done, what is happening, and what someone thinks. Therefore it is not surprising that this verb group has a very high distribution. Table 74 shows its verb form distribution. The BNC summarizes the base form and the present tense form into one category which is found in the

category of the base form. This is not applied for the data representation.

Lexical Verb Form Distribution

		TEACHER						
		base form	-s form	present form	past form	-ing participle	-en participle	to infinitive
Absolute frequency		88	29	104	20	31	25	72
Relative frequency		23.8	7.8	28.1	5.4	8.4	6.8	19.5
		STUDENT						
Absolute frequency		49	11	46	20	4	5	14
Relative frequency		33.8	7.6	31.7	14	2.8	3.5	9.7
		BNC Findings						
BNC % in conversation distribution		27.7	6.0		14.3	10.8	12.9	28.2
BNC % in task-oriented speech		23.3	5.1		9.6	12.4	17.2	32.3

Table 74. Lexical Verb Form Distribution in MARK 2

In the BNC the infinitive and the base form are the most common lexical verb forms. In task-oriented speech the third from is the *-en* participle then the *-ing* participle and last the past form. (Leech 2001: 296-297) In teacher talk the most used form is the base and present tense form with together over fifty per cent followed by the infinitive. The least used verb form is the past tense form. As can be observed the ranking corresponds neither to the one of conversations nor to task-oriented speech. The most interesting facts are that the past tense is very rarely used and that the *-s* form is more common than in the BNC. In student talk one finds yet another order. A similarity is that the present/ base form is the most common with over sixty per cent. Then not the infinitive but the past tense form follows. It is very interesting that the past tense is so common in student talk but not in teacher talk. After the past tense form finally follows the infinitive and then the *-s* from. Both participles, which should be more often used, are very rare in student talk.

4.5.5 Semantic Domains and Lexical Verb Distribution in MARK1 and 2

Now that the verb form distribution has been established another factor worth considering is the semantic domain share and which verbs are most common in marketing lessons. Table 75 shows the absolute share of the individual marketing lessons as well as the total numbers which are compared to the results of the LSWE corpus (Biber, Johansson, Leech, Conrad & Finegan 1999:365-366). The sample verbs given for the individual semantic domains all stem from the analyzed data.

<u>Semantic Domain Distribution</u>					
	MARK 1	MARK 2	Total number	% of distribution	% of LSWE
Activity	57	73	130	56.5	43.8
Mental	17	22	39	12.6	27.9
Communication	12	18	30	13.0	11.9
Occurrence	6	5	11	4.8	1.5
Existence	4	4	8	3.5	11.9
Causative	2	4	6	2.6	1
Aspectual	3	3	6	2.6	1

Table 75. Semantic Domain Distribution in HIST 1 and 2 Compared to the LSWE

As can be seen the order of frequency found in the LSWE is not the same as in the classroom data. In marketing lessons activity verbs, such as *buy*, *pay* and *use* are more frequent than the LSWE suggests (Biber, Johansson, Leech, Conrad & Finegan 1999:365-366). 56.5 per cent of all verbs in MARK belong to the domain of activity verbs but only 43.8 per cent according to the LSWE. While in the LSWE corpus mental verbs, such as *consider*, *find*, *like*, and *see* follow in the marketing lessons mental as well as communication verbs, like *ask*, *speak*, *talk*, and *thank* follow. Existence verbs should also show a fairly high distribution in the classroom data they, however, are very rare. Only verbs like *look*, *include*, *seem*, and *stay* could be found. Causative verbs, such as *let* and *allow*, are very rare as well as aspectual verbs as *finish* and *bring*.

The last point taken into consideration is the ranking of the most common lexical verbs in marketing lessons compared to the LSWE (Biber, Johansson, Leech,

Conrad & Finegan 1999: 373-377) and BNC (Leech 2011: 296-297). The bold verbs in the data section were found in the LSWE and the underlined words in the BNC. Those words which are neither bold nor underlined are not included in the most common verb lists of the referential corpora. Some words of the corpus sections are not found in the list of most common lexical verbs of MARK but it will be mentioned if they, nevertheless, appear in the classroom data.

The most Common Lexical Verbs

	MARK		LSWE		BNC
1.	<u>think</u>	81	get		say
2.	<u>know</u>	71	go		get
3.	buy	58	say		make
4.	<u>make</u>	50	know		go
5.	want	35	think		see
6.	<u>use</u>	29	see		know
7.	go	25	want		take
8.	tell	21	come		think
9.	start	17	mean		come
10.	<u>get</u>	17	take		give
11.	<u>see</u>	15	make		look
12.	find	12	give		use

Table 76. Most Common Lexical Verbs in MARK 1 and 2 Compared to the LSWE and BNC

Six of the marketing lesson's most common lexical verbs can also be found in the LSWE and the BNC, all of the verbs listed by the corpora can be found in the data as well. However, as their frequency is lower than five and as they do not occur in each set of data they cannot be found in the frequency list of marketing lessons. This can be led back to the fact that the classroom corpus consists of a smaller number of verbs. Significant are the verbs which are frequent in the classroom corpus but not in the LSWE and BNC. As these verbs are used so often they seem to be essential in marketing lessons. In order to prove this point further research needs to be done.

4.5.6 Summary of Findings

This section is a summary of the most important findings and results of MARK 1 and MARK 2. The results of the data analysis are compared to the BNC and LSWE and important conclusions are drawn. It is referred to the results LSWE (Biber,

Johansson, Leech, Conrad & Finegan 1999: 359-481) and BNC (Leech 2001: 282-297).

Word class distribution analysis has shown that the most common verb class in teacher' talk is the primary verb class. In student talk lexical verbs are more frequent, but the number of primary verbs is exceptionally high, which has to be considered as being unusual in comparison to the LSWE. The primary verb analysis has shown that in all data sets primary verbs are more common in their main verb function than in their auxiliary verb function. The fact that the primary *be* is the most frequently used primary followed by *have* and *do* can be verified by three out of four data sets. Only in MARK 2 student talk *do* is more often used than *have*. Another discrepancy is that both *be* and *have* are more common in their main verb use, but *do* is more frequent in its main verb function. The verb form distribution has revealed that for the verb *be* the past tense is far more uncommon than it is according to the BNC. The most important finding for the primary *have* is that its past tense form is very seldom used which is unusual as the past tense form of this verb is actually the third in line of the most common verb forms of *have*. The infinitive's fairly high use, suggested by the LSWE, can be verified by MARK 1 but not by MARK 2. The analysis of the primary *do* has exposed that in MARK 1 a relatively low number of different verb forms can be found, in contrast MARK 2 shows a variety of forms of the verb. All in all teachers use a higher variety of grammatical functions than students. The modal verb analysis showed the following findings. Teachers use the past tense modal verbs more often and students rather use the present tense. The most common modals are *can*, *could*, *will*, *would* and *should*. Finally, lexical verbs show an interesting distribution. In all data sets the –ing and –en participles are very rare but the corpus analysis has revealed that these participles are used fairly often. Also the infinitive, which is one of the most common verb forms, is used very seldom in student talk and in teacher talk the number of occurrences is higher but not as high as in the BNC. Additionally, the past tense is very uncommon in contrast to corpus findings. This means that the present tense and base form are the most common verb forms. In terms of semantic meaning the data analysis has brought upon the same results as the LSWE corpus except of existence verbs which are extremely rare in the analyzed data. This may be due to the fact that the classroom corpus is considerably smaller. Summing up the present tense is the predominant tense in

MARK 2, also quite often used is the going to future. The past tense used is extremely rare. The table below shows the percentage of tense, aspect, mood and passive voice share. Also tense frequency in all these categories is given as well as the overall distribution.

Tense, Aspect, Mood and Voice in MARK 1 and 2

		TEACHER							
		Tense		Aspect		Mood		Voice	
Number of occurrence		501		8		104		19	
% of occurrence		79.3		1.3		16.5		3.0	
		<i>Present</i>	<i>Past</i>	<i>Present</i>	<i>Past</i>	<i>Present</i>	<i>Past</i>	<i>Present</i>	<i>Past</i>
Times of occurrence		452	49	6	2	47	57	13	5
% of occurrence		71.5	7.8	0.9	0.3	7.4	9.0	2.2	0.8
Overall number		Present		519		Past		113	
% of overall occurrence				82.1				17.8	
		STUDENT							
		Tense		Aspect		Mood		Voice	
Number of occurrence		169		2		57		2	
% of occurrence		63.0		0.9		24.8		0,9	
		<i>Present</i>	<i>Past</i>	<i>Present</i>	<i>Past</i>	<i>Present</i>	<i>Past</i>	<i>Present</i>	<i>Past</i>
Times of occurrence		145	24	2	0	49	8	2	0
% of occurrence		63.0	10.4	0.9		21.3	3.5	0.9	
Overall number				198				32	
% of overall occurrence				86.1				13.9	

Table 77. Tense, Aspect, Mood and Passive Voice in MARK 1 and MARK 2

In both the teacher and student data tense is the most used grammatical structure with 79.3 and 63 per cent, respectively. Mood is also relatively often used but aspect and passive voice are very rare. Summing up, the variety of grammatical functions is low in marketing lessons and the present tense is the dominating tense throughout the data sets.

4.6 Differences between History and Marketing Lessons

In this final part of the data analysis the similarities and differences of history lessons and marketing lessons are described. The aim is to show if the hypothesis is that history lessons show a predominant use of the past tense and that in marketing lessons the present and future tenses dominate can be verified or not. Furthermore, it is claimed that teachers use a higher variety of grammatical functions such as aspect and passive voice.

Firstly, the verb class analysis has shown that in all sets of data primary verbs show an extremely high appearance. What is more, in teacher talk the primary verbs are the most frequent verb class and in student talk although it is not the most common verb class. Primary verbs are only slightly less often used than lexical verbs. Modal verbs are the least common verb class in all sets of data. This differs to the findings of the LSWE (Biber, Johansson, Leech, Conrad & Finegan 1999: 359-360), in which lexical verbs are the most frequently used verbs followed by primary verbs and modal verbs.

Secondly, although the primary verb frequency is so high, in all sets of data they occur more often as main verbs than as auxiliary verbs. The primary *be* is always the most used primary followed by *have* and then *do*. Only in student's talk of MARK 2 *do* is more common than *have*. Even though in general the main verb use is predominant for the primary verb *do* this is not the case. There the auxiliary verb use is predominant. There is one fact which distinguishes history from marketing lessons. In history lessons the past tense forms of the verb have a far higher distribution than in marketing lessons in comparison to the corpus findings. In marketing lessons the past tense forms are less often used than suggested by the LSWE. Another finding is that in history lessons passive voice is the most common structural form when using the primary as an auxiliary verb. In HIST 1 and HIST 2 this is not the case. Here the *going to* future is the most used grammatical structure. In all the data sets, except MARK 2, when *do* is used in teacher talk it is first and foremost to form question functions. In student talk this is not the case, there it is more commonly used to form negative sentences. The primary verb *have* is mostly used as a substitution verb and not to express aspect or voice, in its auxiliary verb function.

Thirdly, modal verbs occur in teacher talk more often in their past tense form. This is also true for student talk in HIST1 and HIST 2, but not in student talk of MARK 1 and MARK 2. In all the data samples *can* and *could* are the most common modal verbs. Also *would* and *will* are fairly frequent but more in marketing lessons than in history lessons.

Fourthly, the main difference between history and marketing lessons when it comes to lexical verbs is the verb form distribution. In the history data sets past tense forms occur very often, which is uncommon regarding the corpus. In marketing lessons the past tense shows a fairly normal distribution, but the participles are extremely rare, which is due to the fact that passive voice and aspect are less frequently used in marketing lessons than in history lessons.

In conclusion it may be said that past tense is very dominant in history lessons, although it is not the dominant tense, both tenses have approximately same share with 50 per cent. In marketing lessons the present tense is with over eighty percent the clearly predominant tense. It can also be considered to be true that teachers use a higher variety of grammatical functions than students. Aspect and passive voice are by far more common in history lessons than in marketing lessons. This leads to the fact that mood is more frequent in marketing lessons than in history lesson. Another factor is that mood, aspect and passive voice are all more common in teacher talk than in student talk. Furthermore it has been presupposed in the formulated hypothesis that modal verbs are used by teachers and students equally often. This did not prove to be right, teachers use modal verbs twice more often than students. All this is true for both the history lessons as well as the marketing lessons.

Conclusion

In conclusion it may be said that the proficiency of both teachers and students in terms of verb and grammatical structure usage are very high and the results are similar to the findings of the corpus done on native speaker's speech events. The outcome of the analysis is very similar to the results found in the corpora on spoken English which suggests that in CLIL classroom discourse the participants show a high level of proficiency. Furthermore, the hypothesis that in history lessons the past tense is a dominating tense in contrast to marketing lessons could be verified. Also that marketing lessons show a higher number of future tense occurrences. All in all present and past tenses in their simple forms are predominant. Aspect constructions as well as passive voice are very rarely used. It has also been proven that such grammatical constructions are more common in teacher talk than in student talk. While these hypotheses' were verified another one was not. It was suggested that modal verbs are used equally often in student and teacher talk. This is not true as modal verbs are more frequently used in teacher talk. All in all the findings were very similar to the ones of the corpus based on data collected from native speakers the CLIL concept seems to have fulfilled the goal of reaching a high L2 proficiency for students.

Bibliography

- Aarts, Bas. 2006. *The verb in contemporary English*. Cambridge: Cambridge University Press.
- Altenberg, Evelyn P. & Roberta M. Vago. 2010. *English grammar. Understanding the basics*. Cambridge: Cambridge University Press.
- Baker, Paul. 2006. *Using Corpora in Discourse Analysis*. London: Continuum.
- Biber, Douglas and Stig Johansson, and Geoffrey Leech, and Susan Conrad, and Edward Finegan. 1999. *Longman Grammar of Spoken and Written English*. London: Pearson Education Limited.
- Biber, Douglas and Susan Conrad and Randi Reppen. 1998. *Corpus Linguistics: Investigating Language Structure and Use*. Cambridge: Cambridge University Press.
- Brannen, J. 2005. *Mixed Methods Research: A discussion paper*. Southampton: ESRC National Centre for Research Methods. <http://www.ncrm.ac.at/publications/documents/MethodsReviewPaperNCM-005.pdf> (March 25, 2012)
- Comrie, Bernard. 1985. *Tense*. Cambridge: Cambridge University Press.
- Coyle, Do and Philip Hood and Marsh David. 2010. *CLIL: content and language integrated learning*. Cambridge: Cambridge University Press.
- Dalton-Puffer, Christiane. 2007. *Discourse in content and language integrated learning (CLIL) classrooms*. Amsterdam: John Benjamins Publishing Company.
- Dalton-Puffer, Christiane. 2010. *Language use and language learning in CLIL classrooms*. Amsterdam: John Benjamins Publishing Company.

- Dalton-Puffer, Christiane. 2007. *Empirical perspectives on CLIL classroom discourse*. Frankfurt am Main: Peter Lang.
- Declerck, Renaat. 2006. *The grammar of the English verb phrase.1. The grammar of the English tense system: a comprehensive analysis*. Berlin: Mouton de Gruyter.
- Dörnyei, Zoltán. 2007. *Research Methods in Applied Linguistics*. Oxford: Oxford University Press.
- Frawley, William. 1992. *Linguistic semantics*. Hillsdale: Erlbaum.
- Greenbaum, Sidney. 1991. *An introduction to English Grammar*. Harlow: Longman.
- Joos, Martin. 1964. *The English verb: form and meaning*. Madison: University of Wisconsin Press.
- Kennedy, Graeme. 1998. *An introduction to corpus linguistics*. London: Longman.
- Leech, Geoffrey N.. 1974. *Semantics*. Harmondsworth: Penguin.
- Leech, Geoffrey N.. 2009. *Change in contemporary English: a grammatical study*. Cambridge: Cambridge University Press.
- Leech, Geoffrey N.. 1991. *Word frequencies in written and spoken English*. Harlow: Longman.
- Lüdeling, Anke. 2008. *Corpus linguistics: an international handbook*. Berlin: de Gruyter.
- Markee, Numa. 2000. *Conversation Analysis*. London: Lawrence Erlbaum Associates, Publishers.

- McEnery, Tony & Andrew Wilson. 1996. *Corpus Linguistics*. Edinburgh: Edinburgh University Press.
- McEnery, Tony & Andrew Wilson. 2001. *Corpus Linguistics: an introduction*. Edinburgh: Edinburgh University Press.
- McEnery, Tony and Richard Xiao and Yukio Tono. 2006. *Corpus-based language studies: an advanced resource book*. London: Routledge
- Meyer, Charles F.. 2002. *English corpus linguistics: an introduction*. Cambridge: Cambridge University Press.
- Palmer, Frank R.. 1988. *The English verb*. London: Longmans.
- Palmer, Frank. R.. 1966. *A linguistic study of the English verb*. London: Longmans.
- Palmer, Frank 4. 1976. *Semantics: a new outline*. Cambridge: Cambridge University Press.
- Rasinger, Sebastian. 2008. *Quantitative Research in Linguistics: An Introduction*. London: Continuum International Publishing Group.
- Thomas, Jenny & Geoffrey N. Leech. 1996. *Using corpora for language research: studies in the honour of Geoffrey Leech*. London: Longman.
- Teubert, Wolfgang. 2007. *Corpus Linguistics*. London: Routledge.
- Twaddell, William F.. 1960. *The English verb auxiliaries*. Providence: Brown University.
- Walsh, Steve. 2006. *Investigating Classroom Discourse*. London: Routledge Taylor and Francis Group.

Zusammenfassung

Die Diplomarbeit mit dem Titel „Aspects of the English Verb in CLIL Classroom Discourse“ zielt durch die Analyse von Schulstunden in denen Englisch als Arbeitssprache verwendet wird darauf ab den Sprachgebrauch im *Integriertes Fremdsprachen und Sachlernen* (CLIL) zu erforschen. Es werden verschiedene Aspekte des Englischen Verbs, wie Verb- Klassen, Verb- Formen und die Verwendung von verschiedenen Zeiten in Geschichte- und Marketingstunden miteinander verglichen. Ebenso werden Resultate aus dem Feld der Corpus Linguistics herangezogen um festzustellen ob die Verwendung von Verben im Unterricht der von Muttersprachlern produzierten Sprachgeschehnissen ähnelt.

Mehrere Thesen wurden im Hinblick auf dies erstellt. Die SchülerInnen werden von der Lehrkraft in hohem Grad mit Englisch konfrontiert. Dadurch soll die Sprachfertigkeit erhöht werden um dies zu bestätigen werden die Resultate der Datenanalyse mit denen der Korpora verglichen. Zusätzlich wird analysiert ob es Unterschiede im Gebrauch des Verbs zwischen Geschichte und Marketing Stunden gibt. Die These, dass in Geschichte die Mitvergangenheit öfter Verwendung findet als und in Marketing die Gegenwart dominiert wurde aufgestellt. Sowohl die Verb-Verwendung von Lehrkräften und SchülerInnen wird analysiert. Auch wenn die sprachlichen Fertigkeiten der SchülerInnen durch *Integriertes Fremdsprachen- und Sachfachlernen* gesteigert werden sollen, so wird doch angenommen, dass es Unterschiede zwischen dem Sprachgebrauch der Lehrkraft und der SchülerInnen gibt. Folgende Unterschiede werden in Betracht gezogen. Erstens, Lehrer verwenden eine größere Anzahl verschiedener grammatikalischer Funktionen. Zweitens, der Gebrauch des „progressive and perfect aspect“ und des Passives wird öfter von der Lehrkraft verwendet. Modal Verben werden von beiden Gruppen gleichermaßen verwendet.

Die vorliegende Diplomarbeit kann in vier Teile unterteilt werden. Der Erste befasst sich mit der theoretischen Auseinandersetzung mit dem Englischen Verb. Es beinhaltet eine detaillierte Erklärung der einzelnen Verb- Klassen, Verb- Formen und grammatikalischen Funktionen. Im Zweiten Teil wird

der Linustische Bereich der Korpus- Linguistik und deren Resultate zweier Korpora des LSWE (Longman Spoken and Written English) Korpus und BNC (British National Corpus). zum Gesprochenen Englisch vorgestellt. Der Dritte Teil erklärte, dass quantitative Forschungsmethoden für die Datenanalyse verwendet wurden und wie eine Forschung aufgebaut ist. Der Letzte Teil gibt die Resultate der Datenanalyse wieder und vergleicht diese mit den Resultaten des LSWE und BNC verglichen.

Die Datenanalyse hat ergeben, dass es tatsächlich Unterschiede in der Zeitverwendung in Geschichte und Marketing Stunden gibt. Es konnte verifiziert werden, dass die Mitvergangenheit häufiger in Geschichte Stunden vorkommt und Marketing Stunden durch Gegenwartsformen dominiert wird. Weiters verwenden die Lehrkräfte eine größere Anzahl von verschiedenen grammatikalischen Strukturen. Im Bereich der Semantischen Domänen sind starke Parallelen zu finden, allerdings nicht in der Liste der meist verwendeten Wörter. Diese scheinen nach der Datenanalyse stark vom Unterrichtsfach abzuhängen.

ABSTRACT

This thesis aims at proving that *Content and Language Integrated Learning* (CLIL) providing information about the verb uses in CLIL classroom discourse. Various aspects of the English verb, as verb classes, verb forms, the usage of tense and grammatical functions, which occur in history and marketing lessons are compared to each other. Additionally, these results are compared to the findings on spoken English of corpus linguistics in order to evaluate if verb usage in classroom discourse is comparable to the findings of the LSWE and BNC.

As students are exposed to the English language in addition to language class their proficiency should improve. Their verb use and the one of the teacher should be fairly similar to native speakers. In order to confirm this thesis the results of the data analysis and the corpus findings are compared to each other. Furthermore, it is analyzed if there are differences in the use of the English verb in history and marketing lessons. Another hypothesis is that the past tense is a dominant tense in history lessons while the present tense and potentially the future are predominant. Both the teachers' and the students' verb use is analyzed. On the one hand, to show that the teacher is able to confront the students with a high level of language proficiency and on the other hand, to confirm that the students' show a considerably high level of proficiency. Even though it is presupposed that the students' language skills are improved by CLIL it still has to be considered that student and teacher verb usage differs very much. A further hypothesis claims that the teachers use a higher number and variety of grammatical functions, like aspect and passive voice. Finally, modal verbs are supposed to be used equally often by teachers and students respectively.

The diploma thesis at hand can be subdivided into four parts. Firstly, a theoretical description of the English verb and all its aspects is given. Secondly, corpus linguistics is introduced and the findings of both the LSWE and BNC are illustrated. Thirdly, it is described why quantitative methods were used for the data analysis and how research is conducted. Lastly, the results of the data analysis is presented and compared to the findings of the LSWE and BNC.

The results of the data analysis have verified that past tense is dominant in history lessons and present tense and the future are the main tenses in marketing lessons. The analysis has also resulted in proving that teachers use both a higher number and variety of grammatical structures. The use of tense, grammatical forms and verb forms is highly dependent on the subject. This can not be confirmed for the share in semantic domains but not for the most common verb forms which also depend on the subject.

Curriculum Vitae

Persönliche Daten:

Name: Silvia Tiefenbacher
Anschrift: Handlkreuzstraße 26, 3231 St.Margarethen an der Sierning
Geburtsdatum: 24. Oktober 1985
Geburtsort: St. Pölten
Staatsangehörigkeit: Österreich
Familienstand: ledig

Schulische Ausbildung: 1991-1995 Volksschule St. Margarethen
1995-2003 Privatschule der Englische Fräulein, St. Pölten
2004-2012 Lehramtstudium an der Universität Wien

Weiterbildung: 2001 Cambridge Certificate
2002 Ausbildung zur SchulmediatorIn

Beruflicher Werdegang: 2009- 2011 Schülerhilfe St. Pölten
2011-2012 Vienna Business School HAK/HAS Mödling

Sprachkenntnisse: Deutsch, Englisch, Polish

Interessen: Reisen, Geschichte, Musik