Verb complementation by infinitive and *-ing* forms
In British and American English:
A corpus based study

(Bakalářská práce)

Kateřina Dubská

(anglická filologie – španělská filologie)

Vedoucí práce: Mgr. Michaela Martinková, PhD.

OLOMOUC 2010

Prohlášení

Prohlašuji, že jsem tuto bakalářskou práci vypracovala samostatně a uvedla úplný seznam citované a použité literatury.

V Olomouci dne 10.5.2009

....

TABLE OF CONTENTS

1. INTRODUCTION

Verb complementation is one of the most studied and the most complicated areas of English grammar. The field of verb complementation is very extensive therefore I focused only on one part of it – monotransitive complementation.

This thesis deals with the occurrence of the infinitive and the *–ing* form after aspectual verbs, namely *start* and *begin*, in two big corpora of contemporary English – the British National Corpus (the BNC) and the Corpus of Contemporary American English (the COCA). The thesis consists of two main parts – theoretical and practical. First basic facts about each of the corpora are introduced. Detailed information on the composition of the corpora will later allow for a more accurate data analyses. In the second chapter the three most important types of monotransitive complementation are described. As the terminology differs among linguists, mainly Quirk's terminology is used in this thesis. The last chapter of the theoretical part deals with the function of the infinitive and the gerund and it also provides information about aspectual verbs.

The practical part presents a quantitative analysis of the corpus data. The data were retrieved from the COCA and the BNC. In the case of the BNC both Xaira and BYU searching engine were used (BYU searching engine provided the data for the last subchapter and appendices). The practical part focuses on the difference between *start* and *begin* and their complements in American and British English. It analyses spoken and written parts of both corpora separately in order to show the tendencies in spoken and written language. Unlike the BNC, the COCA unfortunately does not provide samples of spontaneous speech and therefore does not allow for the further analyses of speakers. Because of their different structures, the corpora are not always comparable.

2. CORPORA USED

In my thesis I was comparing British and American English and for retrieving the data for my research I used two big corpora – BNC and COCA. In this chapter I would like to mention briefly the history of corpus research and compilation and introduce the corpora I was using.

Although in the 60s the creation of the first corpus of English was accompanied by scepticism, nowadays it is one of the most significant sources for linguistic research and it is widely used by linguists all around the world.

In the 50s Noam Chomsky started to develop his generative grammar theory. 'Generative grammarians increasingly took the position that aims of linguistic theory should not be to record linguistic behaviour but rather describe and account for what users of a language know. This movement 'changed the direction from descriptive studies of performance to the modelling of competence.' (Kennedy 23)

Meyer says that 'explanatory adequacy has always been a high priority in generative grammar, often at the expanse of descriptive adequacy: there has never been much emphasis in generative grammar in ensuring that the data upon which analyses are based are representative of the language being discussed.' (Meyer 3)

In this almost hostile environment Nelson Francis and Henry Kučera started to compile *The Brown Corpus* and became the pioneers of corpora creation and corpora based researches. Since many linguists have followed in their footsteps. Corpora based on *The Brown Corpus* were created (e.g. *Australian Corpus of English*) followed by the corpora for studying spoken English (e.g. *London-Lund Corpus*), diachronic corpora (e.g. *Complete Corpus of Old English*) or specialized corpora (e.g. *Guangzou Petroleum English Corpus*). (Kennedy 23 – 50)

The most notable contemporary corpora are *BNC (British National Corpus), ANC (American National Corpus), COCA (Corpus of Contemporary American English)* or *ICE (International Corpus of English).*

One of the things Chomsky disapproved the most were 'probabilistic models of competence derived from linguistic performance.' But it is the probabilistic aspect that distinguishes corpus-based descriptive linguistics from convention descriptive field work and lexicography and makes corpus linguistics exceptional. (Kennedy 9)

'The purpose of a language corpus is to provide language workers with evidence of how language is really used, evidence that can then be used to inform and substantiate individual theories about what words might or should mean.' (http://www.natcorp.ox.ac.uk/using/index.xml)

Nowadays a 'wide range of research activities have come to be within a scope of corpus linguistics. Analyses can contribute to the making of dictionaries, word lists, descriptive grammars, diachronic and synchronic comparative studies of speech varieties, and to stylistic, pedagogical and other applications. With appropriate software it is easy to study the distribution of phonemes, letters, punctuation, inflectional and derivational morphemes, words, collocations, instances of particular word classes, syntactic patterns, or discourse structures.' (Kennedy 11)

2.1 BNC

British National Corpus is a 100-million-word collection of samples released in 1995. Two other editions have been released since – BNC World Edition (2001) and BNC XML Edition (2007).

'The BNC was designed to be well balanced, with a wide range of genres from written and spoken English, and to be widely accessible for educational, academic and commercial purposes.' (Kennedy 50)

BNC samples come from 90% written and 10% spoken sources. 'The texts in the BNC from written sources consist of about 75% 'informative' prose, all post-1975, and about 25% 'imaginative' (literary works), all post-1960. (Kennedy 50)

'For the written section of the corpus, about 60% is taken from books, about 25% from periodicals, about 5% from published brochures and other ephemera, about 5% from unpublished letters, essays, minutes etc., and the remainder from such sources as plays or speeches which have been 'written-to-be-spoken.''(Kennedy 51)

No written extract included in the corpus has more than 40,000 words.

	texts	word units	%
Imaginative	476	16,496,420	18.75
Informative: natural and pure science	146	3,821,902	4.34
Informative: applied science	370	7,174,152	8.15
Informative: social science	526	14,025,537	15.94
Informative: world affairs	483	17,244,534	19.6
Informative: commerce and finance	295	7,341,163	8.34
Informative: arts	261	6,574,857	7.47
Informative: belief and thought	146	3,037,533	3.45
Informative: leisure	438	12,237,834	13.91

Table1 Written part numbers and percentage

(http://www.natcorp.ox.ac.uk/docs/URG/BNCdes.html#BNCcompo)

The spoken section of the corpus has two parts – demographically sampled part of the corpus and context-governed part of the corpus. For the demographically sampled part, 'established random location sampling procedures were used to select individual members of the population by personal interview from across the country taking into account age, gender, and social group. Selected individuals used a portable tape recorder to record their own speech and the speech of people they conversed with over a period of up to a week.'

(http://www.natcorp.ox.ac.uk/docs/URG/BNCdes.html#BNCcompo)

The context governed part of the corpus consists of four categories: educational and informative events (lectures, talks, educational demonstrations, news commentaries, classroom interaction); business (company talks and interviews, trade union talks, sales demonstration, business meetings, consultations); public and institutional events (political speeches, sermons, public/government talks, council meetings, parliamentary proceedings, legal proceedings) and leisure events (speeches, sport commentaries, talks to clubs, broadcast chat shows and phone-ins, club meetings).

(http://www.natcorp.ox.ac.uk/docs/URG/BNCdes.html#BNCcompo)

Table 2 shows the number of texts and word units in these four categories:

	texts	word units	%
Educational/Informative	169	1,646,380	26.65
Business	129	1,282,416	28.11
Public/Institutional	262	1,672,658	27.08
Leisure	195	1,574,442	25.49

Table 2 Spoken context (http://www.natcorp.ox.ac.uk/docs/URG/BNCdes.html#BNCcompo)

XML (Extensive Markup Language) was used for annotating the texts in the BNC. This encoding system conforms to the TEI (Text Encoding Initiative).[1] 'Automatic word-class tagging was carried out using advanced version of the CLAWS tagging system[2] developed at Lancaster University.' (Kennedy 53)

The biggest disadvantage of the BNC corpus is its size. It poses 'challenges for text capture, tagging, storage and processing capacity.' (Kennedy 53)

'Errors are inevitable in transcription (e.g. there/their), and spelling errors in the original texts, or from optical scanning (clear read as dear) cannot all be corrected in a corpus of that size.' (Kennedy 53)

Also errors or ambiguity occur in Part-of-speech tagging. The compilers of the BNC tested a sample of 50,000 words (45,000 written and 5,000 spoken) to find out how many errors will occur. For the written texts the ambiguity rate was 3.83% and the error rate was only 1.14%. For the spoken part the ambiguity rate was 3.00% and the error rate was 1.17%. From this research can be observed that CLAWS achieved 96-97% accuracy.
(http://www.natcorp.ox.ac.uk/docs/URG/posguide.html#errorRates)

It is impossible to eliminate the errors because 'to proof-read thoroughly the text of 10-million-words, allowing 45 minutes for 1,000 words, would take approximately 75,000 hours, or the work of 40 people for about a year.' (Kennedy 53)

---

[1] 'TEI is a consortium which collectively develops and maintains a standard for the representation of texts in digital form. The TEI guidelines define and document a markup language for representing the structural, renditional, and conceptual features of texts.'(http://www.tei-c.org/Guidelines/)
'The advantage of TEI system is that it provides both a standardized set of tags for insertion in a document and the flexibility for the insertion of texts designed by the corpus compiler.' (Meyer 84)
[2] For tagging the British National Corpus CLAWS5 was used.

The British National Corpus can be used for reference books publishing (dictionaries, grammar books, teaching material), linguistic research (data for studying syntax, morphology, semantic), natural language processing and English language teaching.

(http://www.natcorp.ox.ac.uk/using/index.xml)

2.2 COCA

COCA (Corpus of Contemporary American English), released in 2008, is a 400-million-word corpus created by Mark Davies. It includes 20 million words each year from 1990-2009. It is updated once or twice a year (the most recent texts are from summer 2009) which is according to Mark Davies a big advantage, because BNC is compared to COCA out of date.

COCA, like BNC, consists of two sections: spoken and written. Approximately one fifth of the corpus is spoken. This would have been impossible to achieve by tape recording conversations therefore the transcripts of unscripted conversations already existing in electronic form were used (e.g. TV and radio programs like *All Things Considered* (NPR), *Newshour* (PBS), *Good Morning America* (ABC), *Today Show* (NBC), *60 Minutes* (CBS)).

The written section can be divided into four categories: fiction, popular magazines, newspapers and academic journals. In fiction (79 million words), short stories and plays from literary magazines, children's magazines, popular magazines, first chapters of first edition books 1990-present, and movie scripts are included. The following figure shows the representation of these genres in COCA:

	word units	%
Novels (General)	21,014,258	26.52
Magazine/Journal	31,047,970	39.19
Juvenile	2,946,491	3.72
Science Fiction/Fantasy	15,007,897	18.94
Movie Scripts	9,208,594	11.62

Table 3 Fiction genres in COCA

There are nearly 100 kinds of popular magazines (84 millions words). They cover various topics: African-American (e.g. *Black Enterprise, Ebony*), children (e.g. *Children Today, Parenting*), entertainment (e.g. *People, Rolling Stone*), financial (e.g. *Fortune, Money*), home/food/health/garden (e.g. *Horticulture, Total Health*), news and opinion (e.g. *American Spectator, Washington Monthly*), religion (e.g. *America, Christian* Century), science/technology (e.g. *Astronomy, Science News*), social sciences/fine arts (e.g. *National Geographic, American Heritage*), sports/outdoors (e.g. *Sporting News, Skiing*) and women/men/fashion (e.g. *Cosmopolitan, Redbook*).

The following figure shows the number of words in each category:

	word units	%
African-American	3,397,372	4.21
children	2,263,833	2.80
entertainment	3,689,158	4.57
financial	5,368,557	6.66
home/food/health/garden	13,355,433	16.56
news and opinion	16,354,455	20.28
religion	3,206,581	3.98
science/technology	10,166,010	12.60
social sciences/fine arts	7,746,988	9.60
sports/outdoors	9,828,881	12.18
women/men/fashion	6,284,064	7.79

Table 4 Popular magazines in COCA

The newspaper part (79 million words) consists of texts retrieved from ten newspapers from different parts of the US. The newspapers are: *Associated Press, Atlanta Journal Constitution, Chicago Sun-Times, Christian Science Monitor, Denver Post, Houston Chronicle, New York Times, San Francisco Chronicle, USA Today* and *Washington Post*. In most cases, there is a good mix between different sections of the newspaper, such as local news, opinion, sports, financial, etc.

The last part, academic journals (79 million words), consists of 100 different peer-reviewed journals covering various fields of knowledge: education (e.g. *Education Week, Professional School Counseling*), geography/social science (e.g.

Adolescence, Human Ecology), history (e.g. *American Studies International,*
American Indian Quarterly), humanities (e.g. *African Arts, Poetry*), law/politic
science (e.g. *Journal of International Affairs, Michigan Law Review*), medicine (e.g.
Hospital Topics, Creative Nursing), philosophy/psychology/religion (e.g. *Church*
History, Journal of Psychology), science/technology/agriculture (e.g. *Mercury,*
Natural History) and miscellaneous (e.g. *American Scholar, Teacher Librarian*).

Table 5 shows the number of word units in different journal fields:

	word units	%
education	7,262,956	9.18
geography/social science	14,701,485	18.56
history	11,098,148	14.03
humanities	10,565,526	13.36
law/politic science	8,334,809	10.53
medicine	4,867,413	6.15
miscellaneous	3,330,050	4.21
philosophy/psychology/religion	6,278,466	7.94
science/technology/agriculture	12,687,608	16.04

Table 5 Academic journal fields in COCA

The corpus was tagged by CLAWS7. 'The architecture of the corpus relies on MS
SQL Server[3] relational databases, with n-gram databases that contain contextual information
for each or the 400+million-word in the corpus, as well as other databases containing
information on word forms, part of speech, lemmas, synonyms, customized word lists etc.'
(http://www.americancorpus.org/)

The corpus allows users to search for exact words or phrases, wildcards,
collocates, lemmas or part of speech, semantically-based queries can be carried out.
Users can also compare the frequency of words, phrases, and grammatical
constructions, in at least two main ways:

---

[3] MS SQL Server is a Microsoft data platform by allowing organizations 'to run their most mission-critical applications while lowering the cost of managing the data infrastructure and delivering insights and information to all users.' (for more information go to http://www.microsoft.com/sqlserver/2008/en/us/product-information.aspx)

- By genre: comparisons between spoken, fiction, popular magazines, newspapers, and academic, or even between sub-genres (or domains), such as movie scripts, sports magazines, newspaper editorial, or scientific journals
- Over time: compare different years from 1990 to the present time

(http://www.americancorpus.org/)

Figure 6 shows an example of the COCA query:

Fig. 1 COCA query

The advantages of COCA are that it is four times bigger that BNC so can provide data for lower frequency constructions; COCA is updated every year so it is probably the best source for research on contemporary American English; and also COCA includes about 20% of spoken texts. The only disadvantage of the spoken part is that unlike in BNC there are no recordings of everyday speech (spontaneous dialogues).

3. VERB COMPLEMENTATION

According to Quirk complementation is 'a part of a phrase or clause which follows a word, and completes the specification of a meaning relationship which that word implies. As such, complementation can be either obligatory or optional on the syntactic level.' (Quirk 63)

It is not easy to define the term *complement*. This problem occurs because the word *complement* is ambiguous and linguists use it in different contexts. *Complement*, according to Veselovská, can mean three different things: '1) doplnění (obecně) 2) povinné doplnění transitivního slovesa 3) doplněk (syntaktické funkce)' (Veselovská 17)

Aarts defines *complement* as 'a general term to denote any constituent whose presence is required by another element.' (Aarts 104-105)

Sweet (cited in Matthews) comments that 'transitive verbs…require a noun-word or noun equivalent in the direct object relation to serve as complement to them, that is, complete their meaning.'(Matthews 142) This definition perhaps corresponds to what Veselovská calls 'povinné doplnění transitivního slovesa.'

Matthews gives us many different views and opinions on the scope of the term *complement* in grammars:

'In French grammars (and Romance generally) a complement is any element that follows the predicator, even, and sometimes especially, those that are optional.' (Matthews 142)

'For Hill a complement is any noun or noun construction which is not the subject and which has its normal position immediately after the verb. Hill sees only a difference of meaning between 'object complements' (=objects) and 'non-object complements'.'

'In most transformational work the term refers to complements derived from an embedded sentence.' (Matthews 143)

'In English grammars it applies especially to 'subject (ive)' and 'object (ive)' complements (*happy* and *the treasurer* in *He became happy/treasurer, They made him happy/the treasurer*).

'For Quirk complement and object are different types of elements in 'complementation'.' (Matthews 142)

14

OBJECT ⇒ direct object (*My mother enjoys* **parties**.)

⇒ indirect object (*Mary gave* **the visitor** *a glass of milk.*)

COMPLEMENT ⇒ subject complement[4] (*The country became* **totally independent.**)

⇒ object complement (*Most people consider these books* **rather expensive**)

(Quirk 54-55)

Quirk distinguishes the objects and complements from adjuncts. Adjuncts do not belong to verb complementation.

I gave you (DO) *the book* (IO) *yesterday* (adjunct).

As Matthews says most authors make a distinction between objects and adjuncts, but they use a different terminology. For example Tesniére uses the terms 'actants' and 'circonstants'. 'The 'actants' make up the valency of a verb; so in *I gave you the book yesterday*, the 'actants' are *I* as subject, *the book* as direct object, and *you* as indirect object, which are the dependents required by the trivalent GIVE. An adverb is not an 'actant' but a 'ciconstant'.' (Matthews 122)

I gave you the book (actants) *yesterday (circonstant).*

Another point of view is that of Longacre, 'who distinguishes the 'nucleus' of a clause, such as *I saw him* in *I saw him yesterday*, from its 'periphery'.' (Matthews 122)

I saw him (nucleus) *yesterday* (periphery).

In this thesis Quirk's terminology will be used and discussed in more detail. According to Quirk there are four main types of complementation:

1. Copular (*John is only a boy*.)
2. Monotransitive (*I have caught a big fish*.)
3. Complex transitive (*She called him a hero*.)
4. Ditransitive (*He gave Mary a doll*.)

(Quirk 1170)

---

[4] Quirk notes that 'in place of 'subject complement', the term 'predicative noun' or 'predicative adjective' is sometimes used.' (Quirk 55)

The following table shows Quirk's four complementation categories and their subcategories:

Variants	Example
COPULAR	
Adjectival Cs	The girl seemed restless.
Nominal Cs	William is my friend.
Adverbial complementation	The kitchen is downstairs.
MONOTRANSITIVE	
Noun phrase as object (with passive)	Tom caught the ball.
Noun phrase as object (without passive)	Paul lack confidence.
that-clause as object	I think that we have met.
wh-clause as object	Can you guess what he said?
wh-infinitive as object	I learned how to sail a boat.
to-infinitive as object (without subject)	We've decided to move house.
-ing clause as object (without subject)	She enjoys playing squash.
to-infinitive as object (with subject)	They want us to help.
-ing clause as object (with subject)	I hate the children quarrelling.
COMPLEX TRANSITIVE	
Adjectival Co	That music drives me mad.
Nominal Co	They named the ship 'Zeus'.
O + adverbial	I left the key at home.
O + to-infinitive	They knew him to be a spy.
O + bare infinitive	I saw her leave the room.
O + -ing clause	I heard someone shouting.
O + -ed clause	I got the watch repaired.
DITRANSITIVE	
Noun phrases as direct and indirect objects	They offered her some food.
With prepositional object	Please say something to us.
IO + that-clause	They told me that I was ill.

IO + wh-clause	He asked me what time it was.
IO + wh-infinitive clause	Mary showed us what to do.
IO + to-infinitive	I advised Mark to see a doctor.

Table 6 Verb complementation types (Quirk 1171)

'Many verbs are versatile enough to allow several complementation types. It is therefore likely to be misleading to talk of 'intransitive verbs', 'monotransitive verbs', 'complex transitive verbs', etc. Rather, it is often better to say that verbs have 'monotransitive use', 'monotransitive complementation', etc. Although one verb may belong to a number of different complementation types, it is usually possible to observe a common ground of meaning in the various types.' (Quirk 1168)

Although this division might seem very straightforward, not all linguists agree with it. Mair criticizes Quirk's complex transitive category calling it 'a catch-all.' He argues that many verbs that Quirk classifies as complex transitive are in reality rather monotransitive or ditransitive. Mair mentions for example the case of *cause*, classified by Quirk as complex transitive, which he considers monotransitive. Mair notes that the main criterion for classification used by Quirk, passivization, is not very reliable and should not be used. He offers some reclassifications. (Mair 93-101)

In this thesis I will deal with monotransitive complementation, therefore only monotransitive complementation will be mentioned in greater detail.

3.1 Monotransitive complementation

Quirk distinguishes three main types of monotransitive complementation: 1) by a noun phrase 2) by a finite clause 3) by a nonfinite clause (Quirk 1176)

3.1.1 Noun phrase

According to Quirk a noun phrase in monotransitive complementation can have a function of either a direct object or a prepositional object.

Quirk gives two different types of complementation by noun phrase as a direct object – either with or without passive. The typical monotransitive verbs allowing the passive are e.g. *begin, believe, help, hold, mean, move, need, pass, remember, support, take,win…*

(Quirk 1176)

Example: *Professor Dobbs **won the prize.***

Monotransitive verbs that do not allow passivization are called middle verbs (e.g. have, lack, suit, become, fit, equal…). 'They are all stative relational verbs, and therefore normally do not occur in the progressive.' (Quirk 735)

Example: *They **have a nice house.*** (Quirk 1177)

Huddleston calls the noun phrases related to the verb by preposition *oblique* which he distinguishes from an object. 'The preposition characteristically makes a contribution to identifying the semantic role of the NP.' (Huddleston 216)

Quirk compares a prepositional object to a direct object. He notes there is a resemblance 'in accepting the passive voice, though usually with some awkwardness of style.' (Quirk 1177)

Examples of such prepositional verbs are: *account for, agree with, aim at, apply for, concentrate on, insist on, long for, object to, refer to, rely on, think about/of...* (Quirk 1178)

3.1.2 Finite clause

Quirk describes two types of finite clauses that can participate in monotransitive complementation – that-clause and wh-clause.

'That- clauses have one of three types of verb phrase, depending on the 'governing' verb in the matrix clause.'(Quirk 1180)

There are two major types of superordinate verbs (factual[5] and suasive[6]) and two minor types (emotive and hypothesis verbs).

1. indicative verb

Indicative verb is the most common type and it can occur with factual and emotive verbs in the main clause. In British English also suasive verb in the main clause is sometimes used.

*I suppose that he **is coming** alone.* (Quirk 1180)

2. putative should

---

[5] Factual verbs introduce 'what one might generally describe as factual or propositional information. In term of speech act classes factual verbs are associated with the expression of speech acts concerned with statements. (Quirk 1180)

[6] Suasive verbs 'imply intentions to bring about some change in the future, whether or not these are verbally formulated as commands, suggestions etc. In terms of the speech act classes suasive verbs are associated with directives.' (Quirk 1180)

Putative should is used more in British than American English. It occurs with suasive and emotive verbs in the main clause

*I regret that he **should be** so stubborn.* (Quirk 1180)

3. subjunctive verb

It is more common in American English, in British English it is felt to be rather formal. It occurs with suasive verbs in the main clause.

*I request that she **go** alone.* (Quirk 1180)

The typical verbs that take a that-clause as object are *agree, ask, beg, demand, claim, comment, ensure, intend, order, pray, recommend, remember, suggest, suppose, vote…* (Quirk 1181- 1182)

'Many of the verbs which take a that-clause as object can also take a wh-interrogative clauses. The use of the wh-interrogative clause (which generally implies lack of knowledge on the part of the speaker) is particularly common where the superordinate clause is interrogative or negative. (Quirk 1184)

'There are also some verbs which express uncertainty, such as ask and wonder: these occur with the wh-clause without the non-assertive constraint.' (Quirk 1184)

The verbs taking a wh-interrogative clause as object are *argue, care, demonstrate, depend, explain, learn, know, note, observe, prove, remember, see…* (Quirk 1184)

3.1.3 Nonfinite clause

Quirk distinguishes four main types of monotransitive complementation by nonfinite clause:

1. Subjectless infinitive clause as direct object
2. Subjectless –ing participle clause as object
3. Complementation by to-infinitive clause (with subject)
4. Complementation by –ing participle clause (with subject) (Quirk 1187-1195)

Palmer introduces the term catenatives for these four categories. Palmer suggests that the catenatives share some grammatical characteristics with the auxiliary verbs.

Palmer also points out that some similar constructions should not be mistaken for catenatives. For example infinitives of purpose and infinitives of result:

I ran to catch the train. (I ran in order to catch the train) (Palmer 206)

I ran all the way to find that he had gone. (Palmer 207)

Catenatives can be followed by a bare infinitive (e.g. *He helped wash up*.), to-infinitive (e.g. *He wants to go to London*.), -ing form (e.g. *He keeps talking about it*.) and –en form (e.g *He got shot in the riot*.). A noun phrase can occur between the catenative and the following verb. (Palmer 173)

According to Palmer, catenatives can be divided into nine classes:

1. Futurity

These verbs refer to plans etc. for the future. The verbs that belong to this category are e.g. *choose, wish, desire, long, prepare, invite, expect, lead, teach, promise.* (Palmer 191-194)

Example: *I wish to meet Mary.* (Palmer 191)

2. Causation

Common verbs in this class are e.g. *help, let, make, have, get, want.*

Example: *He had them come early* (Palmer 195)

3. Report

Verbs belonging into this group are e.g. *believe, accept, certify, prove, read, report, state*.

Example: *I believe John to be clever.* (Palmer 196)

4. Perception

Verbs from this class are e.g. *see, feel, hear, smell, notice.*

Example: *I saw the children eat their lunch*. (Palmer 199)

5. Process

The most common verbs from this class are e.g. *keep, finish, start, cease, leave, stop.*

Example: *He kept talking*. (Palmer 200)

6. Achievement

Verbs that belong in this category are e.g. *manage, attempt, struggle, try*.

Example: *He managed to come*. (Palmer 202)

7. Attitude

The most common verbs are e.g. *like, enjoy, mind, resent, welcome, delight.*

Example: *I like going to the theatre.* (Palmer 203)

8. Need

Example: *The boys need to watch*. (Palmer 204)

9. Appearance and change

The verbs belonging in this class are e.g. *seem, happen, appear.*

Example: *John seems to like Mary*. (Palmer 205)

4. *ING* FORM AND INFINITIVE

Ing form and infinitive are nonfinite verb forms. Nonfinite forms unlike finite forms do not contrast in number, person and tense:

*He wanted **to do** it.* (infinitive)

*I remember **putting** the book on the table.* (gerund-participle)

*His father got **charged** with manslaughter.* (past-participle) (Huddleston 1174)

'The modal auxiliaries and supportive do are excluded from all nonfinite clauses.' (Huddleston 1174) Infinitives accept perfect *have*, progressive *be* and passive *be* while gerunds do not admit progressive be.

The past participle has two uses – perfect and passive. Perfect past participle occurs as a complement of auxiliary have (e.g. *She has **done** it*.) and it allows progressive and passive be. Passive past participle does not allow any auxiliaries. (Huddleston 1174)

4.1 *Ing* form

As Kruisinga notes the –*ing* form is one of the most important and interesting structures in English. (Kruisinga 247) Words containing the –*ing* suffix can be used as verbs, adjectives and also substantives. Many authors differentiate gerund and present participle.

4.1.1 Gerund

Gerunds share some features with substantives. Like substantives, gerunds usually do not indicate time (*an account of his coming* – the meaning can be present, past or future). With certain verbs gerunds can refer to time though. In the sentence *I remember locking the door*, the gerund form refers to past. The perfect form of gerund unlike the simple form always expresses time, namely the past. The perfect form can be redundant when following some prepositions (e.g. *after*) (Jesperson 250-251)

Zaandvoort adds that gerund shares also some syntactic properties with a verb. 'It may be qualified by an adverb or adverbial phrase, and in the case of a transitive verb may govern an object. It may also take a subject of its own. It may be used in the perfect tense (having written) and in the passive voice (being written).' (Zaandvoort 25)

He also points out that gerund shares many syntactic properties with the infinitive. They both can appear in the function of the subject, object or nominal predicate; they both can be modified by an adverbial phrase or take an object or subject; both can be used in a passive voice. (Zaandvoort 25)

Also gerund, like the infinitive, may have in the clause the same agent as the finite verb or it can have its own agent. 'In the latter case the gerund is usually preceded by the stem of a noun or indefinite pronoun, or by a possessive pronoun.'(Zaandvoort 30)

4.1.2 Present participle

'As a general rule, present participle expresses an action or a state simultaneous with that expressed by the predicate of the sentence.' (Zaandvoort 32)

One of the most important features of the present participle is the progressive aspect. The progressive aspect of the present participle in predicative use can have several functions. Zaandvoort points out that 'the verbal adjectival character of the present participle makes the progressive more descriptive that the purely verbal character of the simple form.' (Zaandvoort 38) when the progressive aspect has the emotional character, it can express also annoyance or irritation, e.g. *I'm losing keys all the time.* (Zaandvoort 39)

Leech gives more detailed study of the progressive aspect. With present and past tenses, progressive aspect can refer to 'temporary situations, activities, or goings-on.' (e.g. *She is having a bath.*) The progressive aspect indicates duration, limited duration and it indicates that the happening need not be complete. The progressive aspect can also serve as 'temporal frame' (e.g. *I was having a bath when somebody rang.*). Other use of progressive aspect is as Leech adds habitual or iterative use; habitual use means that some action was done temporary over a period of time (e.g. *She is going to work by tram until her car is repaired.*)

Some verbs do not usually occur with progressive aspect. This is the case of 'verbs of inert perception' (feel, hear, see, smell...), 'verbs of inert cognition' (believe, forget, hope...), 'state verbs of being and having' (be, belong to, contain, depend on...) or 'verbs of bodily sensation' (ache, feel, hurt, itch...). (Leech 18-34)

Even though this division of -ing form into present participle and gerund may seem very straightforward and clear, not all authors agree with it. Kruisinga criticizes this division because he says the semantic criterion on which it is based is not

sufficient. He offers his own division based on syntactic criteria. He describes three main functions of verbal ing – the usage in a group with another word for its leading member, in free adjuncts and as grammar subject. (Kruisinga 247-276)

4.2 Infinitive

Like *-ing* form infinitive shares also some common features with substantives. The main reason Jesperson gives is that in past the infinitive was 'a fully inflected verbal substantive.' (Jesperson 329) Even though the infinitive developed into a purely verbal form, it still retains some of the characteristics of substantives, e.g. it can stand as a subject or object. (Jesperson 329)

There are two types of infinitives – the bare (or plain) infinitive and *to* infinitive. Bare infinitives are used after modal auxiliary verbs (*will, shall, would, should, can, could, ma*y, *might, must*); after *had better*, *need* and *dare;* after *let, make hear, feel*; after *why(not)*; after *and, or, except, but, than, as* and *like*; after *do*. (Swan 256-257)

Dušková gives us a list of all the possible forms of *to* infinitive:

infinitiv	přítomný	minulý
prostý činný	to carry	to have carried
průběhový činný	to be carrying	to have been carrying
trpný	to be carried	to have been carried

Table 7 Forms of the infinitive (Dušková 267)

4.3 Infinitive vs *-ing* form

A lot has been written about the competition of the infinitive and the *–ing* form. For the purposes of this paper, let me mention two important points.

As was mentioned before gerund shares many features with infinitive. Poldauf provides a detailed study of the competition between gerund and infinitive. According to Poldauf infinitive 'má za svou základní funkci vyjadřovat děj jako abstrakci, něco pomyslného, neskutečného, nekonkrétního, případně generického.' (Poldauf 205) As a secondary function infinitive can 'označovat i děje skutečné, konkrétní, individuální.' (Poldauf 205) Gerund also denotes 'děje skutečné' and

therefore as Poldauf notes, the competition between gerunds and infinitives is based on this unclear division.

Also I would like to mention the distinction between the plain (or bare) infinitive and present participle. In contrast with the plain (or bare) infinitive, present participle expresses an action that is not completed or is in progress, e.g. *She heard him coming.* x *She heard him come*. In the case of the bare infinitive the duration of the action is considered irrelevant by the speaker. Zaandvoort notes that this difference is in aspect, calling the aspect expressed by the present participle *imperfective or durative,* and the aspect expressed by the plain infinitive *perfective.* (Zaandvoort 33)

4.4 Aspectual verbs – *start* and *begin*

The verbs *start* and *begin* belong to a class of verbs which is called *aspectual verbs*. However not all authors agree on this term. Palmer creates a term 'process verbs'. Within this class various verb groups are included (e.g. group of *begin* and *start*, group of *get, keep, stop*) (Palmer 201) Brinton mentions other authors and their terminology (e.g Kruisinga calls them 'auxiliaries of aspect', Joos speaks about 'quasi-auxiliaries'). (Brinton 60) These terms refer to the syntactic status of these verbs in the English language system. Linguists' opinions on this matter vary significantly, some of them consider these verbs 'fully developed auxiliaries', 'regular lexical verbs' or even something in between. (Brinton 59) Considering the semantic criteria, it is agreed that the aspectualizer and the following verb form a single semantic unit. (Brinton 74)

Brinton divides aspectualizers into four groups: ingressive aspectualizers (*begin, commence, start, proceed*), continuative/itinerative aspectualizers (*keep on, go on, continue to*), egressive aspectualizers (*cease, finish, stop, quit*) and habitual aspectualizers (*used to, be accustomed to*). (Brinton 61)

Being ingressive aspectualizers, *begin* and *start* focus on 'the beginning point or initiation of a situation'. (Brinton 60) Freed (cited in Brinton) describes the meaning of *begin* and *start* in terms of 'presupposition (information which is 'prior knowledge shared by speaker and hearer') and consequence (information which is 'conveyed by the speaker and learned by the hearer').' (Brinton 79) As Freed says '*begin* and *start* do not presuppose prior initiation of an event but entail subsequent

occurence of the event.' She admits that *start* can also have the meaning of non-occurence, because 'one can start something but then not do it.' (ibid. 79) Freed also adds that *begin* 'points to the time segment in which an event is initiated' and *start* denotes 'the initial segment if the event itself, or the initiating activity.' (ibid. 80)

Freed also observes that *begin* and *start* occur freely with accomplishments and activities, but less freely with states and achievements. When occurring with states, aspectualizers are followed more commonly by the infinite than by the gerund (although gerund can occur too). (ibid. 85)

According to Zaandvoort *begin* followed by a gerund expresses a 'deliberate act'. He states that *start* normally takes gerund, because with the infinitive it would be considered rather colloquial. (Zaandvoort 28)

Quirk mentions that gerund and infinitive contrast in 'potentiality' and 'performance'.

He ***started to speak***, *but stopped because she objected.* (potentiality)
He ***started speaking***, *and kept on for more than an hour.* (performance)
(Quirk 1192)

Schibsbye explains the difference between the gerund and the infinitive in terms of intentionality. *Begin* followed by gerund expresses intentional action and begin followed by infinitive non-intentional action. Like *begin* followed by gerund, also *start* followed by gerund expresses intentional action. Non-intentional action in the case of *start* is expressed by both gerund and infinitive. (Schibsbye 63)

Poldauf comments that in the case of aspectual verbs 'řídící sloveso se blíží ke gramatickým zařízením.' (Poldauf 221) He adds that 'I je při konkurenci s G doplněním obecným, kdežto G nastupuje tam, kde jde o objektivně pojatý předmět záměru, a vyskytuje se proto při vypovídání o subjektu označujícím osobu, organisaci nebo orgán.'(Poldauf 221)

Freed (cited in Brinton) argues that the infinitive produces 'a generic (or serial) reading' while gerund produces 'a durative (or itinerative) reading.' (Brinton 92) She defines a generic reading as a 'repetition of events of the same kind on different occasions over an unspecified period of time' and the durative reading can be defined as 'the duration of a single event or the repetition of an event within a single period of time'. (ibid. 92)

Macpherson notes that *begin* is more formal that *start*, and he goes on to say 'the infinitive is preferred when the grammatical subject is lifeless' (e.g. *The ship began to sink.*) and when 'the predicate denotes a state of mind' (e.g. *He began to feel irritable*.) (Macpherson 215)

Swan gives us some cases when the usage of *begin* is not possible – 'start a journey', 'start working' (about machines) or 'make something start'. (Swan 85)

5. CORPUS DATA

5.1 General overview

The biggest problem I was facing when working with the COCA and the BNC was their size. The results have to be compared carefully because every corpus is of different size and also their structures are distinct.

As Table 8 shows in both corpora the most frequent construction is *begin + V infinitive* (18,233 tokens in the BNC and 38,129 tokens in the COCA) and the least frequent is *begin+ V ing* (2,758 tokens in the BNC and 18,487 tokens in the COCA).

	BNC	COCA
start + V ing	5,911	35,271
start + V infinitive	5,979	20,623
begin + V ing	2,758	18,487
begin + V infinitive	18,233	38,129

Table 8 Number of tokens (lexical verbs) in the COCA and the BNC

The rate between *start +V ing* and *begin +V ing* is approximately 2:1 in both the BNC and the COCA. The rate differs in the case of *begin + V infinitive* and *start + V infinitive*, being approximately 3:1 in the BNC and 2:1 in the COCA.

Table 9 presents the occurrence of the verbs *have, be* and *do* after *start* and *begin.*[7]

	BNC				COCA			
	start + V ing	start + V infinitive	begin + V ing	begin + V infinitive	start + V ing	start + V infinitive	begin + V ing	begin + V infinitive
have	51	36	6	153	454	247	193	526
do	180	63	8	46	1174	287	214	368
be	47	62	4	666	229	366	28	1409

Table 9 Number of tokens (have, do and be) in the COCA and the BNC

The prevalence of *begin + V infinitive* appeared in both corpora in the case of *have* (153 tokens in the BNC and 526 tokens in the COCA) and *be* (666 tokens in the

---

[7] These had to be searched separately, because the general tag for –ing forms (VVG) and infinitives (VVI) is only used for lexical verbs. Instead the tags VHG,VDG,VBG and VHI, VDI, VBI were put in the queries.

BNC and 1409 tokens in the COCA). The most frequent construction for *do* was in both corpora *start + V ing* (180 tokens in the BNC and 1174 tokens in the COCA). The construction with the lowest number of tokens was in both corpora *begin + V ing*.

In the following chapters I will consider the results for written and spoken language separately.

5.2 Spoken part

The results in the spoken part differ significantly from the general results I presented before. The prevalence of *begin + V infinitive* did not show up at all. The results changed dramatically in favor of *start* taking the place of *begin.* Table 10 presents the number of tokens (lexical verbs) for queries within the spoken parts of the BNC and the COCA.

	BNC spoken	COCA spoken
start + V ing	1,433	10,749
start + V infinitive	668	5,163
begin + V ing	31	2,150
begin + V infinitive	477	5,937

Table 10 Number of tokens (lexical verbs) in the spoken parts of the COCA and the BNC

The most frequent construction in both the COCA and the BNC was *start + V ing* (1,433 tokens in the BNC and 10,749 tokens in the COCA). The rate between *start + V ing* and *begin + V ing* was approximately 46:1 in the BNC and 5:1 in the COCA. The rate between *start + V infinitive* and *begin + V infinitive* was approximately 14:10 in the BNC and 10:11 in the COCA.

While in the general overview the most frequent construction was *begin + V infinitive*, in the spoken part it was *start + V ing*.

As for competition between the *–ing* form and the infinitive, the results in the BNC and the COCA differ significantly. The rate between *start + V ing* and *start + V infinitive* is approx. 2:1 in both the BNC and the COCA. But in the case of *begin* the rate is approx. 14:1 in the BNC and approx. 3:1 in the COCA in favor of *begin + V infinitive.* It can be concluded that in spoken language *begin + V ing* occurs more in American English.

Table 11 shows the frequency of occurrence of the verbs *be, do* and *have* after *begin/start* constructions:

	BNC spoken				COCA spoken			
	start + V ing	start + V infinitive	begin + V ing	begin + V infinitive	start + V ing	start + V infinitive	begin +V ing	begin + V infinitive
have	18	5	0	3	172	113	35	132
do	108	19	2	6	572	144	51	124
be	14	8	0	23	90	151	7	248

Table 11 Number of tokens (have, do and be) in the spoken parts of the COCA and the BNC

In both the COCA and the BNC the construction *start + V ing* was the most frequent for *have* (18 tokens in the BNC and 172 tokens for the COCA) and *do* (108 tokens in the BNC and 572 tokens for the COCA). On the other hand *begin + V infinitive* was the most frequent construction for *be* (23 tokens in the BNC and 248 tokens in the COCA).

As stated above, *begin* is more formal than *start* (Macpherson 215). When the sources of all the tokens are analyzed in more detail, however it can be seen that the 31 tokens of *begin +V ing* in the BNC occur not only in formal registers, but also in those that could be considered informal. More specifically, 22 tokens occur in formal spoken texts (meetings, documentaries, school lectures, sermons, radio and TV news) and 9 tokens in informal (recorded everyday speech dialogs, interviews). In the case of *begin + V ing* the number of formal speeches is twice as big as the number of the informal ones.

(1) *the best way of organizing your files and just, just take separate bits of paper now, to save yourselves time, is to take separate pieces of paper for the separate topics, that separate people have done, and as you're talking to each other,* **begin making** *notes and adding to, to these different bits.* (informal) (BNC, KPV-555)

Source:
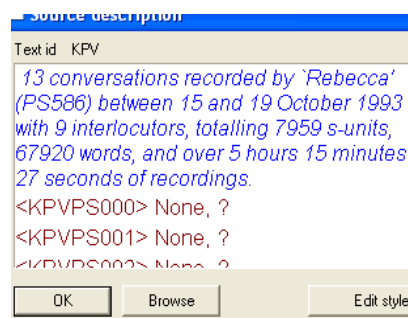
Fig. 2 Source 1

(2) *This company **began searching** for finance four years ago.* (formal) (BNC, HMH-19)

Source:
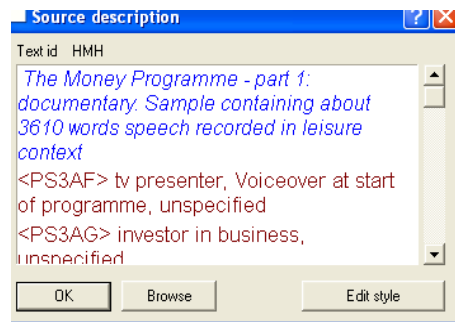
Fig. 3 Source 2

Tokens of *begin + V infinitive* were also found in both formal and informal context:

(3) *It **began to annoy** me.* (informal) (BNC, KCE-5458)

Source:
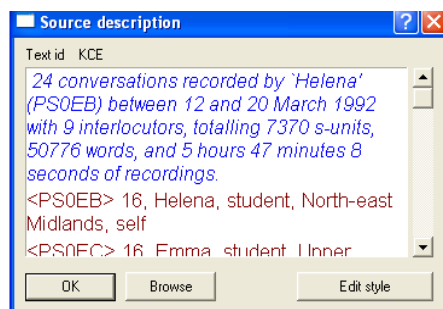
Fig. 4 Source 3

(4)...*to escape that, by fleeing into the countryside, very often into particularly backward and inaccessible areas of the countryside, where they would be safe from*

*the security forces and where they hoped they could **begin to create** a new base among the peasantry.* (formal) (F8R-50)

Source:

Text id F8R

[Leigh Community Centre: lecture on China]. Sample containing about 15816 words speech recorded in educational context

`<PS1PR> 30, Neil, lecturer, unspecified`

`<F8RPSUNK> Unknown speaker, other`

`<F8RPSUGP> Group of unknown speakers, other`

OK   Browse   Edit style
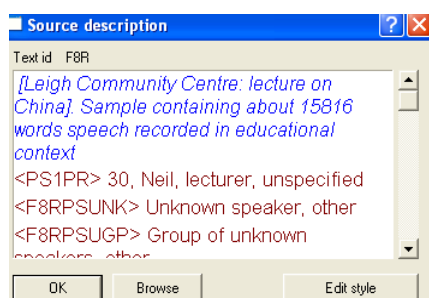
Fig. 5 Source 4

As was mentioned in Chapter 2*,* the spoken part of the BNC consists of context-governed spoken texts (4.30 %) and demographic spoken texts (6.27%).

It follows from Fig. 6 that in the case of *begin + V ing, begin + V infinitive* and *start + V ing* the most frequent context was informative/educational. *Start + V infinitive* occurred mainly in the business contexts:

Context-governed texts

	Informative/Educational	Business	Public/Institutional	Leisure
■ Begin + V ing	12	2	4	6
■ Begin + V infinitive	173	51	82	99
□ Start + V ing	142	103	106	116
■ Start + V infinitive	217	228	116	193
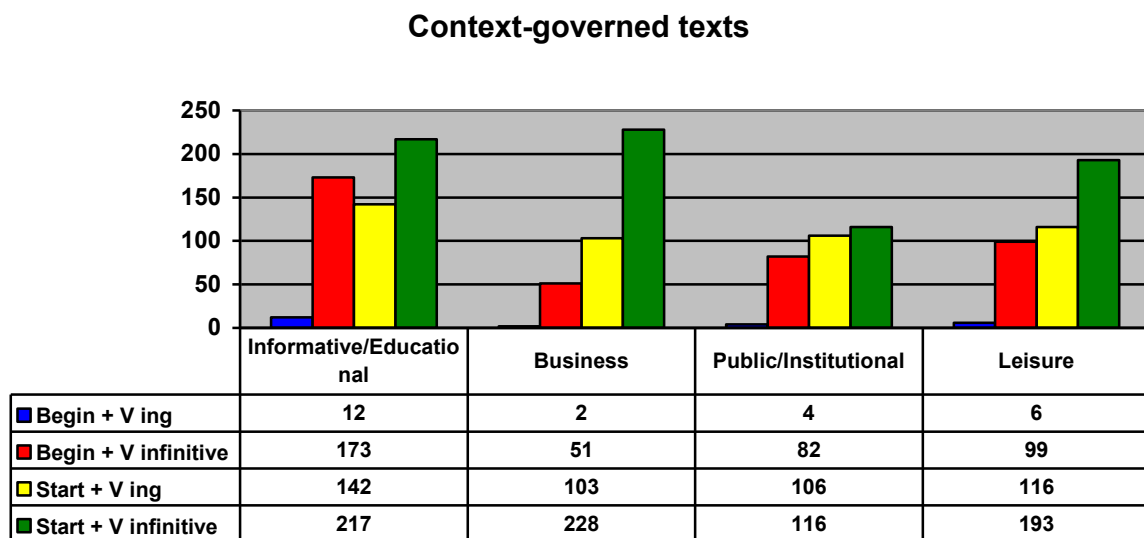
Fig. 6 Graph of context-governed texts in the BNC

The corpus itself composes of 26.65 % educational/informative, 20.76% business, 27.08% public/institutional and 25.49% leisure context, so it is can be considered balanced as far as the distribution of these texts is concerned.

The percentage for all the four constructions were as follows: *begin + V ing* (informative/educational – 50%, business – approx. 8%, public/institutional – 17%, leisure – 25%), *begin + V infinitive* (informative/educational – approx. 43%,

business – approx. 13%, public/institutional – approx. 20%, leisure – approx. 24 %), *start + V ing* (informative/educational – approx. 30%, business – approx. 22%, public/institutional – approx. 23%, leisure- approx. 25%), *start + V infinitive* (informative/educational – approx. 29%, business – approx. 30%, public/institutional – approx. 15%, leisure – approx. 26%).

.

Figure 7 deals with the frequency of the constructions in the speech of speakers with a view to their social background. The highest number of tokens for *begin + V infinitive, start + V ing* and *start + V infinitive* appeared in higher management social class. *Begin + V ing* occurred with the highest frequency in manual skilled social class. However these results might be influenced by the composition of the corpus. The social class in the BNC consists of 32.42% higher management (AB), 26.56% lower management (C1), 25.69% skilled manual (C2) and 14.91% semi-skilled or unskilled (DE). No serious conclusions can therefore be drawn.

Social class

	higher management AB	lower management C1	skilled manual C2	semi-skilled or unskilled DE
■ Begin + V ing	2	0	4	1
■ Begin + V infinitive	27	17	22	5
□ Start + V ing	237	128	196	111
■ Start + V infinitive	65	54	62	19
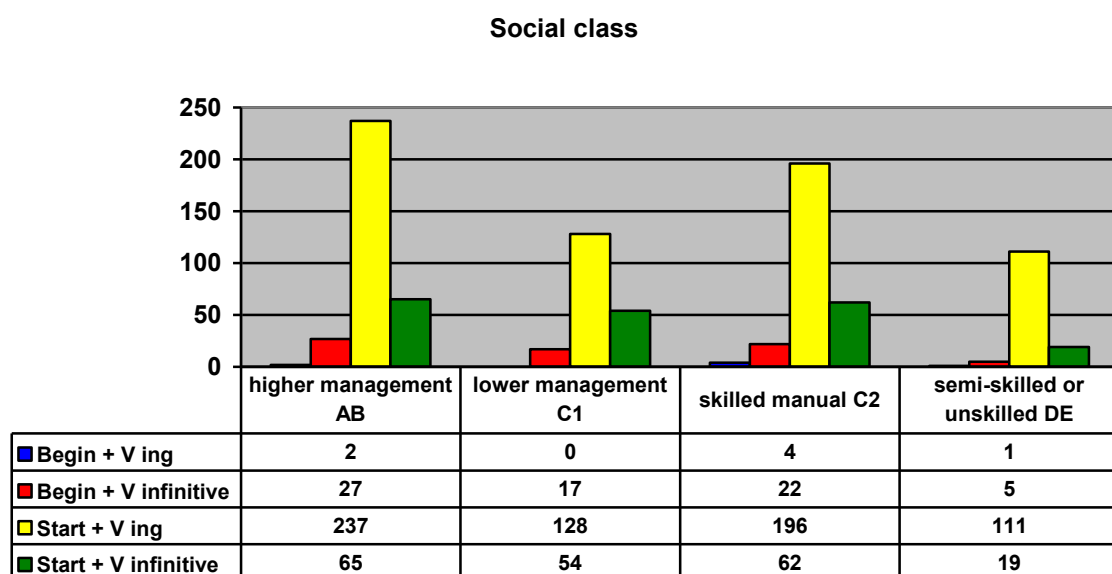
Fig. 7 Graph of social class in demographic part of the BNC

In the case of *begin + V ing* the frequency of occurrence was very low (only 7 tokens – AB – approx. 29%, C1 – 0%, C2 – approx. 57%, DE – approx. 14%). The rest of the data is more reliable: *begin + V infinitive* (AB – approx. 38 %, C1 – approx. 24%, C2 – approx. 31%, DE – approx. 7%), *start + V ing* (AB – approx.

35%, C1 – approx. 19%, C2 – approx. 29%, DE – approx. 17%), *start + V infinitive* (AB – approx. 33%, C1 – 27%, C2 – 31%, DE – approx. 20%).

Figure 8 deals with the occurrence of the constructions in each of the respondents' age categories in the BNC. The corpus is composed of 6 age categories: 0 – 14 category (6.30%), 15 – 24 category (15.71%), 25 – 34 category (20.16%), 35 – 44 category (19.96%), 45 – 59 category (22.75%) and 60+ category (15.09%). The dominance of *start + V ing* showed up again. It was the most frequent construction for all the age categories (approx. 10.9% occurred in the 0 – 14 category, approx. 21.5% in the 15 – 24 category, approx. 19% in the 25 – 34 category, approx 22.5 % in the 34 – 44 category, approx. 15.6% in the 45 – 59 category and approx. 11% in the 60+ category). The percentage more or less corresponds to the distribution of each age category in the corpus. Therefore it could be concluded that the distribution of *start + V ing* is equal in all age categories.

Age

	0 - 14	15 - 24	25 - 34	34 - 44	45 - 59	60+
■ Begin + V ing	0	1	2	1	1	2
■ Begin + V infinitive	1	10	12	15	11	23
■ Start + V ing	71	146	129	153	106	74
■ Start + V infinitive	13	24	50	51	33	30
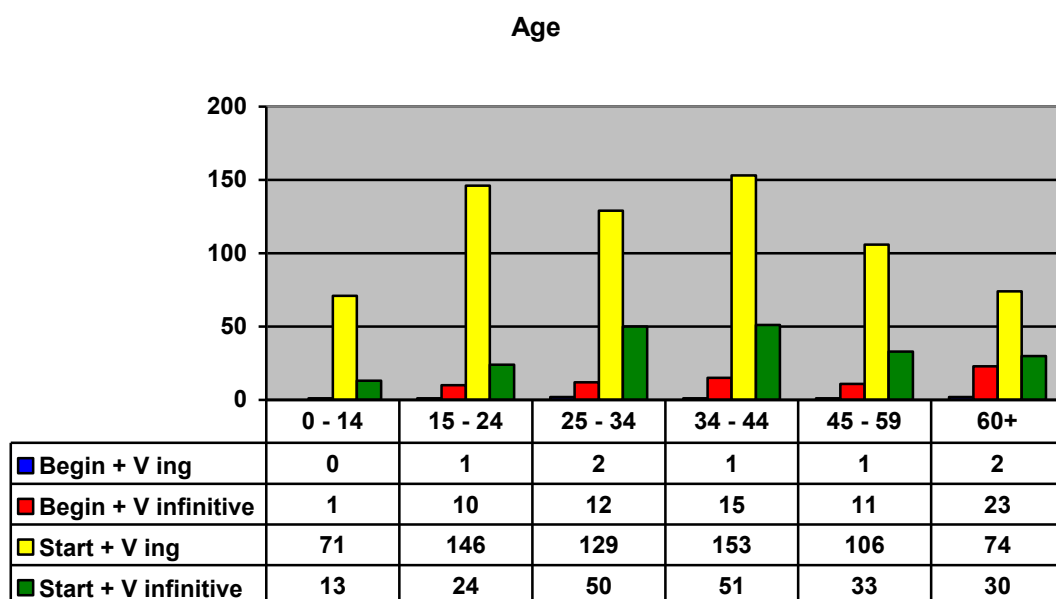
*Fig. 8 Graph of age class categories in the spoken part of the BNC*

The frequency of the construction *begin + V ing* was very low (7 tokens) in all age categories, therefore no reliable percentage can be provided. The construction *begin + V infinitive* was most frequent in the 60+ category (approx. 32%) and the least frequent in 0 – 14 category (only one token). Both *start + V ing* and *start + V*

infinitive occurred with the highest number of tokens in the 34 – 44 category and with the lowest number of tokens in the 0 – 14 category.

Unfortunately the COCA does not allow more specific searches within its spoken part. Besides, there is no demographic component and therefore no spontaneous dialogues. For this reason any further analysis is impossible.

5.3 Written part

The written parts of both the BNC and the COCA differ notably in the occurrence of *start/begin* constructions. While in the spoken part *start + V ing* was the dominant construction in both corpora, in the written part the most frequent one was *begin +V infinitive.*

	BNC	COCA
start + V ing	4,478	24,632
start + V infinitive	5,311	15,597
begin + V ing	2,727	16,392
begin + V infinitive	17,756	34,402

Table 12 Number of tokens (lexical verbs) in the written parts of the COCA and the BNC

As follows from Table 12 the construction *begin + V* infinitive sticks out significantly in the BNC while the results in the COCA are more balanced.

The rate between *start + V ing* and *begin + V ing* was approximately 16:10 in the BNC and 15:10 in the COCA. The form with *start* is still more frequent but the rate is not in favor of *start* as strongly as it was in the spoken part (46:1 in the BNC and 5:1 in the COCA).

The rate between *start + V infinitive* and *begin + V infinitive* was 33:10 in the BNC and 22:10 in the COCA. The results in the BNC changed in favor of *begin* (in the spoken part the rate was 14:10 in favor of *start*).

As for the competition between the *-ing* form and the infinitive, like in the spoken part, the results in both corpora differ considerably. The rate between *start + V ing* and *start + V infinitive* is 10:12 in the BNC, but 16:10 in the COCA.

Therefore, it can be concluded that in written language the infinitive after *start* is more frequent in British English and the *ing* form after *start* is more frequent in American English. In the case of *begin* the prevalence of *begin + V infinitive* appeared in both corpora, but again more significantly in the BNC (approx. 7:1 in favor of the infinitive in the BNC and 2:1 in favor of the infinitive in the COCA).

Table 13 shows the representation of *have, do* and *be* in the written part of the BNC and the COCA:

	BNC written				COCA written			
	start+ V ing	start + V infinitive	begin + V ing	begin + V infinitive	start+ V ing	start + V infinitive	begin + V ing	begin + V infinitive
have	33	31	6	150	278	132	157	392
do	72	44	6	40	600	142	163	244
be	33	54	4	643	137	214	21	1,159

Table 13 Number of tokens (have, do and be) in the written parts of the BNC and the COCA

In both corpora the construction *begin + V infinitive* was the most frequent for *have* (150 tokens in the BNC and 392 tokens in the COCA) and *be* (643 tokens in the BNC and 1,159 tokens in the COCA). For *do* the construction with the highest occurrence was *start + V ing*. The results changed in case of *have* (the most frequent construction in the spoken part was *start + V ing*).

As was mentioned in Chapter 2, the written domain of the BNC consists of imaginative and informative prose (for more information see page 2).

Although the domain with the biggest number of words is the world affairs domain, it follows from Figure 9 that the most frequent one for all the constructions was imaginative domain.

Written domain BNC

	start + V ing	start + V infinitive	begin + V ing	begin + V infinitive
■ **Imaginative**	1433	2042	870	7202
■ **Natural/Pure science**	93	210	36	461
■ **Applied science**	302	391	247	800
■ **Social science**	459	393	212	1955
■ **World affairs**	450	512	605	2900
■ **Commerce/Finance**	271	357	138	797
■ **Arts**	512	334	217	1226
■ **Belief/Thought**	125	168	35	770
■ **Leisure**	833	904	367	1645
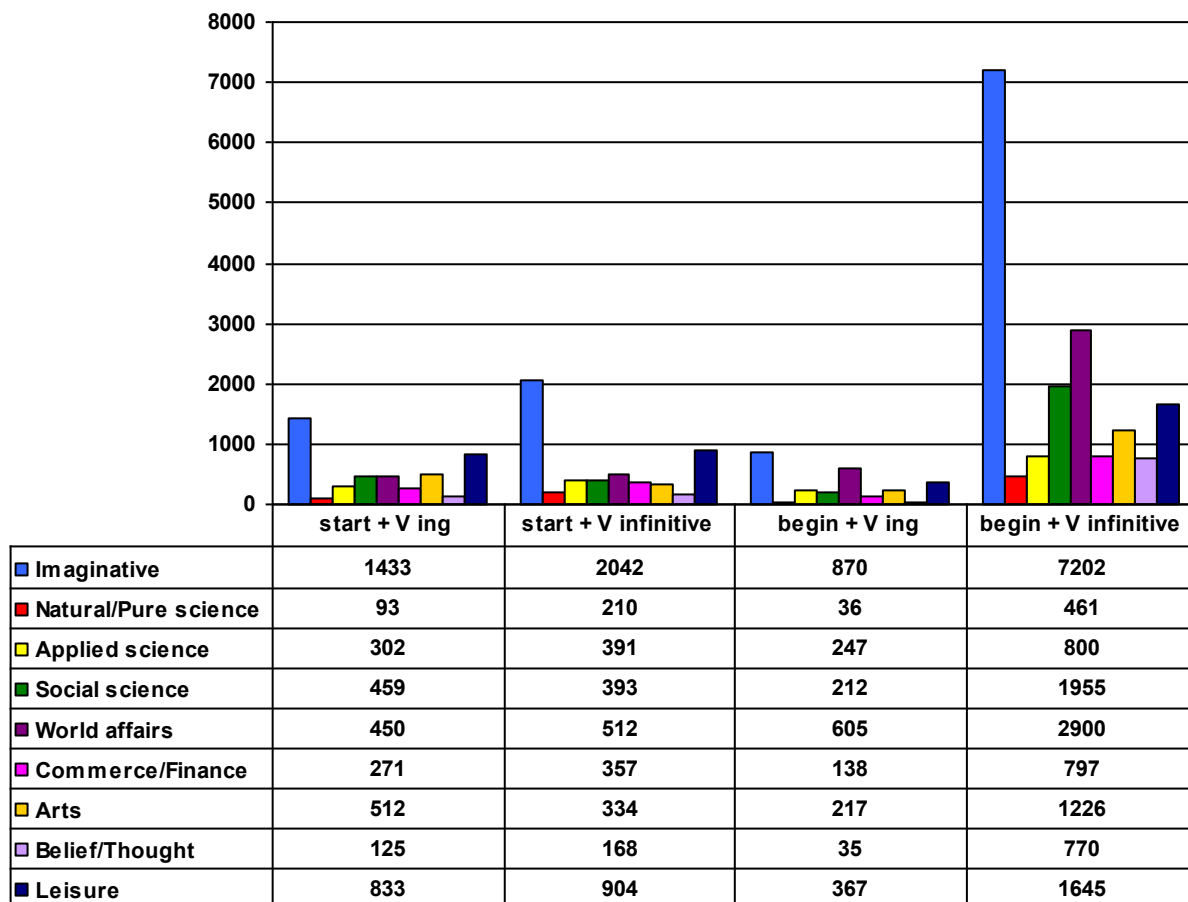
Fig. 9 Graph of written domain in the BNC

Imaginative domain dominates in all four cases: for *start + V ing* it is approx. 32%, for *start + V infinitive* it is approx. 38%, for *begin + V ing* approx. 32 % and for *begin + V infinitive* approx. 41%.

Unfortunately the COCA is composed of different parts than the BNC, so any comparison is impossible again.

Figure 10 shows the representation of the *start/begin* constructions in the four types of written texts in the COCA:

Written domain COCA

|	start + V ing	start + V infinitive	begin + V ing	begin + V infinitive
Fiction	9549	8467	3623	13284
Magazine	7472	3355	5401	8206
Newspaper	6273	2854	5206	4762
Academic	1338	921	2162	615
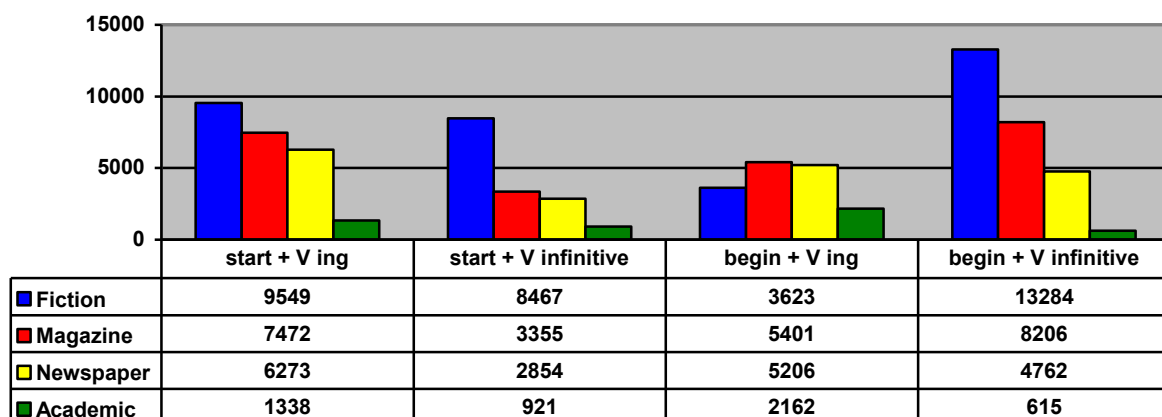
Fig. 10 Graph of the written domain in the COCA

Fiction is the most frequent domain in case of *start + V ing* (approx. 39%)*, start + V infinitive* (approx. 54%) and *begin + V infinitive* (approx. 39%). The most frequent domain for *begin + V ing* is magazine (approx. 33%).

5.4 Activity and state verbs after start/begin construction in the COCA and the BNC

The following tables present the occurrence of activity and state verbs after *start* and *begin* in both the COCA and the BNC, according to frequency. Perhaps it is to be expected that state verbs will be used in the infinitive form, but not necessarily all infinitives will be infinitives of state verbs. For better orientation, activity verbs are marked in red and state verbs in blue.

Spoken part

COCA Spoken	start + V ing	start + V infinitive	begin + V ing	begin + V infinitive
1.	talk (895)	get (578)	take (96)	see (442)
2.	get (626)	see (375)	talk (93)	get (321)
3.	do (573)	come (208)	work (91)	think (268)
4.	look (466)	look (197)	look (86)	look (263)
5.	go (405)	think (188)	make (79)	be (248)
6.	take (377)	feel (178)	write (54)	take (192)
7.	work (366)	go (171)	do (51)	feel (176)
8.	think (356)	be (156)	use (51)	realize (168)
9.	make (286)	talk (149)	arrive (46)	understand (160)
10.	come (276)	do (144)	move (44)	make (148)
11.	play (263)	make (141)	ask (39)	change (140)
12.	write (206)	take (128)	have (35)	talk (136)

13.	see (204)	say (126)	run (33)	come (134)
14.	have (188)	move (116)	come (32)	have (132)
15.	say (188)	have (115)	investigate(32)	do (124)
16.	run (181)	cry (95)	withdraw (32)	wonder (118)
17.	call (180)	work (89)	get (31)	move (113)
18.	ask (172)	realize (80)	serve (31)	go (99)
19.	move (171)	turn (73)	think (29)	emerge (87)
20.	use (166)	happen (71)	play (28)	hear (86)
21.	tell (137)	lose (71)	search (27)	show (85)
22.	put (132)	change (70)	tell (27)	turn (81)
23.	sing (121)	run (68)	call (26)	work (80)
24.	read (120)	become (64)	sell (26)	say (75)
25.	hear (115)	play (64)	plan (24)	lose (72)
26.	scream (112)	fall (59)	shoot (24)	develop (67)
27.	shoot (109)	put (57)	speak (24)	question (65)
28.	pay (107)	show (51)	show (23)	try (65)
29.	be (90)	build (50)	put (22)	focus (61)
30.	give (86)	write (45)	pay (21)	put (59)
31.	feel (82)	pick (44)	try (21)	ask (58)
32.	try (82)	hear (43)	go (20)	find (55)
33.	walk (78)	pay (43)	prepare (18)	tell (54)
34.	cry (76)	walk (42)	sing (20)	fall (53)
35.	yell (70)	worry (38)	test (20)	happen (52)
36.	hit (68)	believe (38)	build (19)	unravel (50)
37.	fall(66)	understand (38)	hear (19)	break (48)
38.	pull(66)	wonder (38)	date (18)	run (48)
39.	sell(66)	find (36)	fall (18)	build(45)
40.	show (66)	develop (35)	see (18)	grow (45)
41.	throw (66)	use (34)	study (18)	believe (44)
42.	laugh (65)	grow (33)	debate (17)	worry (43)
43.	turn (62)	tell (33)	record (17)	use (42)
44.	date(61)	ask (32)	air (15)	fight (41)
45.	speak (59)	break (30)	deliberate (15)	fade (40)
46.	listen (57)	pull (30)	appear (14)	learn (40)
47.	bring (56)	rise (29)	kill (14)	pay (40)
48.	watch (56)	call (28)	pull (14)	pull (39)
49.	act (55)	notice (28)	teach (14)	become (38)
50.	drink (55)	open (27)	focus (13)	cry (38)

Table 14 Occurrence of activity and state verbs after start and begin in the spoken part of the COCA

BNC Spoken	start + V ing	start + V infinitive	begin + V ing	begin + V infinitive
1.	do(10)	get (72)	work (4)	get (36)
2.	get(99)	go (34)	lay (3)	look (25)
3.	talk(87)	come (28)	make (2)	be (23)
4.	go(86)	look (25)	arrive (2)	see (19)
5.	think(62)	do (19)	clear (2)	feel (12)

6.	look(49)	make (17)	do (2)	wonder (12)
7.	make(42)	think ((17)	talk (2)	make (10)
8.	come (38)	feel (16)	set (1)	realize (10)
9.	work (36)	take (15)	send (1)	think (10)
10.	laugh (34)	put (14)	search (1)	become (8)
11.	write (30)	work (14)	say (1)	ease (8)
12.	pay (27)	talk (13)	ring (1)	take (8)
13.	say (26)	be (8)	re-rehearse (1)	go (7)
14.	play (24)	build (8)	recover (1)	happen (7)
15.	use (24)	grow (8)	raise (1)	pick (7)
16.	take (23)	move (8)	pull (1)	put (7)
17.	put (22)	run (8)	nag (1)	understand (7)
18.	give (20)	speak (8)	get (1)	grow (6)
19.	have (18)	use (8)	flutter (1)	find (6)
20.	try (17)	change (6)	direct (1)	do (6)
21.	cry (16)	say (6)	carry (1)	lose (6)
22.	build (16)	see (6)	breach (1)	run (6)
23.	buy (15)	ask (5)	beat (1)	change (5)
24.	run (15)	become (5)		realise (5)
25.	be (14)	fight (5)		show (5)
26.	eat (13)	have (5)		build (4)
27.	cut (12)	laugh (5)		address (4)
28.	sell (12)	pay (5)		come (4)
29.	walk (12)	walk (5)		drop (4)
30.	record (11)	develop (4)		say (4)
31.	sing (11)	fall (4)		smell (4)
32.	ask (10)	overtake (4)		fray (3)
33.	lose (10)	read (4)		help (3)
34.	smoke (10)	ring (4)		have (3)
35.	collect (9)	tell (4)		experience (3)
36.	grow (9)	write (4)		dissolve (3)
37.	speak (9)	drift (3)		discover (3)
38.	throw (9)	ease (3)		open (3)
39.	dance (9)	draw (3)		question (3)
40.	dig (9)	cry (3)		tackle (3)
41.	drink (8)	cut (3)		talk (3)
42.	learn (8)	deteriorate (3)		work (3)
43.	move (8)	buy (3)		achieve (2)
44.	pull (8)	accept (3)		bite (2)
45.	shout (8)	give (3)		believe (2)
46.	bring (7)	happen (3)		bring (2)
47.	plan (7)	learn (3)		challenge (2)
48.	read (7)	lose (3)		clear (2)
49.	ring (7)	pack (3)		annoy (2)
50.	show (7)	play (3)		calm (2)

Table 15 Occurrence of activity and state verbs after start and begin in the spoken part of the BNC

It follows from Table 14 and Table 15 that the occurrence of state verbs in the spoken parts of the COCA and the BNC is very similar. In the COCA 6 state verbs appeared in the *-ing* form after *start* and 15 state verbs in the infinitive form. In the BNC state verbs occur with similar frequency – 4 state verbs in the *–ing* form and 8 state verbs in the infinitive form. The reason why state verbs appear also in the *-ing* form might be the fact that some of them can function as both active and state verbs (e.g. feel, be, have) In the case of *begin*, the infinitive construction is also more frequent for state verbs in both corpora (6 state verbs in the –ing form and 14 state verbs in the infinitive in the COCA; no state verbs in the –ing form and 15 state verbs in the infinitive form in the BNC).

Written part

COCA Written	start + V ing	start + V infinitive	begin + V ing	begin + V infinitive
1.	talk (1230)	get (1101)	work (1022)	feel (1393)
2.	work (1069)	cry (803)	make (603)	see (1231)
3.	look (947)	feel (705)	use (554)	think (1186)
4.	get (902)	look (479)	take (532)	be (1159)
5.	think (834)	say (458)	talk (465)	take (1149)
6.	make (828)	walk(419)	look (415)	understand (895)
7.	take (700)	see (406)	write (383)	look (874)
8.	walk (693)	move (375)	sell (350)	make (807)
9.	play (688)	make (374)	walk (330)	move (795)
10.	go (624)	think (372)	move (325)	cry (761)
11.	do (606)	go (361)	play (294)	get (736)
12.	use (526)	take (342)	offer (279)	change (704)
13.	run (498)	laugh (320)	call (243)	wonder (692)
14.	come (464)	turn (309)	teach (243)	fall (644)
15.	move (462)	come (298)	study (236)	realize (610)
16.	call (436)	run (276)	run (232)	appear (538)
17.	laugh (433)	fall (257)	read (221)	turn (497)
18.	write (359)	be (221)	collect (211)	talk (469)
19.	ask (336)	speak (221)	build (201)	emerge (463)
20.	scream (314)	rise (216)	think (199)	show (460)
21.	sing (302)	leave (205)	search (186)	grow (453)
22.	have (297)	change (201)	appear (180)	speak (442)
23.	see (273)	show (194)	arrive (180)	play (435)
24.	sell (270)	pull (189)	ask (179)	develop (420)
25.	feel (369)	work (189)	speak (170)	work (407)
26.	put (261)	lose (186)	plan (169)	rise (405)
27.	cry (258)	play (183)	do (162)	have (392)
28.	tell (234)	talk (178)	turn (162)	walk (380)

29.	read (224)	become (166)	have (157)	run (370)
30.	date (203)	grow (163)	produce (152)	read (365)
31.	pay (203)	wonder (156)	try (149)	lose (357)
32.	pull (201)	shake (152)	operate (145)	sing (348)
33.	give (200)	put (147)	investigate(141)	come (325)
34.	throw (197)	climb (145)	put (139)	form (301)
35.	eat (190)	do (145)	sing (139)	use (299)
36.	shoot (190)	ask (143)	pull (135)	write (290)
37.	yell (186)	rain (139)	prepare (130)	build (275)
38.	act (182)	realize (38)	receive (128)	notice (275)
39.	show (179)	sing (137)	experiment(126)	fade (274)
40.	collect (169)	have (133)	publish(125)	laugh (271)
41.	teach (163)	open (132)	show (123)	worry (271)
42.	build (156)	build (124)	get (120)	suspect (261)
43.	fall (156)	write (121)	explore (119)	break (257)
44.	buy (155)	read (115)	give (117)	fill (257)
45.	wear (154)	believe (30)	pay (117)	recognize (257)
46.	try (151)	tell (112)	buy (114)	tell (251)
47.	plan (147)	break (106)	see (114)	question (249)
48.	say (147)	appear (103)	test (114)	explore (247)
49.	turn (144)	fade (102)	tell (114)	do (244)
50.	hit (138)	give (100)	come (110)	go (242)

Table 16 Occurrence of activity and state verbs after start and begin in the written part of the COCA

BNC Written	start + V ing	start + V infinitive	begin + V ing	begin + V infinitive
1.	talk (134)	get (107)	ship (55)	feel (439)
2.	look (116)	feel (84)	talk (38)	be (343)
3.	make (108)	cry (76)	work (35)	think (275)
4.	think (94)	look (75)	make (34)	take (222)
5.	go (92)	go (73)	walk (32)	look (208)
6.	play (87)	take (69)	move (29)	move (204)
7.	get (78)	walk (69)	take (26)	get (200)
8.	work (76)	make (64)	write (25)	wonder (200)
9.	take (65)	run (56)	collect (19)	make (190)
10.	use (64)	move (52)	look (18)	see (171)
11.	cry (62)	think (50)	play (18)	walk (142)
12.	come (55)	come (48)	think (18)	cry (134)
13.	do (55)	laugh (43)	use (17)	understand (113)
14.	ship (51)	grow (37)	read (16)	appear (112)
15.	run (50)	play (37)	build (16)	talk (109)
16.	write (45)	say (37)	sell (15)	fall (106)
17.	ask (44)	be (36)	run (14)	have (103)
18.	walk (40)	turn (36)	arrive (14)	run (102)
19.	shout (37)	fall (34)	shout (14)	laugh (99)
20.	move (35)	work (34)	withdraw (14)	show (99)
21.	try (35)	appear (30)	plan (13)	speak (88)

22.	build (33)	read (30)	search (13)	emerge (87)
23.	call (32)	talk (30)	broadcast (12)	play (86)
24.	put (32)	build (29)	offer (12)	realise (81)
25.	be (30)	shake (29)	operate (12)	turn (80)
26.	have (28)	do (27)	pace (12)	grow (78)
27.	laugh (28)	lose (27)	try (12)	work (77)
28.	scream (27)	rain (26)	appear (11)	go (76)
29.	sell (27)	speak (26)	develop (11)	come (75)
30.	sing (27)	write (26)	investigate (11)	lose (72)
31.	pay (26)	show (25)	market (11)	climb (69)
32.	read (26)	rise (23)	trade (11)	read (69)
33.	collect (25)	use (23)	climb (10)	develop (66)
34.	eat (25)	become (22)	draw (10)	rise (66)
35.	tell (25)	have (22)	pull (10)	change (63)
36.	shoot (24)	break (20)	turn (10)	tell (63)
37.	train (24)	eat (20)	clear (9)	find (62)
38.	drink (22)	pick (20)	fill (9)	shake (62)
39.	plan (22)	pull (20)	fire (9)	sing (59)
40.	see (22)	sing (20)	leave (9)	believe (56)
41.	give (21)	climb (19)	prepare (9)	enjoy (54)
42.	buy (20)	give (18)	sort (9)	build (50)
43.	feel (20)	put (17)	speak (9)	fill (50)
44.	pull (20)	roll (17)	come (8)	pull (48)
45.	throw (20)	see (17)	cut (8)	form (47)
46.	fight (19)	wonder (17)	gather (8)	realize (46)
47.	say (19)	worry (17)	publish (8)	tremble (46)
48.	dig (18)	develop (16)	put (8)	sound (44)
49.	fall (17)	tell (16)	ring (8)	suspect (44)
50.	smoke (17)	believe (15)	set (8)	put (43)

Table 17 Occurrence of activity and state verbs after start and begin in the written part of the BNC

As follows from Table 16 and Table 17, in the written part the results are similar for both corpora. In both the COCA and the BNC the infinitive form is more frequent for state verbs. After *start* 5 state verbs appeared in the *–ing* form and 11 state verbs in the infinitive form in the COCA; 6 state verbs occurred in the –ing form and 11 state verbs in the infinitive form in the BNC. After *begin* 5 state verbs appeared in the *–ing* form and 12 state verbs in the infinitive form in the COCA; 3 state verbs occurred in the –ing form and 14 state verbs in the infinitive form in the BNC.

In both the COCA and the BNC the infinitive is used with state verbs more often than with the *–ing* form. Unlike the *–ing* form, the infinitive is used for both

activity and state verbs. Therefore it can be concluded from this analysis that the *-ing* form is marked and the infinitive unmarked.

6. SUMMARY

This thesis analyses the competition of *start* and *begin* in British and American English, more specifically the frequency of the infinitive and *–ing* forms after these verbs. This largely quantitative analysis is based on the data retrieved from two big corpora – British National Corpus (BNC) and Corpus of Contemporary American English (COCA).

The thesis is divided in a theoretical and a practical part. The theoretical part provides information about the structure of both corpora, necessary for later analyses of the data. It also presents the types of monotransitive complementation and the description of the *–ing* form and the infinitive.

In the practical part I strictly distinguished between the spoken and written language in order to show the different tendencies in British and American English.

It follows from the data retrieved from the corpora that in the spoken language the dominant structure is *start + V ing*. As for the competition between the *–ing* form and the infinitive, the rate between *start + V ing* and *start + V infinitive* was approximately the same in both corpora. In the case of *begin + V ing* and *begin + V infinitive,* in both corpora a prevalence of *begin + V infinitive* appeared, but it was more apparent in the BNC. Sociolinuistic variables such as the social background and the age of speakers do not seem to have any influence on the distribution of infinitives and *–ing* forms in the competition of *begin* and *start*; the dominance of *start + V ing* is reflected even in each of the studied sections of the demographic part i.e. for speakers of all ages and from all social backgrounds the most frequent construction was *start + V ing.* The construction with the highest number of tokens for every area of context-governed texts was *start + V infinitive.* Unfortunately COCA does not allow specific searches within the spoken part, therefore, no further analysis is possible.

The data in the written part differ significantly from the data obtained in the spoken part. The construction with the highest number of tokens was *begin + V infinitive*. As for the competition between the *–ing* form and the infinitive, the corpora differ considerably. In the BNC *start + V infinitive* is more frequent then *start + V ing*, while in the COCA it is the other way round. Like in the spoken part, the prevalence of *begin + V infinitive* occurred in both corpora, being more notable in the BNC. The dominance of *begin + V infinitive* is reflected in every field of the

written domain. Although COCA offers a more detailed research within its written part, a proper comparison is not possible because of the different structures of the corpora. In the COCA the construction *begin + V infinitive* is the most frequent one in fiction and magazine text, while *begin + V ing* is the most frequent one in the academic texts and *start + V ing* in the newspaper texts.

The last subchapter presents the frequency of activity and state verbs after *start* and *begin.* In both the written and the spoken parts of the COCA and the BNC, state verbs appeared in the infinitive form more often than in the *–ing* form. It follows from the data retrieved from both corpora that unlike the *–ing* form, the infinitive is used for both activity and state verbs. Therefore, it can be concluded that the *-ing* form is marked while the infinitive is unmarked.

RESUMÉ

Tato bakalářská práce se zabývá infinitivem a *–ing* formou v doplnění fázových sloves, *start* a *begin*. Slovesné doplnění je v současné době jedním z nejstudovanějších a nejsložitějších témat anglické gramatiky. Cílem práce je kvantitativní analýza korpusových dat s důrazem na rozdíly v britské a americké angličtině. Tato analýza je založená na datech, která byla čerpána ze dvou největších současných korpusů – Corpus of Contemporary American English (COCA) a British National Corpus (BNC - XML). Na výzkumy v korpusech je dnešní době kladem velký důraz, poněvadž umožňují jazykovědcům ověření lingvistických teorií na praktickém jazyce. Popis korpusů, jejich struktury a jejich porovnání je zahrnuto do druhé kapitoly. Znalost struktury korpusů je nezbytně nutná pro pozdější přesnou analýzu dat. Každý z korpusů má své výhody a nevýhody. COCA je narozdíl od BNC aktualizován každý rok, takže pravděpodobně představuje nejlepší zdroj pro studium současné americké angličtiny. BNC tuto možnost nenabízí, ale jeho výhoda je v tom, že na rozdíl od COCA jeho mluvená část obsahuje spontánní dialogy.

Třetí kapitola je zaměřena na samotné slovesné doplnění a vychází z odborné literatury. Lingvisté mají různý názor na typy slovesného doplnění, v této práci je upřednostňováno Quirkovo členění.

Čtvrtá kapitola se zabývá neurčitými slovesnými tvary – gerundiem a infinitivem a krátce též porovnáním jejich společných znaků. Dále se zabývá funkcí *–ing* formy a infinitivu po frázových slovesech *start* a *begin*.

Poslední kapitola je už zaměřená na samotnou kvantitativní analýzu dat získaných z korpusů. Největší nevýhodou těchto korpusů je jejich rozdílná velikost a odlišná struktura, která v některých případem znemožňuje jejich porovnání. Kapitola zkoumá odděleně psaný a mluvený jazyk.

Z dat získaných z korpusu vyplývá, že pro mluvený jazyk je nejčastější struktura *start + V ing*. Co se týče výskytu gerundia a infinitivu, poměr mezi *start + V ing* a *start + V infinitive* byl zhruba stejný v obou korpusech (2:1 ve prospěch *start + V infinitive*). V případě *begin* , převládá struktura *begin + V infinitive*, přičemž tento rozdíl v použití *begin + V ing* and *begin + V infinitive* je více patrný v britské angličtině (v BNC 14:1 ve prospěch *begin + V infinitive*; v COCA 3:1 ve prospěch *begin + V infinitive*). Jak již bylo řečeno BNC umožňuje sociolingvistickou analýzu korpusových dat. Tato analýza ukazuje, že ani povolání ani věk mluvčích

neovlivňuje distribuci infinitivu a –*ing* formy po fázových slovesech *start* a *begin*. Převaha *start + V ing* se objevuje v obou analyzovaných částech demografické sekce (v těchto spontánních rozhovorech je pro všechny mluvčí nehledě na věk ani zaměstnání nejčastější konstrukcí *start + V* ing). Nejčastější konstrukcí pro každou ze sekcí kontextově řízených textů je *start + V infinitive*. Bohužel COCA neumožňuje detailnější analýzu mluvené části, a proto nelze srovnat výskyt –*ing* formy a infinitivu v těchto dvou specifických oblastech mluvené části BNC.

Data v psané částí se výrazně liší od dat získaných v mluvené části. Nejčastější konstrukcí pro tuto část je *begin + V infinitive*. Co se týče výskytu –*ing* formy a infinitivu v jednotlivých konstrukcích, korpusy se značně rozcházejí. V britské angličtině se *start + V infinitive* vyskytuje častěji než *start + V ing* (poměr 10:12 ve prospěch *start + V infinitive*), zatímco v americké angličtině je tomu přesně naopak (16:10 ve prospěch *start + V ing*). Stejně jako v mluvené části, se převaha *begin + V infinitive* projevila v obou korpusech, výrazněji v BNC. Převaha *begin + V infinitive* se objevuje ve všech sekcích psané části. Ačkoliv COCA, na rozdíl od mluvené části, umožňuje detailnější analýzu zkoumaných textů, porovnání opět není možné z důvodu odlišné struktury obou korpusů.

V COCA je nejčastější konstrukcí pro beletrii a časopiseckou sekci *begin + V infinitive,* zatímco pro akademické texty je nejčastější *begin + V ing* a pro novinové texty je to *start + V infinitive*.

Poslední podkapitola prezentuje četnost výskytu infinitivních forem a –*ing* forem stavových a dějových sloves po fázových slovesech *start* a *begin*. Podkapitola je opět rozdělena na psaný a mluvený jazyk. V korpusech se neprojevil žádný výrazný rozdíl mezi těmito dvěma částmi. Jak v psané, tak v mluvené části se stavová slovesa objevují častěji ve formě infinitivu než ve –*ing* formě. Z korpusových dat vyplývá, že –*ing* forma je ve většině případů použita u dějových sloves, zatímco infinitiv je běžný jak pro stavová tak pro dějová slovesa. *-Ing* forma se jeví jako tvar příznakový a infinitiv jako nepříznakový.

ANOTACE

Jméno a příjmení: Kateřina Dubská

Katedra: Katedra anglistiky a amerikanistiky

Vedoucí práce: Mgr. Michaela Martinková, PhD.

Rok obhajoby: 2010

Název práce: Doplnění slovesa infinitivem a *-ing* formou v britské a americké angličtině: studie založená na datech z korpusu

Název práce v angličtině: Verb complementation by infinitive and *-ing* forms in British and American English: A corpus based study

Anotace práce: Tato bakalářská práce se zabývá infinitivem a *–ing* formou v doplnění fázových sloves, *start* a *begin*. Cílem práce je kvantitativní analýza korpusových dat s důrazem na rozdíly v britské a americké angličtině.

Klíčová slova: -*ing* form, infinitiv, doplnění slovesa, korpus, fázová slovesa, *start, begin*

Anotace v angličtině: This thesis focuses on the competition of *start* and *begin* in British and American English, more specifically the frequency of the infinitive and *–ing* forms after these verbs. The aim of this thesis is quantitative analyses of the corpus data with emphasis on the difference between British and American English.

Klíčová slova v angličtině: -ing form, infinitive, verb complementation, corpus, aspectual verbs, *start, begin*

Přílohy vázané v práci: Přílohy obsahují přehled sloves po všech tvarech *begin* a *start* a tabulky stavových a dějových sloves vyskytujících se po *start* a *begin* (psaný y mluvený jazyk dohromady)

Rozsah práce: 43 stran + přílohy

Jazyk práce: Anglický jazyk

BIBLIOGRAPHY

Aarts, Bas. *English Syntax and Argumantation.* 2nd ed. Houndmills: Palgrave Macmillan, 2001

Brinton, Laurel J. *The Development of English Aspectual Systems.* Cambridge: Cambridge University Press, 1988.

Dušková, Libuše. *Mluvnice současné angličtiny na pozadí češtiny*. Praha: Academia, 1994.

Huddleston, Rodney, and Geoffrey K. Pullum. *The Cambridge Grammar of English Language.* Cambridge: Cambridge University Press, 2002.

Jespersen, Otto. *A Modern English Grammar*. Vol. 4. London: Allen and Unwin, 1946.

Kennedy, Graeme. *An Introduction to Corpus Linguistics*. London: Longman, 1998.

Kruisinga, E. *A Handbook of Present-day English*. Groningen: P. Noordhoff, 1921.

Leech, Geoffrey. *Meaning and the English Verb*. London: Longman, 2004.

Macpherson, Robin. *Základy anglické stylistiky*. Praha: Academia, 1997.

Mair, Christian. *Infinitival Complement Clauses in English. A study of syntax in discourse.* Cambridge:University Press, 1990.

Matthews, P.H. *Syntax*. New York: Cambridge University Press, 1992.

Swan, Michael. *Practical English Usage*. Third Edition. Oxford: Oxford University Press, 2005.

Palmer, F.R. *The English Verb*. London: Longman, 1997.

Poldauf, Ivan. "O konkurenci infinitivu a gerundu v angličtině." *Časopis pro moderní filologii 37,* 1955.

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech, and Jan Svartvik. *A Comprehensive Grammar of the English Language*. London: Longman, 1985.

Schibsbye, Knud. *A modern English Grammar.* Second edition. London: Oxford University Press, 1970.

Veselovská, Ludmila. *English Syntax: Syllabi, Examples and Exercises*. Olomouc: Univerzita Palackého, 2006.

Zandvoort, R.W. *A Handbook of English Grammar*. London: Longmans, 1969.

Online sources:

http://www.natcorp.ox.ac.uk/using/index.xml

http://www.natcorp.ox.ac.uk/docs/URG/BNCdes.html#BNCcompo

http://www.natcorp.ox.ac.uk/docs/URG/posguide.html#errorRates

http://www.natcorp.ox.ac.uk/using/index.xml

http://www.tei-c.org/Guidelines/

http://www.americancorpus.org/

http://www.microsoft.com/sqlserver/2008/en/us/product-information.aspx

Corpora used:

British National Corpus (BNC-XML)

Corpus of Contemporary American English (COCA)

APPENDICES

Appendix 1

Forms of *start* in the COCA

Gerund:

start – 1.talk(864) 2.look(749) 3.think(686) 4.get(614) 5.make(494) 6.take(448) 7.work(444) 8.go(274) 9.play(266) 10.use(261) 11.move(252) 12.come(210) 13.see(209) 14.pay(193) 15.put(188) 16.ask(177) 17.feel(159) 18.run(156) 19.call(153) 20.walk(139) 21.tell(127) 22.say(129) 23.build(121) 24.shoot(118) 25.write(117) 26.give(116) 27.plan(112) 28.eat(109) 29.sell(109) 30.act(106)

starts – 1.talk(192) 2.walk(143) 3.laugh(114) 4.move(105) 5.get(103) 6.go(95) 7.run(82) 8.look(80) 9.play(74) 10.scream(70) 11.make(68) 12.sing(68) 13.work(64) 14.pull(57) 15.take(56) 16.come(46) 17.cry(42) 18.shake(40) 19.put(39) 20.tell(36) 21.ask(33) 22.shoot(32) 23.hit(31) 24.hit(31) 25.yell(31) 26.fall(28) 27.spin(28) 28.beat(27) 29.say(26) 30.pick(25)

started – 1.talk(1043) 2.work(917) 3.get(798) 4.go(647) 5.play(609) 6.look(570) 7.take(560) 8.make(532) 9.walk(487) 10.come(471) 11.think(470) 12.call(438) 13.run(432) 14.write(432) 15.use(401) 16.ask(296) 17.laugh(283) 18.move(270) 19.see(255) 20.scream(254) 21.sing(253) 22.read(244) 23.cry(240) 24.date(218) 25.sell(216) 26.tell(205) 27.feel(176) 28.say(176) 29.yell(176) 30.put(162)

starting – 1.talk(6) 2.go(5) 3.say(4) 4.call(3) 5.work(3) 6.use(3) 7.yell(2) 8.get(2) 9.make(2) 10.look(2) 11.come(2) 12.chant(2) 13.move(2) 14.take(1) 15.spill(1) 16.show(1) 17.mash(1) 18.run(1) 19.roll(1) 20.read(1) 21.pull(1) 22.pitch(1) 23.neglect(1) 24.murmur(1) 25.lick(1) 26.invite(1) 27.investigate(1) 28.inch(1) 29.hike(1) 30.hang(1)

Infinitive:

start – 1.get(301) 2.see(271) 3.feel(248) 4.think(188) 5.look(126) 6.make(108) 7.move(92) 8.cry(89) 9.go(88) 10.take(87) 11.come(83) 12.talk(76) 13.lose(69) 14.fall(68) 15.say(66) 16.turn(63) 17.work(54) 18.wonder(53) 19.play(52) 20.become(49) 21.grow(46) 22.run(45) 23.believe(42) 24.change(42) 25.walk(42) 26.build(41) 27.finish(41) 28.laugh(40) 29.realize(40) 30.rise(39)

starts – 1.cry(229) 2.get(205) 3.walk(179) 4.move(147) 5.go(105) 6.laugh(89) 7.leave(83) 8.run(83) 9.rise(73) 10.take(72) 11.look(71) 12.say(71) 13.pull(69) 14.turn(63) 15.climb(61) 16.fall(58) 17.come(56) 18.feel(54) 19.open(53) 20.sing(49) 21.play(46) 22.speak(46) 23.shake(45) 24.make(44) 25.fade(42) 26.talk(36) 27.put(33) 28.rain(33) 29.follow(32) 30.roll(30)

started - 1.cry(531) 2.get(509) 3.say(378) 4.feel(257) 5.go(225) 6.walk(205) 7.make(200) 8.laugh(198) 9.think(197) 10.come(182) 11.take(169) 12.turn(168) 13.run(166) 14.speak(164) 15.look(161) 16.see(158) 17.move(149) 18.talk(143) 19.fall(139) 20.change(125) 21.work(124) 22.play(106) 23.write(102) 24.pull(98) 25.put(97) 26.become(93) 27.read(93) 28.rise(93) 29.ask(92) 30.shake(89)

starting – 1.get(661) 2.see(339) 3.feel(321) 4.look(309) 5.come(179) 6.think(164) 7.make(155) 8.take(138) 9.show(126) 10.go(113) 11.move(96) 12.realize(92) 13.sound(84) 14.turn(82) 15.lose(81) 16.change(79) 17.talk(72) 18.wonder(72) 19.say(68) 20.work(68) 21.believe(64) 22.become(63) 23.understand(61) 24.happen(56) 25.build(51) 26.fall(49) 27.catch(47) 28.cry(47) 29.pay(47) 30.grow(45)

Forms of *start* in the BNC

Gerund:

start – 1.think (159) 2.talk(122) 3.look(117) 4.get(110) 5.make(88) 6.use(65) 7.go(64) 8.work(52) 9.play(46) 10.take(45) 11.come(43) 12.pay(39) 13.write(39) 14.plan(38) 15.ask(36) 16.build(36) 17.move(36) 18.put(33) 19.cry(31) 20.ship(31) 21.say(30) 22.try(30) 23.give(27) 24.give(26) 25.buy(26) 26.run(24) 27.shoot(21) 28.call(20) 29.shout(20) 30.walk(20)

starts – 1.go(20) 2.talk(18) 3.get(16) 4.look(11) 5.come(10) 6.make(9) 7.play(9) 8.run(8) 9.cry(7) 10.speak(7) 11.ship(6) 12.try(6) 13.ask(5) 14.beat(4) 15.dance(4) 16.neck(4) 17.rain(4) 18.roll(4) 19.shoot(4) 20.shout(4) 21.sing(4) 22.take(4) 23.think(4) 24.work(4) 25.give(3) 26.feel(3) 27.drip(3) 28.dig(3) 29.clean(3) 30.break(3)

started – 1.go(120) 2.talk(116) 3.work(94) 4.make(86) 5.play(76) 6.get(75) 7.look(64) 8.take(62) 9.come(60) 10.laugh(56) 11.write(56) 12.run(46) 13.cry(46) 14.walk(39) 15.use(38) 16.shout(35) 17.think(33) 18.put(29) 19.sing(28) 20.sell(27) 21.sell(25) 22.ask(24) 23.try(24) 24.drink(23) 25.scream(22) 26.call(21) 27.build(21) 28.smoke(21) 29.read(20) 30.ring(20)

starting – 1.make(2) 2.smoke(2) 3.shoot(1) 4.return(1) 5.put(1) 6.print(1) 7.pay(1) 8.zander(1) 9.yak(1) 10.work(1) 11.use(1) 12.train(1) 13.torment(1) 14.take(1) 15.swim(1) 16.stop(1) 17.start(1) 18.sport(1) 19.sob(1) 20.keep(1) 21.import(1) 22.hit(1) 23.grow(1) 24.go(1) 25.flash(1) 26.contract(1)

Infinitive:

start – 1.get(40) 2.look(36) 3.feel(32) 4.think(31) 5.go(28) 6.make(25) 7.work(21) 8.fall(20) 9.take(20) 10.build(19) 11.see(18) 12.run(16) 13.use(16) 14.lose(15) 15.move(14) 16.come(12) 17.grow(11) 18.put(10) 19.talk(10) 20.arrive(9) 21.break(9) 22.change(9) 23.cry(9) 24.develop(8) 25.walk(8) 26.climb(7) 27.appear(7) 28.improve(7) 29.learn(7) 30.live(7)

starts – 1.get(26) 2.fall(14) 3.look(13) 4.take(13) 5.become(12) 6.go(12) 7.move(12) 8.work(10) 9.build(9) 10.make(9) 11.rise(9) 12.break(8) 13.come(8) 14.grow(7) 15.run(7) 16.climb(6) 17.show(6) 18.speak(6) 19.turn(6) 20.change(5) 21.appear(5) 22.behave(5) 23.cry(5) 24.feel(5) 25.give(5) 26.lose(5) 27.play(5) 28.shake(5) 29.emerge(4) 30.breathe(3)

started – 1.get(99) 2.make(78) 3.go(77) 4.walk(68) 5.cry(65) 6.take(57) 7.laugh(52) 8.run(51) 9.feel(50) 10.come(49) 11.look(48) 12.move(48) 13.work(48) 14.say(45) 15.play(38) 16.write(31) 17.read(30) 18.speak(30) 19.talk(30) 20.think(30) 21.use(29) 22.rain(27) 23.turn(24) 24.pull(23) 25.build(22) 26.shake(22) 27.sing(22) 28.appear(21) 29.eat(21) 30.grow(21)

starting – 1.get(57) 2.feel(36) 3.look(30) 4.come(22) 5.take(21) 6.make(19) 7.show(19) 8.go(15) 9.think(15) 10.grow(13) 11.talk(11) 12.turn(11) 13.appear(10) 14.become(10) 15.pick(10) 16.emerge(9) 17.sound(9) 18.use(9) 19.believe(8) 20.build(8) 21.fall(7) 22.move(7) 23.rain(7) 24.walk(7) 25.work(7) 26.worry(7) 27.write(7) 28.cry(6) 29.eat(6) 30.flow(6)

Forms of *begin* in the COCA:

Gerund:

begin – 1.work(152) 2.take(128) 3.make(120) 4.look(106) 5.use(99) 6.sell(89) 7.talk(82) 8.move(75) 9.build(70) 10.think(59) 11.pay(57) 12.plan(56) 13.play(52)

14.test(52) 15.teach(48) 16.write(48) 17.arrive(46) 18.prepare(42) 19.withdraw(42) 20.collect(41) 21.serve(41) 22.produce(38) 23.offer(36) 24.run(33) 25.operate(32) 26.walk(32) 27.appear(31) 28.ask(31) 29.explore(30) 30.get(30)

begins – 1.walk(52) 2.talk(48) 3.move(36) 4.sing(35) 5.work(32) 6.play(31) 7.look(29) 8.pull(22) 9.search(21) 10.make(20) 11.speak(20) 12.take(20) 13.fire(18) 14.put(17) 15.read(15) 16.shake(15) 17.write(15) 18.back(13) 19.eat(13) 20.prepare(13) 21.roll(13) 22.run(13) 23.shoot(13) 24.climb(12) 25.draw(12) 26.pace(12) 27.serve(12) 28.cut(11) 29.fill(11) 30.hum(11)

beginning – 1.throttle (1) 2.smoke(1) 3.refer(1) 4.operate(1) 5.disconnect(1)

began – 1.work (865) 2.make(479) 3.use(422) 4.take(418) 5.talk(391) 6.write(351) 7.look(322) 8.sell(253) 9.walk(251) 10.play(230) 11.move(223) 12.call(216) 13.study(204) 14.run(203) 15.offer(201) 16.teach(184) 17.ask(176) 18.read(176) 19.arrive(170) 20.collect(158) 21.search(155) 22.appear(153) 23.think(148) 24.investigate(137) 25.speak(132) 26.turn(132) 27.try(125) 28.build(122) 29.plan(120) 30.date(112)

begun – 1.use(84) 2.work(72) 3.make(65) 4.take(64) 5.offer(51) 6.look(44) 7.talk(39) 8.move(38) 9.sell(33) 10.write(28) 11.call(25) 12.build(23) 13.study(21) 14.try(21) 15.experiment(19) 16.teach(19) 17.explore(18) 18.put(18) 19.investigate(17) 20.think(17) 21.collect(16) 22.give(16) 23.run(16) 24.push(15) 25.send(15) 26.plan(14) 27. play(14) 28.speak(14) 29.implement(13) 30.operate(13)

Infinitive:

begin – 1.see(471) 2.understand(414) 3.think(341) 4.make(235) 5.feel(230) 6.look(218) 7.get(215) 8.take(198) 9.move(164) 10.wonder(135) 11.realize(129) 12.develop(120) 13.imagine(112) 14.change(111) 15.tell(107) 16.talk(103) 17.address(99) 18.describe(98) 19.fall(97) 20.work(96) 21.lose(95) 22.turn(94) 23.form(89) 24.appear(81) 25.grow(80) 26.rise(80) 27.hear(77) 28.learn(77) 29.show(74) 30.build(67)

begins – 1.cry(163) 2.move(138) 3.take(129) 4.look(128) 5.feel(102) 6.fall(101) 7.play(96) 8.sing(88) 9.rise(86) 10.speak(77) 11.turn(77) 12.make(70) 13.read(69) 14.get(69) 15.walk(68) 16.run(63) 17.grow(62) 18.pull(58) 19.see(55) 20.talk(55) 21.lose(54) 22.change(53) 23.think(52) 24.form(51) 25.emerge(48) 26.laugh(48) 27.set(48) 28.shake(43) 29.climb(42) 30.fade(42)

beginning – 1.see(443) 2.feel(366) 3.think(355) 4.get(333) 5.understand(313) 6.look(286) 7.take(228) 8.realize(198) 9.show(183) 10.wonder(182) 11.make(172) 12.emerge(143) 13.change(128) 14.come(123) 15.turn(115) 16.move(89) 17.find(78) 18.learn(78) 19.recognize(77) 20.believe(74) 21.worry(72) 22.develop(71) 23.sound(71) 24.fall(69) 25.lose(68) 26.appear(66) 27.fade(63) 28.suspect(61) 29.talk(61) 30.form(60)

began – 1.feel(755) 2.take(619) 3.see(578) 4.think(575) 5.cry(559) 6.move(459) 7.change(450) 8.look(419) 9.wonder(416) 10.get(398) 11.make(383) 12.fall(374) 13.realize(357) 14.talk(342) 15.appear(329) 16.speak(292) 17.play(266) 18.walk(263) 19.run(253) 20.grow(251) 21.walk(263) 22.read(247) 23.emerge(228) 24.turn(225) 25.write(224) 26.sing(222) 27.understand(215) 28.come(205) 29.rise(200) 30.develop(198)

begun – 1.take(170) 2.think(134) 3.see(130) 4.feel(122) 5.change(104) 6.make(99) 7.look(85) 8.show(78) 9.emerge(75) 10.understand(73) 11.realize(71) 12.turn(69) 13.explore(67) 14.appear(62) 15.move(61) 16.fall(59) 17.develop(58) 18.use(58) 19.question(57) 20.recognize(56) 21.grow(55) 22.wonder(54) 23.address(50) 24.work(47) 25.lose(45) 26.talk(45) 27.focus(44) 28.get(42) 29.speak(42) 30.fight(40)

Forms of *begin* in the BNC

Gerund:

begin – 1.ship(16) 2.make(10) 3.move(9) 4.work(9) 5.think(8) 6.read(7) 7.film(6) 8.draw(6) 9.sell(6) 10.train(6) 11.build(5) 12.plan(5) 13.test(5) 14.appear(4) 15.deliver(4) 16.export(4) 17.develop(4) 18.fish(4) 19.learn(4) 20.look(4) 21.market(4) 22.offer(4) 23.operate(4) 24.put(4) 25.talk(4) 26.teach(4) 27.trade(4) 28.withdraw(4) 29.draft(3) 30.evacuate(3)

begins – 1.ship(6) 2.listen(2) 3.count(2) 4.teach(2) 5.work(2) 6.write(1) 7.whiz(1¨) 8.whine(1) 9.try(1) 10.travel(1) 11.transmit(1) 12.tighten(1) 13.think(1) 14.take(1) 15.swell(1) 16.stand(1) 17.sob(1) 18.shovel(1) 19.brush(1) 20.bleed(1) 21.bellow(1) 22.beam(1) 23.interview(1) 24.interrogate(1) 25.hit(1) 26.heal(1) 27.go(1) 28.form(1) 29.follow(1) 30.flow(1)

beginning – 1.train(2) 2.read(1) 3.comission(1)

began – 1.work(48) 2.talk(43) 3.write(41) 4.make(35) 5.walk(35) 6.take(33) 7.move(24) 8.play(22) 9.look(21) 10.collect(19) 11.ship(19) 12.use(19) 13.arrive(18) 14.think(18) 15.build(17) 16.publish(17) 17.run(16) 18.shout(16) 19.try(15) 20.read(14) 21.operate(13) 22.search(13) 23.experiment(12) 24.speak(12) 25.trade(12) 26.clear(11) 27.broadcast(11) 28.climb(11) 29.fire(11) 30.pace(11)

begun – 1.ship(14) 2.work(11) 3.make(7) 4.market(7) 5.use(7) 6.investigate(5) 7.produce(4) 8.write(4) 9.read(3) 10.meet(3) 11.prepare(3) 12.operate(3) 13.sell(3) 14.set(3) 15.take(3) 16.talk(3) 17.withdraw(3) 18.arrest(2) 19.build(2) 20.charge(2) 21.change(2) 22.collect(2) 23.develop(2) 24.eat(2) 25.hunt(2) 26.explore(2) 27.interview(2) 28.leave(2) 29.manufacture(2) 30.offer(2)

Infinitive:

begin – 1.see(97) 2.understand(92) 3.feel(65) 4.look(60) 5.think(49) 6.make(41) 7.take(40) 8.appear(31) 9.develop(29) 10.wonder(28) 11.show(27) 12.get(26) 13.build(19) 14.lose(19) 15.work(19) 16.appreciate(18) 17.change(18) 18.grow(18) 19.use(18) 20.believe(17) 21.imagine(17) 22.emerge(16) 23.move(16) 24.realize(16) 25.rise(16) 26.come(15) 27.fall(15) 28.learn(15) 29.find(14) 30.talk(14)

begins – 1.look(41) 2.take(23) 3.fall(21) 4.feel(18) 5.make(17) 6.emerge(16) 7.move(16) 8.get(15) 9.set(14) 10.develop(12) 11.flow(11) 12.grow(11) 13.lose(11) 14.appear(10) 15.rise(10) 16.build(9) 17.run(9) 18.think(9) 19.become(8) 20.seem(8) 21.work(8) 22.break(7) 23.fail(7) 24.approach(6) 25.fade(6) 26.form(6) 27.go(6) 28.melt(6) 29.open(6) 30.realize(6)

beginning – 1.feel(171) 2.think(152) 3.get(146) 4.take(94) 5.wonder(90) 6.look(87) 7.see(85) 8.show(72) 9.make(59) 10.realize(43) 11.find(41) 12.understand(39) 13.emerge(38) 14.enjoy(34) 15.appear(33) 16.turn(31) 17.come(30) 18.lose(30) 19.sound(28) 20.move(25) 21.realize(25) 22.change(22) 23.fall(22) 24.go(21) 25.become(20) 26.happen(20) 27.fade(19) 28.grow(19) 29.rise(19) 30.believe(17)

began – 1.feel(279) 2.take(192) 3.move(188) 4.make(145) 5.think(145) 6.look(140) 7.walk(136) 8.cry(132) 9.get(126) 10.see(117) 11.appear(115) 12.wonder(114) 13.talk(108) 14.fall(100) 15.laugh(99) 16.run(96) 17.speak(91) 18.emerge(83) 19.play(78) 20.grow(75) 21.climb(70) 22.rise(70) 23.sing(68) 24.read(66) 25.write(65) 26.work(64) 27.come(62) 28.go(62) 29.develop(61) 30.show(58)

begun – 1.feel(39) 2.make(36) 3.develop(32) 4.take(30) 5.change(21) 6.emerge(20) 7.think(20) 8.appear(18) 9.look(18) 10.show(17) 11.wonder(15) 12.fall(14)

13.move(14)   14.see(14)   15.grow(13)   16.realise(13)   17.learn(12)   18.rise(11)

19.play(10)   20.question(10)   21.suspect(10)   22.believe(9)   23.form(9)   24.get(9)

25.notice(9) 26.understand(9) 27.come(8) 28.attract(8) 29.build(8) 30.explore(8)

Appendix 2

Activity and state verbs after start/begin construction in the COCA and the BNC (both written and spoken language)

COCA	Start + V ing	Start + V infinitive	Begin + V ing	Begin + V infinitive
1.	talk (2130)	get (1691)	work (1125)	see (1678)
2.	get (1531)	cry (900)	make (685)	feel (1576)
3.	work (1445)	feel (885)	take (630)	think (1457)
4.	look (1419)	see (786)	use (613)	be (1409)
5.	think (1196)	look (679)	talk (560)	take (1345)
6.	do (1181)	say (584)	write (503)	look (1138)
7.	make (1118)	think (561)	look (501)	understand (1058)
8.	take (1080)	go (535)	walk (383)	get (1057)
9.	go (1034)	make (515)	sell (382)	make (959)
10.	play (960)	come (507)	move (372)	move (911)
11.	walk (771)	move (491)	build (328)	change (846)
12.	come (746)	take (473)	play (327)	wonder (811)
13.	use (700)	walk (461)	run (322)	cry (800)
14.	run (681)	turn (383)	offer (304)	realize (779)
15.	move (635)	be (378)	read (299)	fall (700)
16.	call (618)	run (344)	teach (295)	talk (606)
17.	write (567)	laugh (341)	call (270)	turn (580)
18.	ask (509)	talk (329)	think (268)	appear (566)
19.	laugh (499)	fall (317)	study (254)	emerge (553)
20.	have (489)	do (290)	arrive (230)	show (548)
21.	see (484)	work (278)	ask (224)	have (526)
22.	scream (427)	change (274)	collect (221)	grow (498)
23.	sing (425)	lose (257)	do (213)	develop (488)
24.	put (393)	have (250)	search (213)	work (487)
25.	tell (372)	play (247)	plan (203)	speak (480)
26.	feel (352)	show (247)	speak (196)	play (472)
27.	read (346)	rise (246)	appear (195)	come (461)
28.	sell (338)	speak (235)	have (193)	rise (441)
29.	cry (337)	become (233)	turn (182)	lose (430)
30.	say (336)	leave (227)	investigate (177)	run (420)
31.	pay (311)	realize (220)	try (172)	read (392)
32.	shoot (303)	pull (219)	date (170)	walk (392)
33.	give (287)	put (204)	develop (168)	sing (371)

34.	pull (267)	grow (198)	paint (167)	do (368)
35.	date (265)	wonder (195)	operate (166)	use (343)
36.	throw (263)	shake (176)	produce (162)	go (341)
37.	yell (257)	ask (175)	put (162)	form (326)
38.	show (246)	build (175)	shoot (160)	build (322)
39.	eat (244)	write (166)	sing (160)	write (321)
40.	act (237)	happen (164)	get (151)	fade (316)
41.	try (233)	sing (164)	prepare (151)	question (316)
42.	be (229)	open (159)	pull (150)	worry (314)
43.	fall (223)	rain (159)	test (149)	notice (312)
44.	build (212)	believe (150)	show (146)	break (307)
45.	hit (209)	climb (149)	come (143)	tell (305)
46.	teach (208)	tell (145)	pay (141)	suspect (290)
47.	collect (207)	read (140)	receive (140)	recognize (282)
48.	turn (207)	worry (137)	eat (138)	believe (279)
49.	buy (205)	break (136)	tell (138)	laugh (275)
50.	drink (192)	understand (132)	publish (135)	fill (272)

Table 18 Occurrence of activity and state verbs after begin/start constructions in the COCA

BNC	Start + V ing	Start + V infinitive	Begin + V ing	Begin + V infinitive
1.	talk (260)	get (225)	work (70)	be (668)
2.	go (205)	go (133)	ship (55)	feel (572)
3.	get (203)	make (133)	make (52)	take (379)
4.	think (196)	look (128)	talk (50)	think (375)
5.	look (194)	feel (123)	write (49)	look (346)
6.	make (185)	take (114)	take (40)	get (322)
7.	do (181)	come (92)	walk (40)	see (317)
8.	work (152)	cry (86)	move (34)	make (298)
9.	play (133)	walk (85)	think (29)	move (259)
10.	take (114)	move (84)	use (29)	wonder (251)
11.	come (113)	work (84)	look (27)	appear (207)
12.	use (105)	run (80)	read (26)	understand (187)
13.	write (95)	think (79)	build (24)	show (180)
14.	cry (83)	be (68)	play (24)	emerge (173)
15.	run (78)	do (64)	collect (23)	fall (172)
16.	ask (66)	fall (59)	arrive (21)	have (154)
17.	laugh (66)	build (58)	operate (20)	walk (154)
18.	put (65)	laugh (56)	publish (19)	develop (148)
19.	walk (62)	use (56)	run (19)	cry (145)
20.	shout (60)	grow (53)	sell (19)	grow (136)
21.	try (60)	talk (52)	teach (18)	talk (136)
22.	build (59)	play (51)	train (17)	run (132)
23.	move (57)	become (48)	search (16)	rise (126)
24.	say (57)	speak (47)	shout (16)	realise (123)
25.	pay (56)	turn (47)	trade (16)	come (120)
26.	have (52)	write (46)	try (16)	change (118)
27.	ship)52)	show (45)	plan (15)	lose (116)

28.	be (50)	appear (44)	produce (15)	speak (114)
29.	give (49)	say (44)	withdraw (15)	work (114)
30.	sell (49)	read (43)	broadcast (14)	laugh (108)
31.	eat (46)	lose (40)	appear (14)	turn (106)
32.	sing (45)	rain (38)	develop (14)	go (105)
33.	train (42)	rise (37)	market (14)	play (103)
34.	call (41)	have (36)	offer (14)	find (103)
35.	plan (41)	put (36)	prepare (14)	realize (92)
36.	buy (40)	see (34)	climb (13)	build (88)
37.	smoke (39)	eat (33)	draw (13)	write (86)
38.	tell (39)	climb (32)	experiment (13)	climb (85)
39.	collect (38)	shake (32)	pace (13)	read (83)
40.	read (38)	change (32)	put (12)	tell (82)
41.	drink (36)	emerge (31)	speak (12)	sing (79)
42.	shoot (36)	give (31)	turn (12)	enjoy (77)
43.	scream (34)	develop (30)	clear (11)	form (74)
44.	throw (33)	break (28)	fill (11)	use (73)
45.	cut(30)	pick (28)	fire (11)	shake (68)
46.	feel (30)	pull (27)	investigate (11)	break (67)
47.	dig (29)	arrive (25)	pull (11)	put (67)
48.	pull(29)	sing (25)	set (11)	question (65)
49.	dance(28)	ask (24)	sort (11)	believe (65
50.	fight(28)	worry (24)	cut (10)	fill (61)

Table 19 Occurrence of activity and state verbs after start/begin constructions in the BNC