



Bakalářská práce

The Comparison of Differences between American and British English in Tennis Articles

Studijní program:

B0114A300068 Anglický jazyk se zaměřením
na vzdělávání

Studijní obory:

Anglický jazyk se zaměřením na vzdělávání
Tělesná výchova se zaměřením na vzdělávání

Autor práce:

Sára Kadlecová

Vedoucí práce:

Mgr. Petra Peldová, Ph.D.

Katedra anglického jazyka

Liberec 2023



Zadání bakalářské práce

The Comparison of Differences between American and British English in Tennis Articles

<i>Jméno a příjmení:</i>	Sára Kadlecová
<i>Osobní číslo:</i>	P20000298
<i>Studijní program:</i>	B0114A300068 Anglický jazyk se zaměřením na vzdělávání
<i>Specializace:</i>	Anglický jazyk se zaměřením na vzdělávání Tělesná výchova se zaměřením na vzdělávání
<i>Zadávací katedra:</i>	Katedra anglického jazyka
<i>Akademický rok:</i>	2021/2022

Zásady pro vypracování:

Cíl: Vytvoření seznamu slov odlišně použitých ve sportovních článcích o tenise v americkém a britském tisku, a jejich kvalitativní a kvantitativní analýza pomocí nástrojů korpusové lingvistiky.

Metodologie: Vytvoření dvou korpusů z novinových článků o tenise, jejich porovnání s ohledem na lexikální a gramatickou rozdílnost.

Požadavky: Orientace v odborné literatuře, znalost principů korpusové lingvistiky, konzultace.

Aim: To create a list of differently used words in tennis articles in the American and British press and their qualitative and quantitative analysis using corpus linguistics tools.

Methodology: Compiling two corpora of tennis newspaper articles, their comparison concerning lexical and grammatical differences.

Requirements: Orientation in specialized literature, knowledge of principles of corpus linguistics, consultations.

Rozsah grafických prací:

Rozsah pracovní zprávy:

Forma zpracování práce:

Jazyk práce:

tištěná/elektronická

angličtina

Seznam odborné literatury:

Čermák, František, Patrick Corness, and Aleš Klégr, eds. 2010. *InterCorp: exploring a multilingual corpus*. Studie z korpusové lingvistiky, sv. 13. Praha: NLN, Nakladatelství Lidové noviny.

Čermák, František. 2017. *Korpus a korpusová lingvistika*. Praha: Univerzita Karlova, nakladatelství Karolinum.

McEnery, Tony, and Andrew Hardie. 2012. *Corpus linguistics: method, theory and practice*. Cambridge textbooks in linguistics. Cambridge: Cambridge University Press.

Baker, Paul. 2017. *American and British English: divided by a common language*. Cambridge: Cambridge University Press.

Davies, C. (2007). *Divided by a common language: A guide to British and American English* (1st Houghton Mifflin pbk. ed). Houghton Mifflin.

Vedoucí práce:

Mgr. Petra Peldová, Ph.D.

Katedra anglického jazyka

Datum zadání práce:

27. června 2022

Předpokládaný termín odevzdání: 14. července 2023

L.S.

prof. RNDr. Jan Pícek, CSc.
děkan

Mgr. Zénó Vernyik, Ph.D.
vedoucí katedry

V Liberci dne 27. června 2022

Prohlášení

Prohlašuji, že svou bakalářskou práci jsem vypracovala samostatně jako původní dílo s použitím uvedené literatury a na základě konzultací s vedoucím mé bakalářské práce a konzultantem.

Jsem si vědoma toho, že na mou bakalářskou práci se plně vztahuje zákon č. 121/2000 Sb., o právu autorském, zejména § 60 – školní dílo.

Beru na vědomí, že Technická univerzita v Liberci nezasahuje do mých autorských práv užitím mé bakalářské práce pro vnitřní potřebu Technické univerzity v Liberci.

Užiji-li bakalářskou práci nebo poskytnu-li licenci k jejímu využití, jsem si vědoma povinnosti informovat o této skutečnosti Technickou univerzitu v Liberci; v tomto případě má Technická univerzita v Liberci právo ode mne požadovat úhradu nákladů, které vynaložila na vytvoření díla, až do jejich skutečné výše.

Současně čestně prohlašuji, že text elektronické podoby práce vložený do IS/STAG se shoduje s textem tištěné podoby práce.

Beru na vědomí, že má bakalářská práce bude zveřejněna Technickou univerzitou v Liberci v souladu s § 47b zákona č. 111/1998 Sb., o vysokých školách a o změně a doplnění dalších zákonů (zákon o vysokých školách), ve znění pozdějších předpisů.

Jsem si vědoma následků, které podle zákona o vysokých školách mohou vyplývat z porušení tohoto prohlášení.

Poděkování

Velmi děkuji vedoucí mé práce paní doktorce Mgr. Petře Peldové, Ph.D., za její vedení, připomínky, odborné znalosti, podporu a zpětnou vazbu v průběhu zpracování bakalářské práce.

Anotace

Bakalářská práce „Porovnání rozdílů americké a britské angličtiny ve sportovních článcích o tenise“ se zaměřuje na rozdíly americké a britské angličtiny v tenisových článcích vytvořením dvou korpusů z novinových článků o tenise a jejich porovnání s ohledem na lexikální a gramatickou rozdílnost. Teoretická část poskytuje informace o anglickém jazyce, rozdílech americké a britské angličtiny a korpusové lingvistice. Praktická část obsahuje data, metodologii, pilotní a velkou korpusovou analýzu. Cílem bakalářské práce je identifikovat rozdíly v pravopise a gramatice mezi americkou a britskou angličtinou. Následně se snaží zjistit, jak moc se liší slovní zásoba v amerických a britských tenisových článcích, zejména se zaměřením na rozdíly v příslovcích, přídavných jménech, slovesech a podstatných jménech.

Klíčová slova

korpusová lingvistika, korpus, rozdíly, americká angličtina, britská angličtina, porovnání, analýza, tenis, novinové články

Abstract

The Bachelor thesis "The Comparison of Differences between American and British English in Tennis Articles" focuses on the differences between American and British English in tennis sports articles by compilation of two corpora of tennis newspapers articles and their comparison concerning lexical and grammatical differences. The theoretical part of the thesis provides information about the English language, the difference between American and British English and corpus linguistics. The practical part of the thesis contains data, methodology, pilot, and large corpus analyses. The bachelor's thesis aims to identify differences in spelling and grammar between American and British English in tennis articles. Subsequently, it seeks to determine how the vocabulary differs in American and British tennis reports, mainly focusing on distinctions in adverbs, adjectives, verbs and nouns.

Keywords

corpus linguistic, corpus, differences, American English, British English, comparison, analyses, press, tennis, news articles

List of Contents

List of Tables	6
List of Abbreviations	7
Introduction.....	8
I Theoretical part	9
1 English language	9
2 Differences between American and British English	11
2.1 Spelling	11
2.2 Pronunciation	13
2.3 Grammar.....	15
2.4 Vocabulary	18
2.4.1 Differences in the vocabulary of American and British tennis.....	20
2.4.2 What not to say	20
2.4.3 Examples of words with different meanings.....	21
3 Corpus linguistic	22
3.1 Analysing corpus data.....	23
II Practical part.....	24
4 Data and Methodology.....	24
4.1 Pilot Corpus	24
4.2 Large Corpus	25
5 Analyses.....	27
5.1 Pilot Corpus	27
5.1.1 Grammar.....	27
5.1.2 Adverbs.....	27
5.1.3 Adjectives	29
5.1.4 Verbs.....	31
5.1.5 Nouns.....	33
5.1.6 Spelling	34
5.2 Large Corpus	35
5.2.1 Keywords	35

5.2.2	Verbs.....	36
5.2.3	Nouns.....	39
5.2.4	Adjectives	41
5.2.5	Adverbs.....	43
5.2.6	Grammar.....	43
	Conclusion	45
	References	47
	Appendices	49

List of Tables

Table 1 Distribution of the number of tokens in American and British articles of the pilot study.....	25
Table 2 Distribution of basic information of ALC and BLC	26
Table 3 Distribution of present perfect and past simple in the pilot corpora (raw frequency)	27
Table 4 List of the first twenty adverbs in APC and BPC	28
Table 5 List of the first twenty adjectives in APC and BPC.....	29
Table 6 Distribution of the most used positive adjectives in APC and BPC.....	30
Table 7 List of the first twenty used verbs in APC and BPC.....	32
Table 8 List of the first twenty most used nouns in APC and BPC	34
Table 9 List of the first twenty keywords of ALC and BLC	35
Table 10 List of the five most frequent verbs with -ize ending in ALC and five most frequent verbs with -ise ending in BLC.....	37
Table 11 Number of modal verbs in ALC and BLC.....	37
Table 12 List of the first twenty most used verbs in ALC and BLC.....	38
Table 13 List of the first twenty most used nouns in ALC and BLC	40
Table 14 Instances of identical adjectives with the -less suffix in ALC and BLC	42
Table 15 The five most frequent adjectives with the prefix -un in ALC and BLC.....	42
Table 16 The five most frequent adverbs in ALC and BLC.....	43
Table 17 Distribution of present perfect and past simple in ALC and BLC.....	44
Table 18 List of the American and British sources	49

List of Abbreviations

ALC (American Large Corpus)

AmE (American English)

APC (American Pilot Corpus)

BLC (British Large Corpus)

BPC (British Pilot Corpus)

BrE (British English)

CIC (Cambridge International Corpus)

GA (General American)

KWIC (key words in contexts)

NYT (New York Times)

RP (Received pronunciation)

TTE (type-token ratio)

UK (United Kingdom)

US (United States)

WSE (World Standard English)

Introduction

American and British English differ in spelling, pronunciation, grammar and vocabulary. This bachelor's thesis aims to identify differences in spelling and grammar between American and British English in tennis articles. Subsequently, it seeks to determine how the vocabulary differs in American and British tennis reports, mainly focusing on distinctions in adverbs, adjectives, verbs and nouns, utilizing tools from corpus linguistics.

The theoretical part of the thesis focuses on presenting basic information about the English language. Next, it explores the differences between American and British English, such as spelling, pronunciation, grammar, and vocabulary. The vocabulary chapter provides insight into distinctions in tennis between American and British English, instances of what not to say, and examples of words with different meanings. Finally, the fundamental knowledge of corpus linguistics is presented in the last chapter of the theoretical part.

In the practical part of the thesis, firstly, the pilot corpora are created, one for American English using articles from the American press, and one for British English, using reports from British papers. The analysis of the pilot corpora primarily concentrates on examining grammar, adverbs, adjectives, verbs, nouns, and spelling differences.

In conclusion, the large corpora are created, and their analyses are established based on the findings from the pilot corpora: examining the differences between American and British English in keywords, verbs, nouns, adjectives, adverbs and grammar.

I Theoretical part

1 English language

The English language varies in different parts of the world through significant development, and it is probably one of the most broadly spoken languages worldwide. American and British English are examples of how English has changed in many ways. Each variant has a separate development; they have no superiority or inferiority nowadays. However, English learners need to know the differences between these two variants to know how English is used (Han 2019, 93-94).

The differences between those two variants can be observed in all the aspects of the language system, i.e., lexis, grammar, pronunciation, discourse and function. According to (Davies 2007, 1) the most distinctive features are in lexis; for example, words such as *caboose*, *bleachers*, and *busboy* are common in the US, but they would confuse most British English speakers. In contrast, words such as *mailshot*, *crosspatch*, and *gymkhana* would seem slightly odd to an American speaker.

Webster (1789, in Davies 2007, 2) predicted that American English would be divergent from British English, and his writing of the *American Dictionary of the English Language* evolved a standard for spelling and word usage in America. He also recommended giving every letter in a syllable a proportion of sound, which might contribute to the pronunciation dissimilarities between American and British English.

Furthermore, Darragh (2000, V) asserts that approximately 4000 everyday words in Britain vary in meaning or do not exist compared to the US.

The following paragraphs deal with the fundamental differences between American and British English: spelling, pronunciation, grammar, and contrasts in vocabulary.

2 Differences between American and British English

2.1 Spelling

In general, the term *spelling* pertains to the demonstration of orthographic knowledge (Rapp 2001, 263).

Greenbaum (1996) thinks that proper spelling reflects how high the writer's knowledge of a particular language is; therefore, it is essential to know and use the spelling correctly. People were unaware that each word should have a unique spelling in the past. In 1700, a steady spelling system was invented for print, and Samuel Johnson's dictionary followed the norms of this spelling system. Britons accepted Johnson's dictionary at the end of the 18th century as the authority for personal use. In the USA, Noah Webster got the credit for influencing American spelling with his spelling book in 1783 and his dictionary in 1828 (558).

In the subsequent paragraphs, some examples of differences in British and American English spelling, according to Darragh (2000) and Svartvik and Leech (2006), are presented. The following examples are not exhaustive; they serve to illustrate the variations.

To begin with, there are specific verbs that can only be suffixed with **-ize** (e.g., *seize*, *capsize*), while others can only take **-ise** (e.g., *advertise*, *advise*, *surprise*). The writer's location does not impact the choice of suffix. However, American and British dictionaries generally favour the **-ize** ending for words like *apologize* and *realize*. Despite this, British writers tend to disagree with these conventions and favour the **-ise** ending when they write.

There are also some one-letter differences concerning one group of words spelt differently in only one or additional letter. These dissimilarities affect the pronunciation of words. Examples in British English are *mum*, *potter*, *speciality*, and *mom*, *putter*, *specialty* in American English.

Additional divergences can be observed in the endings, where most words in British English have **-our** ending (*armour*, *favour*, *colour*), whereas words in American English have **-or** ending (*armor*, *favor*, *color*). However, exceptions such as *glamour* used in AmE and *tenor* in BrE also exist. Next, the ending **-tre** is used in British English (*centre*, *litre*, *metre*, *theatre*), and the ending **-ter** is used in American English (*center*, *liter*, *meter*, *theater*). In addition, some words end in **-ense** in AmE (*defense*, *license*), meanwhile **-ence** in BrE (*defence*, *licence*).

There is another distinction concerning the nouns ending in **-ogue** used in British English, whereas Americans use shortened ending **-og**.

(1) BrE: catalogue, dialogue

(2) AmE: catalog, diaolog

Also, British English typically uses a single **-l** in the spelling of certain disyllabic verbs, such as *enrol*, *fulfil*, and *instal*. On the other hand, American English employs a double **-ll** in these verbs, spelling them as *enroll*, *fulfill*, and *install*.

A further example of words that vary in American and British English that have not been mentioned above because they do not belong to a particular group are, for instance:

(3) BrE: aeroplane/AmE: airplane, BrE: disc /AmE: disk, BrE: pyjamas /AmE: pajamas,

(4) BrE: sceptical/AmE: skeptical, BrE: tyre /AmE: tire, BrE: whisky /AmE: whiskey

Eventually, Baker (2017, 50-51), based on corpus research made in 1931, 1961, 1991, and 2006 claims that using *toward/towards* also differs in American and British English. American English tends to use the shortened version *toward*. In contrast, British English favour the full version *towards*.

2.2 Pronunciation

Dalton and Seidlhofer (2000, 3) define *pronunciation* as the production of sounds in two meanings. Firstly, the sound is part of a particular language's code, meaning there are distinctive sounds in English. Secondly, codes are combined with other factors to make communication possible, indicating pronunciation is used to achieve meaning in context.

Pronunciation differences between American and British English encompass various aspects such as stress patterns, consonant sounds, distribution of sounds, articulation, and intonation/tune. Indeed, tune or intonation, which refers to the melodic pattern accompanying sentences, is recognized as one of the noticeable differences between American and British English. However, it is worth noting that when people sing, the melodic structure of the music overrides the intonation patterns of spoken language, diminishing the significance of pronunciation differences between American and British English (Algeo 2006, 2).

Also, Darragh (2000, 8) claims it is essential to realise which pronunciation aspect we are discussing. While dialects are commonly identified by variations in grammar and vocabulary, accents primarily pertain to pronunciation. Various regional accents are often closely linked to specific dialect regions. Phonologists have identified 16 modern dialect regions in England and 26 in the United States.

The following examples are presented by Darragh (2000) and Davies (2007). They are primarily applied to Standard American English or GA (General American) and Standard British English or RP (Received Pronunciation); this list is incomplete.

The pronunciation of the sound /r/ is one of the most notable distinctions between American and British English. In Received Pronunciation (RP), this sound is considered absent. However, before vowels, /r/ is still present. In contrast, in American English /r/ sound is present before vowels and consonants and at the endings of words – this is called *rhoticity* in phonetics.

In addition, the pronunciation of /ʌ/ in such two words as *fast*, *grass*, and *dance* is long and firm in British English but short in American English. In southern England, during the late 18th century, a transformation occurred where the so-called flat /ʌ/ sound started shifting to the broad /ʌ/ sound. This alteration resulted in the pronunciation of words such as *man* having a similar /ʌ/ sound to that found in *father*. This change was also observed in certain regions of New England, although the majority of the area continues to use the original sound.

Furthermore, the pronunciation of /t/ differs. In British English, /t/ is pronounced relatively clearly; however, it might be pronounced as /d/ or disappear entirely after nasal sounds such as /m/, /n/, and /ŋ/ (*dentist/dennist*, *twenty/twenny* etc.) in American English.

Here are a few common words pronounced and stressed differently in British and American English.

(5) accent: AmE: **ak**.sənt/BrE: **ak**.sent

(6) been: AmE: been/BrE: bin

(7) detail: AmE: **dee**.tail/BrE: de.**tail**

(8) schedule: AmE: **shed**.ule/BrE: **sked**.ule

(9) trauma: AmE: **trau**.ma/ BrE: **trah**.ma

Conclusively, Darragh (2000, 14) states that American and British English stress different syllables; however, the distinctions are rare and most understandable. The dissimilarities represent a crucial point in understanding and standing out in conversations. Also, Americans generally tend to speak more slowly, with less intonation.

2.3 Grammar

"*Grammar* is the rules that show how words are combined, arranged or changed to show certain kinds of meaning" (Swan 2016, 23). According to Greenbaum (1996), the first grammar book was published by William Bullokar in 1580. The author wanted to prove that he could grammatically analyse the English Language in the same way as Latin. Nowadays, grammar books; they vary in content and meaning and are created for native speakers and foreign language learners. Grammars focus on various aspects of language, such as morphology, phonology, orthography, vocabulary and syntax (23, 37).

The contrasts between American and British English in grammar are indistinct. They can also be confusing because American and British English frequently affect each other primarily through the Internet, as well as films, television, pop music, and other cultural mediums. However, the influence of the American language on British English is more significant and is getting bigger every day through American films, music, radio, television, among other things. On the other hand, Americans struggle more with an understanding of the British accent (Darragh 2000, 15).

Svartik and Leech (2006) state that American users tend to be more 'grammar conscious' and obey the grammar rules more. One of the examples is the break of the rule of agreement in British English, where singular collective nouns, for instance, *teams*,

audience, government, and the public, can behave as plural; meanwhile, Americans use the usual singular form (167).

The most significant difference in grammar is presumably gotten/got distinction. Gotten in AmE is not used in the sense of possession (have) and is commonly misused by Britons. 'Have got' and 'gotten' are not synonyms. AmE can make differences as in the following example: they have got to leave (they must leave) vs they have gotten to leave (they have managed to leave) (Crystal 1995, 311).

Algeo (2006) and Darragh (2000) highlights various discrepancies in grammar through several points.

Firstly, in American English, a verb's regular past tense (*burned, spelled*) is preferred, whereas, in British English, a verb's irregular past tense form (*burnt, spelt*) is favoured. Nevertheless, there are both variants possible in either variation of the language.

Also, the present perfect tense is usually used when an action in the past resulted in the present, and Americans tend to use Past Simple and Present Perfect interchangeably. In contrast, the Britons mostly rely on Present Perfect.

(10) AmE: Have you cooked dinner yet? /Did you cook dinner yet?

(11) BrE: Have you cooked dinner yet?

Next, there are two verbal expressions when discussing a future time: *will* or *shall* (the modal future) and *be going to* (the periphrastic future). There is a tendency for Britons to utilise *will* and *shall* more frequently than Americans in general. In contrast, Americans are more inclined to use *be going to*, particularly in informal conversations (almost 2.3 times more frequently, according to CIC).

Furthermore, the position of adverbs varies in American and British English. In British English, when emphasising the auxiliary verb, mid-position adverbs such as

sometimes, always, never, often, and definitely are placed before the auxiliary. On the other hand, in American English, mid-position adverbs are often positioned before the auxiliary without the specific intention of adding emphasis to the given information.

(12) AmE: It certainly will lead to a vote (normal).

(13) BrE: She certainly has arrived by now (emphatic).

The following differences refer to modal verbs. For instance, the modal verb *shall* is uncommon in American English, whereas it is commonly used in British English, where it can be substituted for *will*. In contrast, Americans often prefer to use *should* when seeking advice.

(14) AmE: Which bus should I take?

(15) BrE: Which bus should/shall I take?

Next, Britons use *needn't* as a substitution for *don't need to*, which is not a typical modal verb used in America. Other modal verbs, *must* or *ought to*, are more frequent in British than in American English.

Another difference concerns the word *real* as an intensifier. When speaking informally, Americans use *real* before adjectives (That song was a *real* nice.); meanwhile, Britons insist on using *really* (That song was *really* nice).

Prepositions occasionally vary in American and British English. Sometimes an entirely different preposition is used, and occasionally there is no preposition compared to the other varieties. See the following examples.

(16) AmE: What do you do on the weekends? /BrE: What do you do at the weekends?

(17) AmE: I would like for you to go now. /BrE: I would like (-) you to go now.

(18) AmE: He will have to do it over. / BrE: He will have to do it again.

(19) AmE: She arrived at ten after three. / BrE: She arrived at ten past three.

Next, the pronoun *one* is more commonly used in British than American English. In American English, pronouns like *he*, *him*, and *his* are often used later in the sentence instead of *one*. The Britons continue to use *one*.

(20) AmE: One should always be kind to his father.

(21) BrE: One should always be kind to one's mother.

2.4 Vocabulary

"The term *lexicon* is known in English from the early 17th century. The term itself comes from the Greek *lexis* 'word.' It is still used today in this word-book meaning, but it has also taken on a more abstract sense, especially within linguistics, referring to the total stock of meaningful units in language" (Crystal 1995, 118).

When studying lexicology in the English language, exploring all aspects of the vocabulary becomes necessary. This includes examining the development and formation of words, their relationships with other words, their current usage, and their representation in dictionaries or word books (Crystal 1995, 118).

According to Crystal (1995, 119), two highly significant dictionaries in the field of lexicography are *Webster's Third New International Dictionary* and the *Oxford English Dictionary*. These dictionaries are renowned for including approximately half a million lexemes. However, the actual number is expected to be even higher – with the internet.

The well-known writer and journalist H. L. Mencken authored a book titled *The American Language*, in which he extensively explored the linguistic differences between American and British English. In 1936, Mencken published a list of American-British differences. This list comprised American words such as *bakery*, *bank account*, *hardware*, and *raincoat*. However, these words are now common in British English

mainly because the British have actively borrowed them linguistically from Americans. This scenario happened later in the nineteenth century and mainly in the twentieth century (Svartik and Leech 2006, 157).

American and British English are expanding in the amount of impact the two models have had on each other. Crystal (1995, 306) states, "The influence of US films and television has led to a considerable passive understanding of much American English vocabulary in Britain, and some of this has turned into active use, especially among younger people. The reverse pattern is less obvious, but British films and TV are seen sufficiently often in the USA to mean that growth in awareness of UK vocabulary should not be discounted."

Three distinctions have to be made when discussing the vocabulary of American and British English. Some words exist only in American English (Americanism), some only in British English (Britishisms), and some have been established worldwide by taking part in World Standard English (WSE) (Crystal 1995, 308).

In the following paragraphs, examples of distinction in lexis by Crystal (1995) are presented; the list is not exhaustive.

Firstly, the effect of frequency is an important thing to remember. Some words can be used in both types of language, but they are likely to be used more in British or American English. For instance, words such as *shop* vs *store* are used in both varieties, but *shop* is used more often in British English, whereas *store* is more frequent in American English.

Some words reflect cultural distinctions; however, synonyms in the other varieties do not exist. These words are not part of WSE.

(22) AmE: Ivy League, Groundhog Day, revenue sharing

(23) BrE: A-levels, giro, VAT

There are also specific words with a single definition and synonyms in the other variety.

(24) AmE: checking account, station wagon

(25) BrE: current account, estate car

Also, some words have different meanings in both varieties and no WSE meaning at all.

(26) AmE: flyover = BrE: flypast

however

(27) BrE: flyover = AmE: overpass

2.4.1 Differences in the vocabulary of American and British tennis

In tennis, there are variations in the terms used between American and British English. In British English, the server is determined as *rough* or *smooth*, while in American English, the terms *up* or *down* are used to determine the server. In American English, balls fall in the *alley*, compared to British English, where they fall in the *tramlines*. Americans use the term *toss* when a player serves, but Britons call it *throw up*. In British English, one has *knock-up* before play, whereas in American English, one has *warm-up*. The name: *Australian doubles* or *Canadian doubles* is a tennis game for three players, used primarily in American English (Davies 2007, 50-51).

2.4.2 What not to say

Even though Americans and Britons can usually understand each other, a few words and expressions could cause problems in conversation. Among the British lexical

items that can suggest something different in American English belongs, for instance, the adjective *homely*, which means *cosy* in BrE but *unattractive* in AmE, or the phrase *potty about* (= be very interested in something in the BrE) would not be understood in American English. *Homey* and *crazy about* would be the American alternatives for these phrases (Davies 2007, 101-102).

Next, according to Davies (2007, 102-104), among the American terms which could confuse or insult the British citizen belongs, for example, the phrase *I am stuffed*, which can be perceived as vulgar in the UK, *I am full* would be a better option. Another example of American slang is the word *bummer* (= huge disappointment in the US), which should be replaced by *nuisance* or *disappointing* in the UK.

2.4.3 Examples of words with different meanings

The following words have relatively different meanings in the United Kingdom and the United States. The list is not exhaustive; these are examples only.

(28) basket case: AmE meaning: an exhausted person/BrE meaning: an insane person

(29) casket: AmE meaning: coffin/BrE meaning: small ornamental box

(30) davenport: AmE meaning: sofa/BrE meaning: writing desk

(31) jumper: AmE meaning: pinafore dress/BrE meaning: sweater

(32) knickers: AmE meaning: knee-length trousers/BrE meaning: panties

(33) loft: AmE meaning: open, elevated in a room/BrE meaning: attic

(34) yard: AmE meaning: lawn around a house/BrE meaning: paved area

(Davies 2007, 101-107).

3 Corpus linguistic

As McEnery and Hardie (2012, 1-2) claim, *corpus linguistics* is "an area which focuses upon a set of procedures, or methods, for studying language." It could be characterised as dealing with a group of text read by machines which are considered a suitable basis for studying a concrete group of research questions. The term *corpus*, in other words, the set of computerised texts, is the scale of used data clarifying the usage of machine-readable texts.

When working with a large amount of data, it is not practicable to search it by hand because of the time a human or team analyst would spend searching through the text unless a computer is used. Another potential issue that arises from manual searching is the possibility of errors, as highlighted by McEnery and Hardie (2012, 2).

Subsequently, McEnery and Hardie (2012, 2) mention that another generalisation followed from observation is that "corpora are invariably exploited tools which allow users to search through them rapidly and reliably." Among these tools, for instance, belong *concordancers* or *word frequency lists*. *Concordancers* permit users to examine the word in context. The *word frequency list* is used for listing all words that appeared in the corpus, giving us specific information about the number of occurrences of each word in the corpus.

Next, one of the easy measures of linguistic diversity is *TTR* (type-token ratio) which is also understood as vocabulary size divided by text length (Kettunen 2014). According to Dax (2005, 1), the TTR ratio is calculated by dividing the number of different words (types) in a text or utterance by the total number of words (tokens) in that

text or utterance, where a low TTR indicates a low degree of lexical variation, while a high TTR suggests the opposite.

In this bachelor thesis, the TTR is used to determine the frequency of adverbs, adjectives, verbs and nouns in American and British corpora.

3.1 Analysing corpus data

Čermák (2017, 51) claims *concordance* is the primary and most used manner for representing results in the specific corpus. It is a list of all occurrences of searched form in all contexts. *KWIC* is the abbreviation for a keyword in context, giving us a hint of concordance. Scott and Tribble (2006, in Čermák et al. 2010, 177) define *keywords* as "items of unusual frequency in comparison with a reference corpus."

Sketch Engine defines a *reference corpus* as a "corpus to which focus corpus is compared and is used in keyword extraction and term extraction."

However, *concordance* can be replaced by a more complex instrument called *word sketch*, the lexicographical tool. It can search in a specific corpus its word's outline and simultaneously represents statistically evaluated types of collocations according to its common grammatical function (Čermák 2017, 51).

II Practical part

4 Data and Methodology

4.1 Pilot Corpus

This bachelor's thesis aims to identify differences in spelling and grammar between American and British English in tennis articles. Subsequently, it seeks to determine how the vocabulary differs in American and British tennis reports, mainly focusing on distinctions in adverbs, adjectives, verbs and nouns.

In order to find out the most frequent distinctions between the American and British tennis articles, three news articles from the British source (The Guardian) and the American sources (US News, The Wall Street Journal) were selected and downloaded into Microsoft Word under the abbreviations A (number of the article) and B (number of the article).

The marking "A" indicates that a specific article is American, whereas "B" suggests that the article is British. These articles were published on July 9, July 10, and June 28, 2022, and are similar in topic and word count, which is why these specific sources were used. In total, six articles were compared, see Table 1. All the British articles were compiled into the British Pilot Corpus (BPC) and American Pilot Corpus (APC).

Firstly, the articles were read and compared numerically, see Table 1. Consequently, the articles were manually examined in terms of spelling, grammar, and vocabulary, aiming to detect notable differences.

The findings are discussed in Chapter 5.1.

Article	American Pilot Corpus	British Pilot Corpus
A_3, B_3	972	935
A_7, B_7	924	934
A_9, B_9	1060	910
Tokens in total	2 956	2 779

Table 1 Distribution of the number of tokens in American and British articles of the pilot study

4.2 Large Corpus

Two Large Corpora were created to analyse the differences between American and British articles about tennis. The procedure was as follows. From June 2022, tennis news articles from British and American online sources (see Appendix 2) were collected; the specific articles without advertisements were manually copied into a Microsoft Word document and saved under the appropriate name (A_1, B_1, A_2, B_2, A_3, B_3...).

The complete overview of all downloaded articles in a Microsoft Excel document was created. This document contains information about each report, such as caption, source, date, author and links to the articles. See Appendix 1 for more details on the sources.

Overall, one hundred and twenty articles for each Large Corpus were downloaded regularly from June until February 15. Some websites had a separate section for tennis; however, most articles were located in the *Sports* section. When searching for articles, the following were considered: date, title and sometimes even keywords.

The program called *Sketch Engine* was used for the creation of the corpora. See Table 2 for basic information about American Large Corpus (ALC) and British Large Corpus (BLC.) One can see that the size is comparable, so normalised frequency in findings was not used.

American Large Corpus		British Large Corpus	
Tokens	81,867	Tokens	84,306
Words	69,178	Words	72,339
Sentences	3,292	Sentences	3,358
Documents	120	Documents	120

Table 2 Distribution of basic information of ALC and BLC

The focus was mainly on distinctions between AmE and BrE in keywords, verbs, nouns, grammar, adjectives, adverbs and spelling. The results can be found in Chapter 5.2. The functions such as *keywords*, *word list*, and *concordance* in the Sketch Engine program were utilised to analyse these distinctions—also, the function *TTR* was frequently used.

5 Analyses

5.1 Pilot Corpus

As mentioned in Chapter 4.1, the pilot corpora consist of 2 956 (US) and 2 779 (UK) tokens. Based on the theoretical part and the information on the differences between American and British English. It was decided to look for the following differences in grammar, adverbs, adjectives, verbs, nouns and spelling.

5.1.1 Grammar

The most evident difference in the Pilot Corpora was in grammar, namely in the frequency of the present perfect tense occurrence. In all six articles, the present perfect tense is used more frequently in the British reports than in the American reports, see Table 3 and examples (35) and (36). This finding complies with the theory stated in Chapter 2.3 Grammar.

	American Pilot Corpus	British Pilot Corpus
Present perfect	9	22
Past simple	112	123

Table 3 Distribution of present perfect and past simple in the pilot corpora (raw frequency)

(35) But the experienced 35-year-old **fought** his way back to win his seventh Wimbledon. A_9

(36) Wimbledon is his second most successful grand slam, yet Djokovic **has now won**. B_9

5.1.2 Adverbs

Another distinction was detected in the number of adverbs used in both Pilot Corpora. Even though almost identical in size, the American Pilot Corpus (APC) shows a lower type-token ratio (TTR) in the use of adverbs (40 per cent) than the British Pilot Corpus (48 per cent). It could mean that the British Pilot Corpus (BPC) relies on more

ways to describe the manner of happening in the tennis world. This might be an issue worth analysing in the Large Corpora.

When further analysing the lists of adverbs detected, see Table 4, it can be noticed that the British authors of the articles slightly more apply adverbs with -ly suffixes (30 per cent) than the American authors (27 per cent). To see in which context these adverbs were used, see examples (37) – (40).

American Pilot Corpus		British Pilot Corpus	
1. not	11. also	1. not	11. finally
2. so	12. however	2. so	12. instead
3. here	13. early	3. back	13. around
4. just	14. much	4. then	14. most
5. back	15. only	5. well	15. forward
6. most	16. afterwards	6. gradually	16. never
7. really	17. yet	7. as	17. extremely
8. again	18. down	8. now	18. too
9. eventually	19. never	9. even	19. yet
10. now	20. more	10. maybe	20. definitely

Table 4 List of the first twenty adverbs in APC and BPC

- (37) At the end of a wonderful two weeks when her form and confidence **gradually** blossomed. B_7
 (38) She has ensured that she will be stimulated when her career **finally** comes to an end. B_3
 (39) Needing to fend off multiple break points before **eventually** taking a two-game lead in the set. A_7
 (40) But the comeback ended **early** when the duo pulled out before their semi-final. A_3

In addition, a point worth examining is the position of the adverbs in sentences. In the American articles, the mid-position adverb is placed before the auxiliary verb even if there is no intention to emphasise something. Whereas, in the British English articles, the adverb usually stands after the auxiliary verb and is always emphatic; see example (41) and (42) for a comparison of the two variations.

- (41) It **always** has been and will be the most special one in my heart. A_9
 (42) It has **evidently** returned. B_9

This discovery aligns with the theory put forth by Darragh (2000, 19), which states that specific adverbs like *sometimes*, *always*, *never*, *often*, and *certainly* in British English are positioned before the auxiliary verb to add emphasis. However, in American English,

these adverbs are placed before the auxiliary with no intention to emphasise. Algeo (2006, 148-149) claims that discoveries in the CIC (Cambridge International and Corpus) text support the generalisation that American English favours placing the adverbs of frequency and probability before the auxiliary. In contrast, British English has the opposite preference. So, both assertions of these two authors confirm the result of this pilot study.

5.1.3 Adjectives

Regarding vocabulary, focusing on adjectives, the TTR in the BPC is 58 per cent and 51 per cent in the APC; there is a visible quantitative difference, but not significant. The first three most frequently used adjectives in the corpora are *grand*, *first*, and *final*; it is due to the fact that articles are reports from tennis tournaments therefore *grand* is a part of Grand Slam. *First* is also connected to the Grand Slam or set, and the adjective *final* is used because the news is about the final match. See the negligible differences in Table 5.

Only a relatively small portion of adjectives is used repeatedly in each corpus (25 per cent in the BPC and 30 per cent in the APC), the remaining adjectives are *hapaxes*. They appear just once in each corpus. See Table 5 with the top twenty adjectives in BPC and APC.

American Pilot Corpus		British Pilot Corpus	
1. first	11. more	1. grand	11. most
2. final	12. strong	2. first	12. short
3. grand	13. few	3. final	13. last
4. second	14. fourth	4. more	14. tight
5. young	15. amazing	5. new	15. second
6. happy	16. strong	6. good	16. other
7. other	17. short	7. few	17. third
8. big	18. same	8. great	18. long
9. early	19. much	9. many	19. forehand
10. next	20. third	10. fourth	20. solid

Table 5 List of the first twenty adjectives in APC and BPC

Examining the whole list of adjectives further, it can be seen that in both corpora, the first three adjectives (*first*, *final* and *great*) are the same; only their position differs. These adjectives are ordinary because they mainly describe tennis tournaments. Also, it can be seen that there are positively loaded adjectives in both corpora, yet they differ. See Table 6 with the most used positive adjectives and following nouns.

American Pilot Corpus	British Pilot Corpus
1. amazing – game, talent, to be a winner is amazing	1. best – tennis, players
2. great – match, one of the greatest	2. great – team, results, years
3. special – tennis court, one	3. startling – revelations, titles
4. powerful – serve	4. delicious – drop shot
5. remarkable – title	5. wonderful – weeks
6. inspiring – players	6. wondrous – lob

Table 6 Distribution of the most used positive adjectives in APC and BPC

For example, the adjective *happy* was found in American articles five times but not in the British news. After looking at the context, it was found that the use of this positive adjective occurred only in direct speech of American articles. See example (43).

(43) I was super nervous before the match, during the match and I'm **happy** it's finished. A_7

Next, the adjectives ending in -less, such as *priceless*, *fearless*, *nerveless*, and *faultless*, appear in British news; even though these words have a negative suffix, their meaning is positive; see examples (44) – (47). In American articles, no adjective with a -less ending was detected. This might be an issue to focus on in the Large Corpus analysis.

(44) She would have risen to a new ranking of around sixth, but this is a **priceless** victory. B_7

(45) Yet it was unclear if he would be as **fearless** with so much on the line. B_9

(46) Throughout a **nerveless** first set, Kyrgios served extremely well. B_9

(47) Djokovic absorbed a **faultless** start from Nick Kyrgios before raising his level and gradually smothering his opponent. B_9

The American news, on the other hand, employs several adjectives starting with the prefix -un, such as *unseeded*, *unreturnable*, *unable*, *unforced*, *unbelievable*, and *untouched*. In contrast, such adjectives occurred only three times in the British news, namely in *unforced* and *unbelievable* (*unforced* appeared twice in the lexical expression *unforced errors*). These words might seem negative because the prefix -un stands for negation; however, most adjectives express positive meanings in the corpora, see examples (48), (49) and (50).

(48) The Australian broke Djokovic's service game and, behind his almost **unreturnable** serve, clinched the first set to settle his nerves. A_9

(49) William's high-risk, high-reward tactics left her with 61 winners in the match and more than 50 **unforced** errors. A_3

(50) He played an **unbelievable** return game at 2-1, breaking Kyrgios's serve to love after four otherworldly first-serve returns. B_9

So, according to the Pilot Corpus sample, we might suggest that Britons are more likely to use the suffix -less than the prefix -un, whereas Americans favours using prefix -un than the suffix -less. We will see later in the Large Corpus analysis if this presumption is correct.

5.1.4 Verbs

Focusing on verbs, it was discovered that the APC demonstrates a lower type-token ratio (TTR) in the use of verbs (34 per cent) compared to the BPC (42 per cent), the difference is here more notable than in the previous sections. When further analysing the verbs detected, it was found that the British Pilot Corpus contains 123 verbs used only

once (27 per cent), whereas the American Pilot Corpus has 101 verbs used only once (21 per cent). This finding supports the statement about higher TTR.

In both corpora, the verb *be* dominates; yet, in the APC, it was used a hundred and three times, whereas in BPC only sixty-six times. See Table 7 for the first twenty used verbs in BPC and APC. The table demonstrates that both corpora contain more or less the same verbs without significant distinctions.

American Pilot Corpus		British Pilot Corpus	
1. be	11. look	1. be	11. recover
2. have	12. make	2. have	12. know
3. do	13. know	3. say	13. rise
4. win	14. serve	4. do	14. become
5. say	15. ask	5. play	15. go
6. take	16. need	6. serve	16. break
7. play	17. get	7. win	17. fall
8. come	18. lose	8. take	18. achieve
9. break	19. see	9. give	19. think
10. show	20. begin	10. come	20. continue

Table 7 List of the first twenty used verbs in APC and BPC

In the BPC, twenty-six modal verbs tokens are detected, whereas, in APC, only twenty-one used. Modal verbs *can*, *would*, *could* and *will* appear in both corpora, while the modal verb *might* is detected three times in APC but is not found at all in BPC. On the other hand, the modal verb *must* is used in BPC but not used at all in APC; see examples (51) – (54).

(51) If I **can** win one game or two games, it would be really good for me. A_3

(52) Kyrgios has built a name for himself as a big match player who rises against the best, yet it was not clear if he **would** be as fearless with so much on the line. B_9

(53) Williams **might** come back to the Centre Court, but it won't be this year. A_3

(54) After his holiday, Kyrgios **must** answer to far more serious matters. B_9

The most used modal verb in the American Pilot Corpus is *would*, used seven times. In contrast, in the British Corpus, the most used modal verb is *will*, used eight times. Bonyadi (2011, 1) argues that based on the comparison of the American newspapers (The New York Times) and Persian English newspapers (Tehran Times), in terms of modal verbs, the predictive modal verbs such as *would* and *will* are the preferred ones. However, he mentions that a higher number of these predictive modal verbs occurred in NYT (The New York Times), indicating that the American writers of NYT are more interested in identifying what happens in the future.

Finally, looking at collocations of verbs *take* and *have* the following findings are detected. In American articles, *take* collocates with words like *lead*, *control* and *game*. In British articles, collocations such as: *take the set*, *take a step* or *take control* are noted. The verb *have* is, in American news, used with words such as *potential*, *team*, *weapon* or *idea*, whereas in British articles, phrases such as *have a drink*, *have a recognition*, *have a plan* or *have an interest* are found. So, it can be noticed that *take* collocates with words (*set*, *game*, *lead* and *control*) connected with tennis more in both corpora than the verb *have*.

5.1.5 Nouns

Analysing the nouns, focusing on TTR, it was discovered that the TTR of the British Pilot Corpus is relatively higher (48 per cent) than in the American Pilot Corpus. (41 per cent). Ten names had to be excluded to find out the twenty most frequently used nouns of the British Pilot Corpus, and eight names had to be excluded from the American Pilot Corpus. See Table 8 for the top twenty used nouns in the BPC and APC.

American Pilot Corpus		British Pilot Corpus	
1. set	11. time	1. set	11. shot
2. game	12. lead	2. week	12. time
3. title	13. tournament	3. point	13. serve
4. court	14. serve	4. match	14. break
5. tennis	15. crowd	5. year	15. single
6. world	16. point	6. court	16. career
7. match	17. no.	7. game	17. winner
8. break	18. year	8. player	18. thing
9. player	19. pair	9. tournament	19. return
10. centre	20. trophy	10. title	20. lever

Table 8 List of the first twenty most used nouns in APC and BPC

The tables demonstrate that both corpora include the most frequent nouns that are similar and have some connection to tennis or sport.

5.1.6 Spelling

Spelling differences were insignificant. One of the spelling differences was the adjective *favored* used in the American article, which confirms that American English has -or ending, whereas most British words have -our ending. This finding confirms the theory introduced by Darragh (2000, 3). The thesis's theoretical part on page 11 includes more information about the spelling differences.

A similar thing stands for -tre and -ter ending. British English newspapers prefer to use -tre ending, which was proven by the name *Centre Court*, a tennis court placed in London. So, most sports journalists used this name with British spelling in American and British articles because it is a proper name for one of the world's most famous tennis courts. However, American English prefers -er ending, and there was one instance of mentioning it in an American article, see example (55).

(55) "I was super nervous before the match, during the match and I'm happy it's finished," she told Sue Barker on Center Court. A_7

5.2 Large Corpus

Based on the information obtained in the Pilot study (Chapter 5.1 Pilot Corpus), it was decided to focus on the following features: keywords, verbs, nouns, grammar, adjectives and adverbs.

5.2.1 Keywords

First, it was decided to examine the *keywords* of each corpus. Keywords are single-token items which appear more frequently in the focus corpus than in the reference corpus. American and British Large Corpora were mutually used as a reference corpus, and twenty names had to be excluded from the list to determine the first twenty keywords in the American Large Corpus (ALC) and twenty-two names from the British Large Corpus (BLC). See Table 9 for the first twenty keywords of the American Large Corpus and British Large Corpus.

American Large Corpus		British Large Corpus	
1. semifinal	11. all-white	1. rubber	11. reminder
2. quarterfinal	12. pump	2. cancer	12. decisive
3. organizer	13. top-seeded	3. protester	13. badly
4. organization	14. defense	4. momentum	14. road
5. sexual	15. online	5. verdict	15. criticise
6. cardiac	16. goodbye	6. protest	16. defence
7. global	17. worry	7. fuller	17. continent
8. center	18. americans	8. coverage	18. immense
9. Olympian	19. virus	9. uncle	19. aspect
10. realize	20. organize	10. breast	20. clean

Table 9 List of the first twenty keywords of ALC and BLC

After looking at the table divided into twenty keywords of American and British Large Corpus, it can be noticed that there are visible distinctions in spelling.

The word **defence** occurred in both corpora, but in ALC, it was used with the letter "s", whereas in BLC, "c" was used, confirming the differences in spelling between

American and British English. Another example of American spelling can be seen in words such as *organizer*, *organization*, *organize* and *center*. On the other hand, the word **criticise** in BC was written in British spelling with the letter "s". These findings prove Svartik and Leech's assertion (2006, 154–155).

No other spelling differences were detected.

Additionally, the word *rubber* appears in first place in the British Large Corpora. After looking at a concordance, it was found that it means an *individual match*.

In conclusion, it can be seen that ALC has more words connected with tennis (*semifinal*, *quarterfinal*, *organizer*, *organization*, *center*, *top-seeded*, *defense*) than BLC (*rubber*, *verdict*, *coverage*, *defence*).

5.2.2 Verbs

After the calculation, the TTR (type-token ratio) of the verbs used in BLC is 9.5 per cent, whereas, in ALC, the TTR is lower, 8.9 per cent. So, there is a difference in verb frequency, but less significant than in the pilot study.

Subsequently, verbs were analysed, focusing on the spelling differences. At first, it was focused on -ize/-ise ending. The function world list in the Sketch Engine was used for the detailed analysis. See the following findings.

In the American Large Corpus, only two verbs: *realise* and *apologise*, were used with uncommon (British) spelling. This result complies with the fact that American English prefers to use the suffix -ize.

In the British Large Corpus, only three verbs with the -ize suffix were detected. Namely: *vandalize*, *specialize* and *depressurize*.

However, they were not part of direct speech. So, American spelling prevails in these sentences, but overall, it is not typical for the BLC. See Table 10 for the five most frequent verbs with -ize ending in the American Large Corpus and five most frequent verbs with -ise ending in the British Large Corpus, and examples (50) – (61).

American Large Corpus	British Large Corpus
1. realize	1. criticise
2. organize	2. finalise
3. symbolize	3. recognise
4. apologize	4. apologise
5. politicize	5. depressurise

Table 10 List of the five most frequent verbs with -ize ending in ALC and five most frequent verbs with -ise ending in BLC

(56) He **realizes** how fortunate he is. A_40

(57) In November, fellow inmates managed to **organize** three chocolate cakes. A_63

(58) We designed the champion's Billie Blue Jacket to **symbolize** your incredible win. A_46

(59) Often **criticised** for squandering his talents, Kyrgios reached the final. B_109

(60) Tennis Australia **finalises** mixed team event plans. B_33

(61) He is now reliably **recognised** on the streets of the UK. B_17

Consequently, the following numerical findings were obtained by examining the occurrences of modal verbs in both corpora. See Table 11 for the number of modal verbs in the American and British Large Corpus.

	American Large Corpus	British Large Corpus
Would	142 x	129 x
Will	233 x	269 x
Can	129 x	157 x
Might	34 x	15 x
Must	14 x	9 x
Could	80 x	97 x

Table 11 Number of modal verbs in ALC and BLC

The analysis reveals that the modal verb *will* is the most commonly used modal verb in both the American and British Corpora. However, this modal verb is more frequent in BLC, which complies with the result of the Pilot Corpora. On the other hand, *would* is more frequent in ALC than in BLC, which can also be noticed in the pilot study.

There is a relatively big quantitative difference in the frequency of *might*; no such distinction with other modal verbs is detected. Next, *might* is followed by the verb *be* in both corpora six times. See examples (62) – (63).

(62) Federer decided it **might be** nice to grab tea at the tennis club. A_59

(63) Davis Cup campaign but he **might be** enticed back to the team. B_61

Subsequently, the twenty most used verbs were detected, see Table 12 for the list of twenty most used verbs in American Large Corpus and British Large Corpus.

American Large Corpus		British Large Corpus	
1. be	11. know	1. be	11. think
2. have	12. come	2. have	12. make
3. say	13. think	3. say	13. see
4. do	14. tell	4. do	14. know
5. win	15. see	5. play	15. feel
6. play	16. want	6. win	16. give
7. go	17. reach	7. go	17. want
8. get	18. lose	8. get	18. reach
9. make	19. feel	9. take	19. lose
10. take	20. beat	10. come	20. beat

Table 12 List of the first twenty most used verbs in ALC and BLC

The twenty most used verbs are almost identical. Only one verb in each table differs: *tell* in ALC and *give* in BLC. Even though the articles are about tennis, a few verbs commonly related to tennis or sport are detected: *win*, *play*, *lose* and *beat*. Next, insignificant distinctions in frequency were found.

Focusing on the most frequent collocations of the verb *take* in ALC and BLC, it was discovered that the three most common collocations in British Large Corpus are: *take place*, *take care* and *take part*.

In comparison, collocations such as *take time*, *take place* and *take step* occurred most frequently in the American Large Corpus. See examples (64) – (67).

(64) Great Britain's tie against Colombia is one of 11 qualifying matches **taking place** this week. B_105

(65) The 29-year-old Garcia, who grew up in Lyon, **took time** to thank her home crowd. A_106

(66) **Taking** care of everything. B_17

(67) It is the reason we've **taken steps** in that direction, with a lot of important work still ahead. A_92

5.2.3 Nouns

As in the Pilot Corpora, the type-token ratio is more significant in the British (16 per cent) than American Large Corpus (15 per cent).

Concerning the spelling differences, the nouns were analysed. Again, the function word list was used in the Sketch Engine programme to analyse spelling distinctions of the words' endings. First, the endings -or/our were examined. It was found that both corpora contain more nouns with the -or endings, 44 in ALC and 43 in BLC. Nevertheless, a detailed analysis has yet to be made.

After examining the list of nouns that vary in AmE and BrE spelling, the following results were detected. In ALC, nouns with -or ending, such as *honor* (8x), *color* (5x), *favor* (4x), *behavior* (2x) and *parlor* (1x), were used. In comparison, nouns with -our ending, such as *favour* (1x), *behaviour* (1x) and *honour* (1x), were utilised. These nouns were used at the very least compared to the noun with -or ending. After looking at the concordances, it was discovered that all these three nouns were used in indirect speech.

In BLC, nouns ending with -our ending, such as *honour* (7x), *behaviour* (4x), *neighbour* (4x), *colour* (3x), *favour* (2x), *tumour* (1x), *demeanour* (1x), *armour* (1x), and *flavour* (1x) were found. In contrast, only three nouns with -or unusual American spelling were found: *honor* (2x), *labor* (2x), and *neighbor* (1x).

Next, the nouns with -ter and -tre ending were analysed. In ALC, words with the usual American spelling (-ter), such as, *center* (10x), *meter* (1x) and *centimeter* (1x), were detected. In BLC, the -tre ending was used with identical words as in ALC, like *centre* (33x), *centimetre* (1x) and *metre* (1x). The most interesting noun concerning the spelling is *center/centre*. It was used with the usual American spelling (-ter) ten times in ALC but

eighteen times with the unusual British spelling. After looking at the concordances, it was discovered that it was used in all examples as *Centre Court*, an official name for one of the world's most famous tennis courts.

Finally, among spelling differences that need to be noted is the noun with the American spelling *program*, used eight times in ALC but only once in BLC. The British spelling of the noun *programme* was found in BLC four times, but it was not detected in ALC. All these examples above comply with Darragh's theory (2000, 3,7).

Subsequently, the twenty most frequent nouns were detected. In order to identify the twenty most used nouns, eleven names had to be excluded from ALC and BLC.

See Table 13.

American Large Corpus		British Large Corpus	
1. tennis	11. sport	1. tennis	11. point
2. player	12. point	2. year	12. day
3. year	13. set	3. player	13. sport
4. court	14. game	4. court	14. game
5. match	15. champion	5. match	15. week
6. time	16. man	6. world	16. way
7. tournament	17. week	7. time	17. career
8. no. ¹	18. people	8. title	18. champion
9. world	19. career	9. set	19. team
10. title	20. woman	10. tournament	20. month

Table 13 List of the first twenty most used nouns in ALC and BLC

It can be noted that the first twenty nouns in the list are almost identical with very few different items, i.e., *people*, *man*, and *women* in ALC and *team*, *month* and *day* in BLC. These findings suggest that there is little difference between the American and British reports. Looking further at the list, it is evident that many nouns are related to tennis (*player*, *tournaments*, *set* etc.).

¹ number

The word *tennis* occurs four hundred eighty-seven times in ALC, while only four hundred and five times in BLC, one of the most significant frequency differences in the top twenty nouns. On the other hand, the word *team* appears ninety-one times in BCL but only fifty-one times in ALC; however, the difference is more negligible in this case.

After comparing the most frequent verbs and nouns, it can be noticed that mainly because of the quantitative difference, nouns are related to tennis or sport more than verbs. See Tables 12 and 13 for comparison.

5.2.4 Adjectives

After concentrating on the TTR of adjectives in both corpora, it was found that TTR in ALC is about 19 per cent, and in BLC, about 20. As in the pilot study, there is an insignificant quantitative difference. The most frequent adjective in both corpora is adjective *first*; however, it is slightly more frequent in ALC. Subsequently, in ALC 533, adjectives are used only once (9.8 per cent), whereas in BLC 620 are used only once (10.5 per cent), meaning that writers of the American news tend to repeat some adjectives more than British writers do.

After analysing the adjectives in terms of spelling differences, it was found that the most considerable distinction is with the adjective *favourite/favorite*. The *favorite* was used six times in the American Large Corpus; meanwhile, it was not used in the Large British Corpus. In contrast, the *favourite* occurred fourteen times in the Large British Corpus and only four times in the American Large Corpus and from that, only one time was used in indirect speech. These results comply with the theory of Svartvik and Leech (2006).

Subsequently, the adjectives were examined, focusing on the adjectives with -less ending. In BLC, seventeen adjectives with the suffix -less were located, whereas, in ALC, only eight adjectives were found. See Table 14 for identical instances of adjectives with the -less suffix and their collocations in each corpus.

American Large Corpus	British Large Corpus
1. faultless - integrity (1x)	1. faultless - integrity, start (4x)
2. endless – queue (1x)	2. endless - service, toolbox (2x)
3. helpless - she felt (helpless) (1x)	3. helpless – cramping (1x)
4. countless – millions (1x)	4. countless - times, jokes (2x)
5. baseless - allegations (1x)	5. baseless – allegations (1x)

Table 14 Instances of identical adjectives with the -less suffix in ALC and BLC

As a result, the number of instances can suggest that British writers use adjectives with -less suffixes more often than American writers.

Next, the adjectives were analysed, concentrating on the prefix -un. Forty-seven adjectives starting with -un were located in ALC; in comparison, forty adjectives with the prefix -un occurred in BLC. However, BLC contains twenty-two adjectives used more than once, while ALC comprises only thirteen. See Table 15 for the five most frequent adjectives with the prefix -un of each corpus.

American Large Corpus	British Large Corpus
1. unseeded	1. unable
2. unforced	2. unforced
3. unable	3. unbelievable
4. unbelievable	4. unvaccinated
5. unvaccinated	5. unseeded

Table 15 The five most frequent adjectives with the prefix -un in ALC and BLC

The table demonstrates that the five most frequent adjectives with the prefix -un are identical in both corpora; however, the difference is that they take a different place in frequency, except for the adjective *unforced*, which takes second place in both corpora.

5.2.5 Adverbs

The quantitative difference between adverbs is negligible in both corpora, TTR being in ALC at 9.5 per cent and 9.7 per cent in BLC. After looking at the list of adverbs, the five of the most used adverbs were located. See Table 16 for the five most frequent adverbs in American and British Large Corpus.

American Large Corpus	British Large Corpus
1. not	1. not
2. so	2. just
3. just	3. so
4. also	4. now
5. really	5. also

Table 16 The five most frequent adverbs in ALC and BLC

It can be noticed that the top five adverbs are almost identical; they take only different positions in terms of frequency. In the first place is the adverb, *not* in both corpora; however, it was used 55 times more in ALC than in BLC.

As in the Pilot Corpora, the dominance of adverbs ending with *-ly* is in the British Large Corpus, which comprises 236 adverbs with the suffix *-ly*, whereas the American Large Corpus includes only 200. After analysing the list of adverbs in more detail, it was found that the adverbs such as *extremely*, *barely*, and *initially* dominate quite significantly in the British Large Corpus. The adverb *extremely* was found twenty-two times in BLC, whereas only five times in ALC, *barely* was located fourteen times in BLC, while only three times in ALC and *initially* was detected eight times in BLC but only three times in ALC.

5.2.6 Grammar

Finally, the corpora were analysed in grammar, more precisely in the number of past simple and present perfect uses. The following results were found thanks to the function concordance and use of CQL language. See Table 17 for the results.

	American Large Corpus	British Large Corpus
Present perfect	510	616
Past simple	2 475	2 412

Table 17 Distribution of present perfect and past simple in ALC and BLC

It can be seen that, as in the pilot study, the present perfect tense is more frequent in BLC (8.5 per cent) than in ALC (7.3 per cent). On the other hand, past simple is used more often in ALC (35 per cent) than in BLC (33 per cent). Both of these findings conform to the theory claimed by Darragh (2000), who argues that Americans use these two tenses interchangeably, while in some conditions, only the present perfect tense can be used in British English.

Conclusion

In conclusion, this bachelor thesis aimed to identify differences in spelling and grammar between American and British English in tennis articles. Subsequently, it sought to determine how the vocabulary differs in American and British tennis reports, mainly focusing on distinctions in adverbs, adjectives, verbs and nouns.

The theoretical part provided basic information about the English language and corpus linguistics. It also delivered the overall overview of fundamental differences between American and British English. The following distinctions between American and British English were detected in the practical part.

Firstly, spelling differences were the most significant in the use of nouns, especially in the suffixes: *ter/tre* or *our* and endings in the word *program/programme*. It was found that in both American and British English prevails, the typical American or British spelling only with few exceptions, so the theories stated in the theoretical part were confirmed. Also, there were minor differences in keywords of the Large Corpora, such as *defence/defense*, *organization*, *organizer*, *center* and *critisice*. Finally, there were negligible spelling differences in the use of adjectives and verbs.

Subsequently, one of the most significant differences between American and British English was detected in grammar, specifically in using the present perfect and past simple—the present perfect dominated both Pilot and Large Corpora with clear superiority in British English, agreeing with the theoretical part. Conversely, past simple was more frequent in American English but only in Large Corpora, while in Pilot Corpora, the past simple dominated British English.

Additionally, the following differences between American and British English in vocabulary are as follows. The analyses of adverbs showed that British authors tend to use adverbs with the suffix *-ly* more often than American writers, but overall, adverbs were more or less identical.

Next, the adjectives were similar in number and meaning; however, the research showed that adjectives with the suffix *-less* were more frequent in British English in both Pilot and Large Corpora. On the other hand, adjectives with the prefix *-un* occurred more often in American English in Pilot and Large Corpora.

Furthermore, there was quite a significant difference in the use of verbs in Pilot Corpus, where verbs dominated British English. Nevertheless, this distinction was less dominant in Large Corpus. Also, the modal verbs were analysed, detecting that *will* was most used in American and British English in both Pilot and Large Corpora. However, the most noteworthy difference was the use of *might* dominating in American English.

In conclusion, analysing and comparing verbs and nouns in Large Corpora revealed that nouns are related to tennis more than verbs. This statement was detected mainly due to the quantitative difference.

References

- Algeo, John. 2006. *British or American English? A Handbook of Word and Grammar Patterns*. Studies in the English Language. Cambridge; New York: Cambridge University Press.
- Baker, Paul. 2017. *American and British English: Divided by a Common Language?* Cambridge, England; New York: Cambridge University Press.
- Bonyadi, Alireza. 2011. "Linguistic Manifestations of Modality in Newspaper" 3: E30.
- Crystal, David. 1995. *The Cambridge Encyclopedia of the English Language*. Cambridge [England]; New York: Cambridge University Press.
- Čermák, František. 2017. *Korpus a korpusová lingvistika*. Vydání první. Praha: Univerzita Karlova, Nakladatelství Karolinum.
- Čermák, František, Patrick Corness, Aleš Klégr, Eva Hajičová, and Alena Macurová. 2010. *InterCorp: Exploring a Multilingual Corpus*. Studie z Korpusové Lingvistiky 13. Praha: NLN, Nakladatelství Lidové noviny Ústav Českého národního korpusu.
- Dalton, Christiane, and Barbara Seidlhofer. 2000. *Pronunciation*. 1. publ., 3. impr. Language Teaching. Oxford: Oxford Univ. Pr.
- Darragh, Glenn. 2000. *A to Zed, A to Zee: A Guide to the Differences between British and American English*. 1st ed. Irun, España: Editorial Stanley.
- Davies, Christopher. 2007. *Divided by a Common Language: A Guide to British and American English*. 1st Houghton Mifflin pbk. ed. Boston: Houghton Mifflin.
- Greenbaum, Sidney. 1996. *The Oxford English Grammar*. London: Oxford University Press.
- Han, Ligang. 2019. "A Review of the Major Varieties of English Language." *International Education Studies* 12 (January): 93. <https://doi.org/10.5539/ies.v12n2p93>.
- Kettunen, Kimmo. 2014. "Can Type-Token Ratio Be Used to Show Morphological Complexity of Languages?" *Journal of Quantitative Linguistics* 21 (3): 223–45. <https://doi.org/10.1080/09296174.2014.911506>.
- Kilgarriff Adam, Baisa Vít, Bušta Jan, Jakubíček Miloš, Kovář Vojtěch, Michelfeit Jan, Rychlý Pavel, Suchomel Vít. The Sketch Engine: ten years on. *Lexicography*, 1: 7-36, 2014.

McEnery, Tony, and Andrew Hardie. 2012. *Corpus Linguistics: Method, Theory and Practice*. Cambridge Textbooks in Linguistics. Cambridge; New York: Cambridge University Press.

Rapp, Brenda. 2001. *The Handbook of Cognitive Neuropsychology: What Deficits Reveal about the Human Mind*. Philadelphia: Psychology Press.

Scott, Mike, and Chris Tribble. 2006. *Textual Patterns: Key Words and Corpus Analysis in Language Education / Mike Scott and Christopher Tribble*. Studies in Corpus Linguistics, v. 22. Amsterdam; Philadelphia: John Benjamins Pub.

Svartvik, Jan, and Geoffrey N. Leech. 2006. *English: One Tongue, Many Voices*. Houndmills, Basingstoke, Hampshire [England]; New York: Palgrave Macmillan.

Swan, Michael. 2016. *Practical English Usage*. Fourth edition. Oxford, United Kingdom: Oxford University Press.

Thomas, Dax. "Type-token Ratios in One Teacher's Classroom Talk: An Investigation of Lexical Complexity." (2005).

Webster, Noah. 1998. *The Elementary Spelling Book: Being an Improvement on the American Spelling Book*. Rev. and Reprinted. United States: Weisman Publications.

Appendices

Appendix 1: Overview of all article

Appendix 2: List of the American and British press sources

American sources	British sources
U.S. News	The Guardian
CNN Sports	Metro News
The New York Times	Daily Mail
The Wall Street Journal	BBC News
USA Today	The Independent
NBC News	Mirror
ABC News	Reuters
Fox News	GB News
AP News	The Week
The Washington Post	Express
ESPN	
Tennis World	
Nine.com	

Table 18 List of the American and British press sources