A Small-Scale Corpus-Based Study for The Varietal Differences in American and British English: Implications for Language Education

Ibrahim Halil Topal*

Abstract
This small-scale corpus-based study delineates the most common and significant dialectal variations between the two most commonly spoken English varieties: American English (AmE) and British English (BrE). As a result of the corpus analysis, four main areas have emerged as to where dialectal variations take place: pronunciation, vocabulary, grammar, and orthography/punctuation. A total of 26 variations (\(f=10\) in pronunciation, \(f=5\) in vocabulary, \(f=6\) grammar, and \(f=5\) in orthography/punctuation) was identified by analyzing a variety of sources, including books, articles, online dictionaries, and websites. The significance of the variations in the abovementioned language areas and their implications for language teaching were discussed empirically and pedagogically. Notwithstanding the limitations, the research is expected to contribute to our understanding and awareness of the dialectal variations and assist language learners and teachers with the learning and teaching of these variations pedagogically and systematically since it might serve as a guide or a framework of reference.

Keywords: language variety, American English, British English, corpus, language skills

* Corresponding Author: MA Gazi University, College of Foreign Languages; Email: ibrahimtopal@gazi.edu.tr, ORCID ID: https://orcid.org/0000-0003-4220-3706
1. Introduction

With around 1.35 billion native and nonnative speakers worldwide (Szmigiera, 2021), it is well-established that English is the most prevalent international language. It is used in manifold spheres of life (Rao, 2019), from communication on social media (Lantz-Andersson, 2018), to global business (Neeley, 2012) and science (Elnathan, 2021). Even, most scientific publications are in English, guaranteeing its monopoly in the academia (Di Bitetti & Ferreras, 2017). This predominance also manifests itself in the field of language education.

As Beare (2019) narrates, the British Council maintains that the number of English language learners are nearly 1.13 million (750 m EFL and 375 m ESL learners), with a conceivable growth in the future. It goes without saying that almost a vast majority of educational materials are also published in English (Rose & Galloway, 2019), especially in American English (AmE) and British English (BrE), indicating their hegemony (Young et al., 2016). The fact that native varieties of English are preferred by teachers (He, 2017) and learners (Rindall, 2014), and are used in learning and teaching materials (Rose & Galloway, 2019) highlights the significance of the two prominent varieties: AmE and BrE. But what is a variety?

A variety, also known as language variety, is a sociolinguistic concept that is hard to define but refers to “any system of linguistic expression whose use is governed by situational variables” (Crystal, 2008, p.509). It is also used interchangeably with a dialect defined as “a regional variety of language distinguished by features of vocabulary, grammar, and pronunciation from other regional varieties and constituting together with them a single language” (Merriam-Webster, n.d.). Some common English varieties are AmE, BrE, Canadian English, Australian English, and New Zealand English (Crystal & Potter, 2020). Of these, AmE is estimated to be spoken by approximately 329 million speakers (U.S. Census Bureau, 2019), while BrE speakers are reckoned to be around 67 million (Office for National Statistics, 2021), making both the most widely spoken English varieties. Another point that these two varieties have in common is that they are considered the standard forms of the English language, or in other words, standard English referring to “a prestige variety of language used within a speech community” (Crystal, 2008, p.450). Given the large number of varieties and dialects existing in the English language (Grieve et al., 2019), its standardization becomes crucial in that “standard languages/dialects/ varieties cut across regional differences, providing a unified means of communication, and thus an institutionalized norm
which can be used in the mass media, in teaching the language to foreigners, and so on” (Crystal, 2008, p.450).

Given the prevalence of AmE and BrE varieties in language education contexts and teaching materials (Rose & Galloway, 2019), along with the excessive exposure to these two on the internet and social media (Johnson, 2021), it becomes essential that language learners and teachers be aware of the differences between the two. In this sense, this small-scale corpus-based research aimed to present a comprehensive description of the varietal differences between the two most prevalent English varieties for pedagogic purposes.

Previous studies focused on different aspects of these two varieties. For instance, in their corpus-based study, Mindt and Weber (1989) analyzed the distribution of prepositions in two varieties and revealed, by and large, a close distributional correlation. In another study examining the stress patterns in both varieties, Berg (1999) found divergence with stress placement stemming from lexical type, word endings, compounds, and zero-derived items. Scott (2000) examined the differences in AmE and BrE vocabulary and urged the teaching of the differences for successful business interactions. In another study, Scott (2004) explored the AmE and BrE spellings in business-related documents and provided some teaching ideas to expedite the understanding of these spelling variations in the mentioned field. Bock et al. (2006) investigated the verb number agreement patterns and indicated convergence with sensitivity to variations and divergence with number agreement for the usage of collective nouns. In the same year, Tottie and Hoffmann (2006) looked into tag questions in AmE and BrE and found differences in polarity types, operators, and pragmatic functions. Hundt and Smith (2009) inspected the use of present perfect and simple past tenses and concluded that present perfect was employed more frequently in BrE than in AmE. In a recent study, Carrie (2017) explored the attitudes of Spanish university students and revealed a strong bond toward AmE accent and a tendency towards BrE accent in terms of emulation.

Considering the cited studies, research on AmE and BrE have mainly focused on the examination of grammar (Mindt & Weber, 1989; Bock et al., 2006; Tottie & Hoffmann, 2006; Hundt & Smith, 2009), vocabulary (Scott, 2000), pronunciation (Berg, 1999; Carrie, 2017), and orthography/punctuation (Scott, 2004) in the two varieties. However, very few studies outlining the differences between the two cited varieties are present to serve as a teaching framework. To this end, the present
research is aimed at delineating the most fundamental and significant variations between AmE and BrE with regard to pronunciation, grammar, vocabulary, and orthography/punctuation. It is, therefore, hoped to contribute to raising learner and teacher awareness of AmE and BrE in learning and teaching the previously mentioned language areas.

2. Literature Review

2.1. A Historical Perspective on BrE

English pertains to the West Germanic branch of the Indo-European language family (Crystal & Potter, 2020). Divided into three main historical periods (i.e., old, middle, and modern), English has been influenced by many other languages based on the invasions of other countries. The Celts were the ancient tribes living in Britain for a long time until the Romans arrived. Various Celtic languages were therefore spoken on the island. Some examples of words of Celtic origin in modern-day English are *ambiacotos* (ambassador), *wlana* (flannel), and *carpentom* (carpenter). Until the 4th century, the Romans were in power in Britain for about four centuries, which meant the domination of Latin in Britain during this period. Words such as *vinum* (wine), *balteus* (belt), and *candela* (candle) were of Latin origin. In the 5th century, three Germanic tribes (i.e., Angles, Saxons, and Jutes) occupied Britain through Northern Germany – what is now Denmark, and expelled the prevailing Celtic (a.k.a. oldest language in the UK) dwellers in Scotland and Ireland.

This led many Norse words (e.g., *agnra* (anger), *qfugr* (awkward) to enter the English language. Owing to such types of invasions, the English language is made up of a variety of dialects. In the 10th century took place the Norman invasion – the French conquest of the island. The 11th century bore witness to conversations in French between the upper-class and in English between the lower class. The French invasion brought such words as *coroune* (crown) and *castel* (castle) into the English language. This period was called as Middle English. A momentous event that occurred during this period was the Great Vowel Shift, leading to a great change in long vowels. Later in the 14th century, English took control of Britain. This predominance of English was followed by the birth of Modern English period in the 15th century, leading to the emergence of new vocabulary and grammar. In the early modern period, Shakespeare alone introduced around two thousand words, such as *bandit* (Henrey VI, Part 2, 1594) and *critic* (Love’s Labour’s Lost, 1598). The
period from the early 16th to the 17th century was called the Elizabethan era, during which the culture (i.e., art, music, literature, and theater) flourished. The Industrial Revolution, which took place from the mid-18th century to early 19th century, led to the creation of new technical words (e.g., telegraph and electricity). Britain was an empire between the 18th and 20th centuries, ruling over a large majority of the world that brought with itself the coinage of numerous linguistic items, such as kangaroo (from Australian native word gaNurru) and jungle (from Hindi word jangal). The dominion of the British Empire reached to many parts of the world, including North America. This issue was discussed further in the following part.

2.2. A Historical Perspective on AmE

The first documented European colony (Vinland, what is now Newfoundland) to the Americas was established by the Norsemen led by the Icelandic explorer Leif Erikson circa 1000 AD (Wonders, 2003). Later came the expedition of Christopher Columbus endorsed by the Spanish monarchs. The Americas were inhabited by native Americans at the time. A century later following the Columbus’s voyages started the English invasion in Virginia, followed by the colonization of other parts of North America (13 coastal states in the Atlantic) between the 17th and 18th centuries. Meanwhile, the Anglo-Dutch war left North America divided between the British, Spanish, and French empires. In the War of Independence that took place in 1175, many North American colonies gained their independence from the British Empire with the help of the Spanish and French.

All these developments formed what is called the Colonial Period (www.encyclopedia.com). It might be summarized that there were imperial clashes for the colonization of the Americas among the Spanish, Portuguese, French, Dutch, and British empires between the 16th and 18th centuries (Harvey & Rivett, 2017). This two-century period undoubtedly left profound impacts on the linguistic diversity in the Americas. The Dutch invasions, for instance, resulted in the entrance of one-percent Dutch vocabulary into English (Lugarić, 2015). Mentioning the Dutch settlers, van der Sijs (2009) holds that “in their English they retained many Dutch words, some of which were adopted by English speakers, and eventually these became part of the American English vocabulary” (p. 48). Some Dutch words used in modern-day American English are waffle, cookie, and noodles (Lugarić, 2015). The Spanish also influenced the linguistic diversity of English,
with such words as tequila, tortilla, and Los Angeles entering the language. Though mostly impacted the history of Britain, the French had significant effects on the English language. The British Empire invaded North America and left certain traces (including linguistic ones) there. It is no wonder that some of these traces might be of French origin. Words such as occasion, novel, and ricochet are the presents from the French language among many others. In fact, around 30% of the words in English are claimed to be of French origin (Williams, 1986), though there no consensus on the exact amount.

The period between the late 1700s and 1800s is known as the National Period, during which linguistic nationalism was promoted by excluding foreign impacts from the uniform AmE (www.encyclopedia.com). The settlement of the West, the expansion of railroads, “the growth of industry, the labor movement, the invention of the telegraph and telephone, the burgeoning of journalism, the expansion of education at all levels, and the publication of textbooks and dictionaries were all factors” (n.d.) that shaped the language of Americans during the 19th century (www.encyclopedia.com). From the late 1800s onward is called the International period, when Americans were actively involved in the global politics and economics, creating significant impacts on AmE usage (www.encyclopedia.com). Meanwhile, several important events including the Hollywood movie industry, post-war technological advancements, and the activities and products of major US publications have contributed to the dissemination of ‘Americanism’ all over the world (www.encyclopedia.com).

2.3. The Use of Corpora in Language Learning and Teaching

According to leading experts in corpus linguistics, a corpus is a collection of real language (either written or spoken) assembled for a specific purpose (Flowerdew, 2012). Despite the assertions that corpora should include millions of words, particularly with the advent of computerized corpora (Sinclair, 1991), smaller corpora might be used to investigate more frequent language aspects such as grammatical structures (Flowerdew, 2012) and corpora that are pedagogically important for teaching certain vocabulary (Braun, 2005). Biber (1990) emphasizes that smaller corpora fit best for such purposes. Römer (2011, p.207), on the other hand, provides a framework for the pedagogical utilization of corpora that consists of direct and indirect applications, wherein the first one is hands-on for learners and teachers that affect teaching syllabus, reference works, and teaching materials while
the latter is hands-on for researchers and materials writers influencing teacher- and learner-corpus interactions. Braun's (2005), Biber's (1990), and Römer's (2011) indirect applications are used as a starting point for this research.

Relevant literature provides abundant examples of utilizing corpora for vocabulary. The Academic Word List (AWL) developed by Coxhead (2000) may be the most well-known academic wordlist comprising 570 word families and 3000 terms that are regularly found in written academic materials, and it attempts to help students accomplish their academic goals. However, corpus-based studies were also conducted in other language skills as well. A study, for example, investigated the learners’ attitudes toward using concordancing for grammar teaching and revealed its potential for text writing (Vannestål & Lindquist, 2007). As regards pronunciation, Levis and Cortes (2008) investigated minimal pairs in spoken corpora and found four categories that define them in general and suggested these for future work. In her corpus-based study, Larsson (2012) found that Swedish, Bulgarian, and Italian college students preferred BrE spelling to AmE one. Given these studies, it is plausible to claim that there is a need for a descriptive corpus-based study exploring the AmE and BrE variations in pronunciation, grammar, vocabulary, and orthography/punctuation.

3. Method

3.1. Research Design

The current study utilizes corpus linguistics as the research methodology. Corpus linguistics, according to Kennedy (2001), “encompasses the compilation and analysis of collections of spoken and written texts as the source of evidence for describing the nature, structure, and use of languages” (p. 2816). Corpus-based studies conventionally contain quantitative descriptions of languages with information about the frequency of certain linguistic areas in particular contexts. Despite a predominant disposition toward the use of electronic resources, corpora might vary in size and design (Kennedy, 2001). In addition to general corpora, there are also specialized corpora referring to the corpora of “texts of a particular type, such as newspaper editorials, geography textbooks, academic articles in a particular subject, lectures, casual conversations, essays written by students etc.” (Hunston, 2002, p.14). The corpus employed in the present study is a specialized one comprising the dialectal variations in four language areas (i.e., vocabulary,
pronunciation, grammar, and orthography/punctuation). Amongst the diverse use of corpora, corpora in this study were employed for language learning and teaching because of the possible pedagogical benefits explored in earlier studies (Salsbury & Crummer, 2008).

3.2. Criteria and Guidelines for Creating a Specialized Corpus

The criteria for selecting a corpus structure should be few, clearly distinct from one another, and effective as a group in defining a corpus typical of the language or variation under study (Sinclair, 2005). According to McEnery and Xiao (2014), a balanced corpus includes various text categories reflective of the language variety under study. The following broad criteria for corpus selection were developed based on Sinclair's (2004) recommendations: (i) text mode (written), (ii) text type (i.e., websites, articles, and books), (iii) text domain (academic), (iv) text language (English), and (v) representativeness (varieties). Furthermore, Flowerdew's (2004, pp. 25-27) recommendations for creating a special-purpose corpus addressing the following issues were also considered in the formation of the corpus: (1) What is the purpose for building a specialized corpus? (2) What genre is to be investigated? (3) How large should the specialized corpus be? (4) Is the specialized corpus representative of the genre? (5) How will data be collected? (6) How will the specialized corpus be tagged/marked up? (7) What kind of reference corpus would be suitable to contrast with the specialized corpus?

3.3. Procedures for Compiling the Corpus

Following the criteria and guidelines aforementioned, all variations between AmE and BrE were searched through various sources, including but not limited to websites, blogs, journal articles, and books. The collected data were transferred to an electronic file and then analyzed in terms of content. The analysis revealed four areas in which dialectal variations take place: vocabulary, pronunciation, grammar, and orthography/punctuation. The irrelevant and infrequent variations were removed and the remaining ones were grouped into the four areas above-mentioned. The acquired differences were then cross-checked in the searched sources and online dictionaries to determine the ultimate version of the variations. Upon finalizing the categories, examples were provided for each instance using various sources such as online dictionaries, books, and websites. The summary of the
procedures was illustrated in Figure 1 below.

**Figure 1.**
The Procedures for Creating the Special-Purpose Corpus

<table>
<thead>
<tr>
<th>Step</th>
<th>Procedure</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Online search</td>
</tr>
<tr>
<td>2</td>
<td>Electronic file transfer</td>
</tr>
<tr>
<td>3</td>
<td>Content analysis</td>
</tr>
<tr>
<td>4</td>
<td>In-source check</td>
</tr>
<tr>
<td>5</td>
<td>Ultimate categorization/check</td>
</tr>
</tbody>
</table>

4. Variations between AmE and BrE

Following the corpus collection and finalization, the most common variations between the two varieties in relation to pronunciation, vocabulary, grammar, and orthography/punctuation were presented below. The examples were gleaned from multiple sources, including books, articles, online dictionaries, and the internet.

4.1. Phonetic/Phonological Variations

Variations exist between the two varieties at the segmental and suprasegmental levels discussed successively. One of the biggest articulatory difference between AmE and BrE is the (non)-articulation of the sound “r” (also referred to as (non) rhoticity). This sound is not pronounced when (i) a word ends in “r” that comes after a vowel, and there is no other sound after “r” (e.g. brother), and (ii) “r” appears before a consonant and after a vowel (e.g. bird). Further examples were provided in Table 1 below:
Table 1

**Different Pronunciations of “r” in AmE and BrE**

<table>
<thead>
<tr>
<th>Word</th>
<th>AmE pronunciation</th>
<th>BrE pronunciation</th>
</tr>
</thead>
<tbody>
<tr>
<td>air</td>
<td>/ə/</td>
<td>/ɛ/</td>
</tr>
<tr>
<td>creature</td>
<td>/ˈkriːtʃər/</td>
<td>/ˈkriːtʃər/</td>
</tr>
<tr>
<td>further</td>
<td>/ˈfɜːrðər/</td>
<td>/ˈfɜːrðər/</td>
</tr>
<tr>
<td>important</td>
<td>/ɪmˈpɔːrtənt/</td>
<td>/ɪmˈpɔːrtənt/</td>
</tr>
<tr>
<td>teacher</td>
<td>/ˈtiːtʃər/</td>
<td>/ˈtiːtʃər/</td>
</tr>
</tbody>
</table>

In another phonological variation, BrE uses a back open rounded sound /ɒ/, while AmE uses a back open unrounded sound /ɑ/ instead. Examples are present in Table 2 below:

Table 2

**Sound Shift /ɒ/ to /ɑ/ in AmE and BrE**

<table>
<thead>
<tr>
<th>Word</th>
<th>AmE pronunciation</th>
<th>BrE pronunciation</th>
</tr>
</thead>
<tbody>
<tr>
<td>follow</td>
<td>/ˈfɑː.loʊ/</td>
<td>/ˈfɒ.əʊ/</td>
</tr>
<tr>
<td>lost</td>
<td>/lɑːst/</td>
<td>/lɒst/</td>
</tr>
<tr>
<td>shop</td>
<td>/ʃɑːp/</td>
<td>/ʃɒp/</td>
</tr>
<tr>
<td>watch</td>
<td>/wɑːtʃ/</td>
<td>/wɒtʃ/</td>
</tr>
<tr>
<td>what</td>
<td>/wɑːt/</td>
<td>/wɒt/</td>
</tr>
</tbody>
</table>

A third phonological variation occurs when a group of words including the spelling of “arr” in them are pronounced with a /æ/ sound in BrE and /ɛ/ in AmE. Some examples are available in Table 3.

Table 3

**Sound Shift /æ/ to /ɛ/ in AmE and BrE**

<table>
<thead>
<tr>
<th>Word</th>
<th>AmE pronunciation</th>
<th>BrE pronunciation</th>
</tr>
</thead>
<tbody>
<tr>
<td>arrogant</td>
<td>/ˈer.ə.gənt/</td>
<td>/ˈer.ə.gənt/</td>
</tr>
<tr>
<td>carry</td>
<td>/ˈker.i/</td>
<td>/ˈker.i/</td>
</tr>
<tr>
<td>garrison</td>
<td>/ˈɡer.ə.sən/</td>
<td>/ˈɡer.ɪ.sən/</td>
</tr>
<tr>
<td>marry</td>
<td>/ˈmar.i/</td>
<td>/ˈmær.i/</td>
</tr>
<tr>
<td>narrative</td>
<td>/ˈnɛr.ə.tɪv/</td>
<td>/ˈnær.ə.tɪv/</td>
</tr>
</tbody>
</table>

Fourth, the open-mid central unrounded sound (i.e., /ɜ:/) in BrE turns into a high central lax vowel (i.e. /ɜ:/) in AmE, some examples of which are present in Table 4. Common spelling variations of these sounds include “er”, “ear”, “ur”, “ir”, and “or”.

Common spelling variations of these sounds include “er”, “ear”, “ur”, “ir”, and “or”.

**Language Related Research** 13(3), (July & August 2022) 199-226
As a fifth variation, words in BrE that include the sound /ɪə/, with such spellings as “ea”, “ee”, “ei”, and certain “e”, are pronounced with /ɪr/ sound in AmE.

The phonological variation between AmE and BrE also manifests itself in diphthongs, wherein /əʊ/ sound in BrE occurring in such spellings as “o”, “ol”, “oa”, “ow”, and “ou” are pronounced as /oʊ/ in AmE.

Dialectal variation also occurs when the /t/ sound comes between two vowels, usually in word-medial positions in BrE is frequently pronounced as /t/ in AmE in connected speech, which is one of the most distinguishing phonological features between the two varieties.
Table 7

<table>
<thead>
<tr>
<th>Word</th>
<th>AmE pronunciation</th>
<th>BrE pronunciation</th>
</tr>
</thead>
<tbody>
<tr>
<td>computer</td>
<td>/kəmˈpjuː.tɚ/</td>
<td>/kəmˈpjuː.tə/</td>
</tr>
<tr>
<td>daughter</td>
<td>/ˈdə.tɚ/</td>
<td>/ˈdə.tə/</td>
</tr>
<tr>
<td>hospital</td>
<td>/ˈhaː.спι.təl/</td>
<td>/ˈhos.pι.təl/</td>
</tr>
<tr>
<td>matter</td>
<td>/ˈмаt.ə/</td>
<td>/ˈmaet.ə/</td>
</tr>
<tr>
<td>water</td>
<td>/ˈwɔː.tɚ/</td>
<td>/ˈwɔː.tə/</td>
</tr>
</tbody>
</table>

In BrE, /j/ appears after alveolar consonants (i.e., /t/, /d/, /n/, /l/, /s/, and /z/) and is omitted in AmE, which is often referred to as yod (/j/) dropping.

Table 8

<table>
<thead>
<tr>
<th>Word</th>
<th>AmE pronunciation</th>
<th>BrE pronunciation</th>
</tr>
</thead>
<tbody>
<tr>
<td>due</td>
<td>/duː/</td>
<td>/dʒuː/</td>
</tr>
<tr>
<td>endure</td>
<td>/ɪnˈdʊr/</td>
<td>/ɪnˈdʒʊə/</td>
</tr>
<tr>
<td>mature</td>
<td>/məˈtɚ/</td>
<td>/məˈʃʊə/</td>
</tr>
<tr>
<td>new</td>
<td>/ˈnjuː/</td>
<td>/ˈnjuː/</td>
</tr>
<tr>
<td>tune</td>
<td>/ˈtuːn/</td>
<td>/ˈtjuːn/</td>
</tr>
</tbody>
</table>

Certain words that end in “-ary, -ery, and –ory” tend to be reduced by way of the schwa sound (i.e., /ə/) in BrE and prolonged by using the /ɔː/ and /e/ sounds in AmE, as seen in Table 9.

Table 9

<table>
<thead>
<tr>
<th>Word</th>
<th>AmE pronunciation</th>
<th>BrE pronunciation</th>
</tr>
</thead>
<tbody>
<tr>
<td>inventory</td>
<td>/ˈɪnv.ən.tə.rɪ/</td>
<td>/ˈɪnv.ən.trɪ/</td>
</tr>
<tr>
<td>monastery</td>
<td>/ˈmə.nə.stər.i/</td>
<td>/ˈmə.nə.strɪ/</td>
</tr>
<tr>
<td>preparatory</td>
<td>/ˈpre.ə.tə.rɪ/</td>
<td>/ˈpre.ə.trɪ/</td>
</tr>
<tr>
<td>secondary</td>
<td>/ˈsek.ən.dər.i/</td>
<td>/ˈsek.ən.dri/</td>
</tr>
<tr>
<td>secretary</td>
<td>/ˈsek.rə.tər.i/</td>
<td>/ˈsek.rə.trɪ/</td>
</tr>
</tbody>
</table>

Some words, particularly those of French origin, are emphasized differently in AmE and BrE, with Americans and Britons emphasizing the first or last syllable, the examples of which can be seen in Table 10.
4.2. **Lexical Variations**

An estimated of 4000 words in BrE are not used or have different meanings in AmE (Davies, 2005, p. 1). The use of different words exists in many domains from education (e.g., proctor-invigilator) to hospitality (e.g., front desk-reception) (Davies, 2005). Some examples are visible in Table 11.

**Table 11**

<table>
<thead>
<tr>
<th>Variations in Individual Words of Diverse Domains in AmE and BrE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>AmE</strong></td>
</tr>
<tr>
<td>fall</td>
</tr>
<tr>
<td>cookie</td>
</tr>
<tr>
<td>hood</td>
</tr>
<tr>
<td>pacifier</td>
</tr>
<tr>
<td>movie</td>
</tr>
<tr>
<td>yard</td>
</tr>
<tr>
<td>vacation</td>
</tr>
<tr>
<td>mail</td>
</tr>
<tr>
<td>railroad</td>
</tr>
<tr>
<td>candy</td>
</tr>
<tr>
<td>faucet</td>
</tr>
<tr>
<td>flashlight</td>
</tr>
</tbody>
</table>

Another lexical difference between AmE and BrE is in the use of verbs *have* and *take*, wherein Americans prefer *take* while Britons prefer *have* in certain expressions, such as *have a bath* and *take a bath*. Further examples are: *have/take a break, have/take a shower, have/take a rest, have/take a holiday, have/take a walk, and have/take a look*.

The English suffix indicating direction tends to be used as *-ward* in AmE and *-wards* in BrE. Further examples are: *southeastward/southeastwards,*
upward/upwards, afterward/afterwards, toward/towards, downward/downwards, leftward/leftwards, backward/backwards, and forward/forwards.

Among the lexical variations between AmE and BrE are certain words that mean differently in each variety, such as academician (an academic in AmE) and (a member of an academy in BrE). Some examples can be found in Table 12.

Table 12
Same Words with Different Meanings in AmE and BrE

<table>
<thead>
<tr>
<th>Word</th>
<th>AmE</th>
<th>BrE</th>
</tr>
</thead>
<tbody>
<tr>
<td>à la mode (adj.)</td>
<td>served with ice cream</td>
<td>fashionable</td>
</tr>
<tr>
<td>bathroom (n.)</td>
<td>a room with a toilet in it</td>
<td>a room with a bath and/or shower and often a toilet</td>
</tr>
<tr>
<td>casket (n.)</td>
<td>a coffin</td>
<td>a small box used to keep jewels</td>
</tr>
<tr>
<td>homely (adj.)</td>
<td>unattractive</td>
<td>plain or ordinary, but pleasant</td>
</tr>
<tr>
<td>fix (v.)</td>
<td>cook or prepare food or drink</td>
<td>arrange or agree a time, place, price, etc.</td>
</tr>
</tbody>
</table>

A final dialectal difference might be seen in the use of idioms and phrases such as in drunk as a skunk (AmE) and drunk as a lord (BrE). Further examples are: knock on wood (AmE) – touch wood (BrE), a skeleton in the closet (AmE) – a skeleton in the cupboard (BrE), a drop in the bucket (AmE) – a drop in the ocean (BrE), beating a dead horse (AmE) – flogging a dead horse (BrE).

4.3. Grammatical Variations

The varietal differences manifest themselves in grammar as well. The first one of these is the verb conjugations in the past tense. The past conjugations of certain verbs in AmE and BrE are different as in burned (AmE) – burnt (BrE), fit (AmE) – fitted (BrE), quit (AmE) – quitted (BrE), wet (AmE) – wetted (BrE), gotten (AmE) – got (BrE), learned (AmE) – learnt (BrE), dove (AmE) – dived (BrE), dreamed (AmE) – dreamt (BrE), leaned (AmE) – leant (BrE), lit (AmE) – lighted (BrE), spelled (AmE) – spelt (BrE), and spilled (AmE) – spilt (BrE).

Another grammatical difference between the two varieties is using collective nouns as the subjects of sentences, whereby AmE tends to utilize a singular and BrE employs a plural verb conjugation.

The government intends to cut taxes (AmE).
The government intend (or intends) to cut taxes (BrE).
Ariana’s crew was amazing at last night’s concert (AmE).
Ariana’s crew were amazing at last night’s concert (BrE).

The varietal differences can also be observed in using prepositions in such phrases as *on the street* (AmE) – *in the street* (BrE), *on weekends* (AmE) – *at the weekend* (BrE), *in the back of* (AmE) – *behind* (BrE), and *play on team* (AmE) – *play in a team* (BrE).

A fourth grammatical difference in AmE and BrE shows itself in using simple past tense and present perfect tenses. While Americans prefer to use simple past tense, Britons use present perfect tense to speak about a past action considered to be relevant to the present. For instance:

I’m not hungry. *I just had* breakfast (AmE).
I’m not hungry. *I’ve just had* breakfast (BrE).

Another difference between the two varieties can be observed in using certain modal verbs, wherein both varieties prefer a different modal verb to signify the same thing. For instance:

*It is hot in here. Can I open the window?* (AmE)
*It is hot in here. Shall I open the window?* (BrE)
*You do not have to come until later* (AmE).
*You need not come until later* (BrE).

A final grammatical difference to be mentioned might be the tendency of AmE to use present subjunctive form in formal imperative constructions with clauses followed by “verbs, adjectives, and nouns of requiring and urging” (e.g., suggest)”(https://encyclopedia.com).

I suggest he *come* with us (AmE).
I suggest he *should come* with us (BrE).
It is vital that they *be* stopped (AmE).

It is vital that they *are* stopped (BrE).

### 4.4. Orthographical/Punctuational Variations

Varietal differences also exist in writing, with particular reference to orthography and punctuation. Some of the most common spelling variations are available in Table 13.

#### Table 13

<table>
<thead>
<tr>
<th>Orthographical Variations in AmE and BrE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corresponding spelling</td>
</tr>
<tr>
<td>o-ou</td>
</tr>
<tr>
<td>er-re</td>
</tr>
<tr>
<td>ize-isè</td>
</tr>
<tr>
<td>yze-yse</td>
</tr>
<tr>
<td>l-ll</td>
</tr>
<tr>
<td>e-oe</td>
</tr>
<tr>
<td>e-ae</td>
</tr>
<tr>
<td>e-ea</td>
</tr>
<tr>
<td>se-ce</td>
</tr>
<tr>
<td>og-ogue</td>
</tr>
<tr>
<td>dg-dge</td>
</tr>
<tr>
<td>one-letter difference</td>
</tr>
</tbody>
</table>

Dialectal variations are also present in punctuation usage. Double quotation marks (""”) have always been preferred in AmE, while single quotation marks (‘’) are used in BrE. For instance;

The word *“hombre”* means *“man”* in Spanish (AmE).

The word ‘*hombre’* means *‘man’* in Spanish (BrE).

In another punctuational difference, AmE tends to use periods in abbreviations (e.g., Mr., Ms., Dr.) unlike BrE (e.g., Mr, Ms, Dr). Punctuational variations also occur in indicating the time. Americans prefer to use a colon (e.g., 10:20), while Britons tend to use a period (10.20) when showing the time. A final variation occurs in using hyphens to separate prefixes. BrE tends to employ a hyphen after prefixes (e.g., co-operations), while AmE does not (e.g., cooperation). However, this rule disappears when a prefix is used before an adjective that is capitalized (e.g., anti-American).
5. Discussion and Implications

This research aimed to delineate the varietal differences between the two most prominent and widespread varieties of English: AmE and BrE. Major variations in four areas of language (i.e., pronunciation, vocabulary, grammar, and orthography/punctuation) were described for pedagogical utility. It is, therefore, anticipated that the corpus-based descriptions serve as a guide for teachers in teaching the differences between AmE and BrE. 10 (nine segmental, one suprasegmental) variations were listed in the pronunciation area. These variations are (non) rhoticity or different pronunciations of “r” sound, sound shifts from /θ/ to /θː/, /æ/ to /ɛː/, /ɑː/ to /ɔː/, /æ/ to /æː/, /ʊ/ to /ʊː/, /ʌ/ to /ɜː/, /ʊ/ to /ʊː/, /ʌ/ to /ɜː/, yod /j/ dropping, vowel lengthening/reduction, and stress placement in French loanwords. In his study, Demirezen (2006a) investigated the case of flapping in North AmE and provided a sample lesson to remediate its pronunciation difficulty. In another study, Demirezen (2006b) examined the two problematic sounds (i.e., [æ] and [ε]) for Turkish learners of English and presented remedial teaching materials for these sounds. In her thesis examining the rhoticity (the pronunciation of “r”) in English, Kobák (2017) found that non-rhotic speech is less understandable and difficult to comprehend for Czech English learners. Maeda (2009) reported a lack of awareness of yod dropping (/j/) for Japanese students. In their study exploring the perception and production of BrE vowels and consonants by Arabic learners of English, Evans and Alshangiti (2018) reported that some phonemic contrasts still continued to be problematic. Byers and Yavas (2017) argued that a lack of vowel reduction was a predictor of foreign accent and therefore examined vowel reduction in early and late Spanish-English bilinguals and found that vowel articulations of early bilinguals were inconsistent. Last but not least, Procházková (2011) reviewed in her thesis the pronunciation of the French loanwords in English and found certain tendencies, albeit with some exceptions, some of which are intricately related to word stress. This might imply the accurate placement of stress in English words of French origin for clear and intelligible pronunciation. The studies conducted in the related dialectal variations in pronunciation between AmE and BrE suggest that learners from different linguistic backgrounds might have difficulty with pronunciation at segmental and suprasegmental levels. It is, therefore, reasonable to assert that the knowledge of and competence in these phonetic/phonological differences might assist learners with clear and intelligible pronunciation.

The present research introduced lexical variations in five areas: individual words,
the use of *have* and *take*, the suffix *–ward*, same words with different meanings, and idioms and phrases. Previous studies examining the dialectal variations at the lexical level had different focuses. Regarding variations in individual words, Scott (2000) revealed that lexical differences in AmE and BrE might be confusing for international business communication (IBC) and thus affect understanding. He, therefore, suggested the teaching of these differences by IBC teachers and trainers to create awareness and hamper any potential misunderstanding in IBC contexts where English is the predominant means of communication. Another study investigated Swedish students’ awareness about the lexical differences between the two varieties and concluded that they could moderately distinguish, notwithstanding habitually mixing the two (Lindell, 2014).

Parjas and Supraba (2018) explored learners’ comprehension of vocabulary in AmE and BrE and found moderate understanding. A recent study by Novari et al. (2021) comparatively studied the lexical variations between the two varieties. They reported 150 variations in word usage, spelling, and pronunciation in the hope of diminishing any misunderstandings in communication. With regard to the use of suffixes in the two varieties, Lindquist (2007) investigated the use of *-wise* suffix in spoken and written AmE and BrE corpora and found twice as much prevalence of its use in spoken language. However, no studies, to the best knowledge of the researcher, examining the *-ward* suffix or the use of same words with different meanings in the two varieties were encountered in the literature. This study might, therefore, be the first to present them. In terms of idioms, Rafatbakhsh and Ahmadi (2020) presented a corpus of the most frequent idioms in AmE for the benefit of teachers, learners, and other stakeholders. The present study, on the other hand, presented the different idioms used for the same concept/expression in AmE and BrE. It might, thus, be summarized that both convergence and divergence exist in the findings. 4000 words are estimated to exist in BrE that have disparate meaning or usage in AmE (Davies, 2005). Given the predomination of these two varieties across foreign language education contexts (Young et al., 2016; Rose & Galloway, 2019), learners are likely to encounter those words. For that reason, they should receive pedagogical exposure to the target lexical items. Furthermore, some of the words might be high-frequency. For instance, the word “*program*” (AmE) is listed as the 223rd (395 occurrences) and “*shall*” (BrE) as the 346th (268 occurrences) in the Brown corpus (Kučera & Francis, 1967). To serve a more pedagogical utility, the lexical items in both
varieties might be categorized into high- and low-frequency and presented to learners.

Six dialectal variations were presented in grammar: past tense verb conjugations, usage differences in collective nouns, prepositions, simple past tense (SPT) and present perfect tense (PPT), modal verbs, and subjunctive mood. With regard to verb conjugations in past tense, Gray et al. (2018) investigated verb regularization in textbooks and tweets and revealed that tweets geotagged in the United States have higher regularization than AmE books, but tweets geotagged in the United Kingdom have lower regularization than BrE books. About collective nouns, Levin (2001) explored them in AmE, BrE, and Australian English and suggested “dialectal, stylistic, diachronic, syntactic and semantic factors” (n.d.) were at play for variations.

Álvarez-Alonso (2019) examined in her thesis the use of collective nouns in AmE and BrE in particular and found that the British considered collective nouns (e.g., family) comprising a group of people, while they were a single unity or entity for the Americans. Zhang and Jiang (2008) outlined dialectal variations in the use of three prepositions (i.e., in, on, and at), which concurs with the present study. In another study, Abakumova (2006) explored the use of to and with in the verb talk and revealed that talk with was less used than talk to, and that the first one implied a two-way communication, while the latter referred to both one-way and two-way communication. In regard to the use of SPT and PPT, Yao and Collins’s (2013) corpus-based study indicated a strong preference for non-present-perfects (past, modal, to-infinitival and ing-participial) in BrE and a moderate decrease in past and modal perfect frequencies in AmE. Sheen’s (1984) research demonstrated that SPT tended to be used more in AmE in contexts where PPT was deemed appropriate, which accords with the present study.

Previous studies about modal verbs focused on various aspects. Murphy and Felice (2018), for instance, found that can, could, and would are used less in AmE and more in BrE in interrogatives. Another study indicated the overrepresentation of shall and should in BrE and will and would in AmE (Krovgv & Johansson, 1984), which was also included in this study. A study investigating the use of subjunctive in Pakistani English, AmE, and BrE demonstrated that AmE preferred the mandative subjunctive more than BrE, which was in line with the present study. Given that language textbooks are mostly published in AmE and BrE (Ayawad-
Bishara, 2020), learners and teachers are most likely to encounter the grammar of these varieties. It is, therefore, plausible to assert that language teachers need to be supplied with the grammatical knowledge including the differences between the two varieties, and learners should be exposed to these two varieties in terms of grammar. It might further be urged that more corpus-based studies might be conducted to reveal the most frequent dialectal variations in grammatical usage.

Last but not least, five issues (i.e., spelling variations, the use of double/single quotes, periods in abbreviations, colon/period in dates, and hyphen to separate prefixes) were discussed in relation the varietal differences in writing, more precisely in orthography and punctuation. Scott (2004) examined the spelling differences in business-related documents and suggested their teaching for preparing more culturally-sensitive business documents. Tottie (2002) conceded that the spelling variations between the two dialects were systematic but they should be learned individually. Van Gelderen (2014) explained that the relatively Standard English assisted with mutual intelligibility between the two varieties despite the spelling variations.

The variations in using quotation marks, periods, colon, and hyphen were also described in previous studies (Ramilevna & Aleksandrovna, 2016; Kumar, 20017;). The use of hyphens was claimed to be more frequent in BrE than in AmE (Quaile et al., 2019), which supports the current research. Based on the findings, it is plausible to assert that learners and teachers should possess the knowledge of orthography and punctuation because orthographic knowledge and competence indicate learners’ awareness, understanding, and ability to discern the differences between AmE and BrE in terms of orthography, and a lack of punctuation might lead to writing errors (e.g., run-on sentences) and ambiguity (Kirkman, 2006) and harsh evaluation of learners’ writing (Ferris, 2002). For these reasons, punctuation accuracy is needed for better writing. An awareness of the different use of punctuation marks in AmE and BrE might, in this sense, help learners avoid any ambiguity or stylistic errors.

On the whole, 26 variations in pronunciation, vocabulary, grammar, and orthography/punctuation were described as the most frequent dialectal differences in AmE and BrE. The variations outlined in this study both accorded and discorded with previous studies. However, it is safe to say that this research somewhat varies from the earlier ones in that it presents high-frequency dialectal variations in the two major English varieties in the hope of serving as a guideline for teachers to teach and learners to gain awareness. It must, however, be emphasized that more
corpus-based studies (especially large-scale ones) regarding the dialectal variations are needed to determine high-frequency phonological, lexical, grammatical, and orthographical/punctuational variations in AmE and BrE.

6. Conclusion

A total of 26 variations ($f=10$ in phonology, $f=5$ in vocabulary, $f=6$ in grammar, and $f=5$ in orthography/punctuation) was delineated between AmE and BrE—the two widely encountered English varieties in educational materials and contexts—for pedagogical purposes. The research further intended to provide language teachers with a pedagogical framework for the teaching of phonological, lexical, grammatical, and orthographical/punctuational variations in AmE and BrE. Given the prevalence of these two varieties in academic texts and educational materials (Rose & Galloway, 2019), the teaching of these dialectal variations becomes prominent. It is also essential to teach the varietal differences in AmE and BrE, with specific reference to pronunciation, lexis, grammar, and orthography/punctuation because these sub-skills are correlated with each other and significant for language proficiency (Nation, 2001; Richards, 2015).

It goes without saying that this study has several limitations. First, this is a small-scale corpus and descriptive-in-nature study aimed at specifying the important and prominent varietal alterations in AmE and BrE. No pedagogical interventions were, therefore, made. Future studies might explore the levels of awareness of language learners and teachers from different educational and cultural backgrounds or evaluate their competence levels concerning the differences between the two varieties. Second, not all variations were included in the research since only high-frequency and significant ones were included. Further research might include other variations considered to be important for language learning/teaching. Despite these limitations, the present research might still be viewed salient in that it might contribute to the teaching and awareness-raising of the dialectal variations in AmE and BrE.
References


Braun, S. (2005). From pedagogically relevant corpora to authentic language learning contents. ReCALL, 17, 47-64. https://doi.org/10.1017/S0958344005000510


A Small-Scale Corpus-Based … Ibrahim Halil Topal


contexts for emerging Sociopragmatic competence. *Education and Information Technologies*, 23(2), 705–724. https://doi.org/10.1007/s10639-017-9631-0


Murphy, M. L., & De Felice, R. (2019). Routine politeness in American and British English requests: use and non-use of please. *Journal of Politeness...


Rindal, U. (2014). Questioning English standards: Learner attitudes and L2 choices


Van der Sijs, N. (2009). *Cookies, coleslaw, and stoops: The influence of Dutch on*


**About the Author**

İbrahim Halil Topal is an English instructor at the College of Foreign Languages of Gazi University with more than ten years of teaching experience. He is currently doing his Ph.D. at the ELT Department of Hacettepe University. He has published in several international journals and presented at many international conferences. His main research interests include but are not limited to educational technology, teacher education, program evaluation, and second language pronunciation.