

# **A Corpus-based Study of Linking Adverbials in Written American English**

*Lilit Pipatanusorn*

*Raksangob Wijitsopon*

*Chulalongkorn University, Bangkok, Thailand*

*Corresponding author's email: edelweiss.ern@gmail.com*

Received 14 January 2019; revised 15 March 2019; accepted 18 March 2019; online 31 May 2019

## **Abstract**

In this study, we investigate uses of linking adverbials in present-day written American English, in response to an observation from a previous study that occurrences of linking adverbials may vary according to varieties of English. The American English 2006 corpus (Potts & Baker, 2012) was utilized for the purpose of the study. Distribution patterns of linking adverbials were examined in the light of their categories and text genres, namely general prose, academic prose, fiction, and media texts. The results showed that among the four text genres, linking adverbials occur most frequently in academic prose and least frequently in fiction and media texts. The distribution patterns of linking adverbials by category

revealed that the additive group occurs most frequently across the board while the sequential group the least. These overall findings correspond to previous studies of linking adverbials in other varieties of English, thereby suggesting that uses of linking adverbials are associated with text genres and categories, rather than with varieties. However, it is also found that such major distribution tendencies are not shared by a large number of linking adverbials in each group. This suggests that individual LAs have their idiosyncratic usage patterns, and these are associated with their frequencies in each genre. Analysis of some individual LAs revealed that there are other contextual factors, apart from text genres, including textual positions and co-occurrence patterns, which deserve attention in future research on linking adverbials.

**Keywords:** linking adverbials, American English,  
corpus-based

## 1. Introduction

Linking adverbials (henceforth LAs) are linguistic features that signal specific kinds of logical relationship between propositions in spoken and written communication, such as *however*, *therefore*, and *for example* (Biber, Johansson, Leech, Conrad, & Finegan, 1999). LAs are especially common in writing, where participants in the discourse do not share spatio-temporal contexts. Hence, writers would have to rely heavily on linguistic resources to generate propositional links and flows in texts while at the same time readers more or less need signals of relationship among propositions to comprehend and interpret textual meanings (Brown & Yule, 1984; Swales & Feak, 2012).

A number of previous studies of LAs concentrated on frequency and usage of LAs especially in academic writing. A great deal of research addressed the ways in which LAs were used in learner academic writing (e.g., Gilquin, Granger, & Paquot, 2007; Granger & Paquot, 2009; Leedham & Cai, 2013; Lei, 2012; Shaw, 2009). At the same time, a number of studies investigate LAs in various genres (e.g., Biber et al., 1999; Conrad, 1999; Liu, 2008). While previous research tends to focus on the relationship between LAs and genres, an

observation has been raised by Liu that the use of LAs may vary across varieties:

Furthermore, given that this [Liu's] study focused exclusively on British English and given that Biber et al.'s (1999) data have shown some differences between British and American English in the use of some LAs, it will be necessary in the future to examine other varieties of English to determine any specific differences among them in the use of these adverbials. (Liu, 2008, p. 509)

As mentioned above, most descriptive LA studies investigated British English data, with an exception of Biber et al. (1999), which made use of an American English dataset but focused on a few LAs in spoken data. Motivated by an incomplete picture of description of LAs in terms of language varieties, this study attempts to take an initial step to investigate LAs in American English, so that it can complement the picture of LA usage. The research questions the present study seeks to answer are:

1. What are distribution patterns of linking adverbials in written American English?
2. In what ways are distribution patterns of linking adverbials in written American English related to text genres?

We will first explain related concepts about linking adverbials, including their definitions, categories, and related studies; and then describe the methodology adopted. Then, the results will be reported and discussed, both quantitatively and qualitatively, followed by concluding remarks.

## **2. Literature review**

### **2.1 Defining English linking adverbials**

LAs are just one type of linguistic devices that connect information in texts. According to Liu (2008), there are two broad categories of terminologies that refer to linguistic features that serve this function: generic and specific. Generic terms (e.g., *adjuncts*, *connectives*, *connectors*) cover adverbials and conjunctions-both coordinating and subordinating; specific terms (e.g., *adverbials*, *adverbs*, *conjuncts*) include only adverbials.

As for the term *linking adverbial*, it was put forward by Biber et al. (1999) to refer to clausal elements whose main function is “to state the speaker/writer’s perception of the relationship between two units of discourse” (p. 875). In Biber et al.’s (1999) taxonomy, LAs can be categorized

into five grammatical forms: (a) single adverbs (e.g., *anyway*, *however*, *nevertheless*), (b) adverb phrases (e.g., *even so*, *first and foremost*, *more precisely*), (c) prepositional phrases (e.g., *by the way*, *for example*, *in addition*), (d) finite clauses (e.g., *that is*, *that is to say*, *what I mean is*), and (e) non-finite clauses (e.g., *to cap it all*, *to conclude*, *to put it another way*).

In the same way, Liu (2008) explains the properties of LAs in functional terms as linguistic elements consisting one or more words that “connect spans of discourse semantically at different lengths or levels (e.g. clause, sentence, and paragraph)” (p. 492) which “provide only semantic connections” (p. 492), unlike conjunctions that also perform a grammatical link. Based on Liu’s definition, conjunctions are therefore not a member of LAs. Despite this distinction, it was found that some conjunctions, namely, *so*, *but* and *yet*, were included in Liu’s (2008) LA list. Also, the conjunctions *so* and *yet* are also included in Biber et al.’s (1999) analysis even though conjunctions are not listed as one of the five grammatical categories of LAs in their account.

To develop a systematic and operationalizable definition of LAs for analysis in the present study, we use two criteria

for identification of LAs: form and function. For the latter, we still apply the definitions generally given in the literature of Las-linguistic devices whose function is to generate a specific reading of link between pieces of information in discourse. We also adopted Liu's (2008) LA functional categories and list of LAs due to its clarity and the fact that it contains more items than other lists (e.g., Biber et al., 1999; Celce-Murcia & Larsen-Freeman, 1999).

In terms of form, only adverbials are considered as LAs in this study; conjunctions are excluded. By adverbials, we refer to words or phrases that behave like adverbs, being independent in terms of movement (Celce-Murcia & Larsen-Freeman, 1999); that is, the items may be moved to other positions within the domain governing it, either an independent clause or a phrase such as a noun phrase. In other words, an LA does not move to higher or prior clauses or move out of a particular phrase. Based on this criterion, it can be said that we opt for a *specific* type of definitions of linking words, as explained earlier.

To illustrate the working definition of LAs used in this study, let us compare the behavior between LAs and conjunctions. In (1) the sample LA *nevertheless* behaves like

an adverb in that it semantically links the whole sentence (in this case an independent clause) with prior materials in discourse and can occur in different positions, whether initial, medial, or final. Note that in this study, when an LA is said to be at an initial position, it is the first constituent of a sentence and is capitalized in written texts. LAs in the medial position are those that occur at any positions between the first and the last word of a sentence, whose boundary is marked by a capital letter and a full stop. An LA is said to be in the final position when it is the last constituent of a sentence, immediately followed by a full stop.

- (1) a. **Nevertheless**, they carved out a 5.7 per cent share of the overall vote.
- b. They, **nevertheless**, carved out a 5.7 per cent share of the overall vote.
- c. They carved out a 5.7 per cent share of the overall vote, **nevertheless**.

(Biber et al., 1999, p. 80)

In contrast, coordinating conjunctions have a fixed position in an independent clause, which is at the left edge and linked with the prior independent clause in discourse, as illustrated in (2a). A movement of a coordinating conjunction



to other positions, as illustrated in (2b) and (2c) below, results in ungrammaticality.

- (2) a. **But** they carved out a 5.7 per cent share of the overall vote.  
b. \*They, **but**, carved out a 5.7 per cent share of the overall vote.  
c. \*They carved out a 5.7 per cent share of the overall vote, **but**.

(Biber et al., 1999, p. 80)

On the other hand, the position of a subordinating conjunction is also grammatically constrained due to the structure of a subordinating clause (Celce-Murcia & Larsen-Freeman, 1999), governed by an independent clause. As illustrated in (3), each contains a subordinating conjunction *although* which is a part of a dependent clause [*although* he was sick]. We can see that a movement of the word *although* within its local domain alone—to other positions in the clause [he was sick]—is not allowed. The only possible condition for *although* to move is when the whole clause containing *although* is moved within possible positions of the independent clause by which it is governed.

- (3) a. [**Although** he was sick], Tom went to work.  
b. Tom [**although** he was sick] went to work.  
c. Tom went to work [**although** he was sick].

(adapted from Liu, 2008, p. 492)

That is to say, the behaviour of a subordinating conjunction is different from LAs, so they are excluded from a group of LAs.

An issue of the current definition of LAs seems to arise with some LAs such as *i.e.* Consider the following examples:

- (4) a. [...], what foretold another harsh event was actually [an attack by Mother Nature], *i.e.*, Hurricane Katrina.

(adapted from AmE06, F03, emphasis added)

- b. Additives can be either [synthetic] (*i.e.*, chemically produced, artificial) or natural in origin.

(adapted from AmE06, J46, emphasis added)

It can be seen that *i.e.* generates a link between the noun phrase [an attack by Mother Nature] headed by a noun *attack* with the noun phrase *Hurricane Katrina* in (4a) and an adjective phrase [synthetic] having *synthetic* as a single material as a head with an adjective phrase *chemically produced, artificial* in (4b). The domain that immediately governs *i.e.* is immediate materials precede it—the noun phrase [an attack by Mother Nature] in (4a) and an adjective phrase [synthetic] in (4b).

Although *i.e.* appear to possess a characteristic like conjunctions having a specific position constrained by the structure. However, *i.e.* itself is not a conjunction for conjunction binds linguistic materials from words to clauses, but *i.e.* functions as linking elements at phrasal level, as illustrated in (4). Additionally, if *i.e.* in (4) is a conjunction, it must be seamlessly substituted by a conjunction, but a substitution with conjunctions like *and* suggests another distinct matter, yielding a different kind of link. Compare *synthetic (i.e., chemically produced)* and *synthetic and chemically produced* in (4a). While the former being the same thing from different perspectives, the latter being same perspective but distinct matters.

On the other hand, *i.e.* can be substituted with another adverbial such as *that is to say* or *in other words* without ungrammaticality. This means that theoretically *i.e.* can be moved within its domain although it may not be frequently used, that is, to another possible position—final position of a phrase containing it: a noun phrase *Hurricane Katrina* in (4a) and an adjective phrase *chemically produced, artificial* in (4b). Hence, *i.e.* is argued to be an LA for the fact that it is not a conjunction and behaves more like an LA—can

be substituted by an LA and theoretically possible to move within its local domain. See also Biber et al. (1999).

As mentioned earlier, Liu's (2008) LA functional categories were adopted, so this study categorized LAs into five major categories: *additive*, *adversative*, *causal/resultative*, *sequential*, and *multi-meaning*. Additive LAs are those that signal additional information related what is mentioned earlier (e.g., *in addition*, *that is to say*, *for example*). Adversative LAs show that propositions being linked are opposite, different from expectation or suggest different directions (e.g., *however*, *in fact*, *instead*). Causal/resultative LAs connect propositions that are causes and results of one another (e.g., *thus*, *therefore*, *hence*). Sequential LAs can signal a list or an order of items (e.g., *first*, *second*, *finally*) or temporal relationship (e.g., *meanwhile*). The last category is multi-meaning LAs which are those that can be classified into more than one category mentioned earlier; these include *at the same time*, *of course*, and *then*.

In fact, Liu (2008) proposed that *as a matter of fact* is also a multi-meaning LA, indicating either adversative or additive relationship. However, the data from AmE06 in

this study suggests that this phrase signals a strong adversative reading, as illustrated in (5) below. Hence, in this study *as a matter of fact* was not considered as multi-meaning and categorized as an adversative LA instead.

- (5) According to the president, “Freedom is God’s gift to everybody in the world.”

He told Bob Woodward, “**As a matter of fact**, I was the person who wrote the line, or said it.”

(AmE06, D07, emphasis added)

## 2.2 LA across genres and varieties of English

In this section, major findings in previous corpus-based LA studies are summarized and discussed. Based on the Longman Spoken and Written English Corpus (the LSWE Corpus), which is a corpus of spoken and written British and American English, Biber et al. (1999) reported that across genres LAs are most common in academic prose followed by conversation, and noticeably less frequent in fiction and news. Their findings also revealed that distribution patterns of LAs vary by LAs categories.

Some minor differences between LAs in British and American English in terms of frequency were also reported; for example, *though* and *anyway* were found in British and American conversation in similar proportions

while *then* occurred more frequently in British conversation. Nevertheless, the treatment of LA uses in American English in Biber et al.'s study was relative limited as they only addressed some LAs in conversation, including *so*, *then*, *though*, and *anyway*.

Conrad (1999) reported her corpus findings on English LAs particularly in conversation and academic prose based on British English sub-corpus of the LSWE corpus (Biber et al., 1999). Although LAs occur more frequently in academic prose and conversation, Conrad found that usage patterns of LAs in these two genres are not the same. Most LAs in spoken conversation are single adverbs. In academic prose, although single LAs are dominant, prepositional phrases are more common than in conversation. In terms of positions, in conversation, LAs occur in the initial and final positions more frequently but rarely in the medial position, while in academic prose, LAs occur more frequently in initial and medial positions in British English.

Liu (2008) on the other hand investigated 110 LAs in British English across five genres: conversation, academic prose, fiction, news, and a miscellaneous group of written texts, by investigating the British National Corpus (BNC).

LAs were found to be most common in academic prose, and they are more common in conversation than in the remaining three written genres. LAs in news are relatively few when compared with other genres. These findings are in accordance with Biber et al. (1999) and Conrad (1999) cited above. Another important point made in Liu's study is a suggestion that some LAs indicate multiple textual relationships. These are *of course*, *at the same time*, and *as a matter of fact* and hence should be put in a separate category. This category of LAs will be adopted in the analysis of LAs in American English in the present study, as mentioned earlier.

### **3. Methodology**

#### **3.1 Corpus data**

This study was conducted upon the American English 2006 corpus (henceforth AmE06) (Potts & Baker, 2012). The AmE06 is a corpus of present-day written American English consisting of 1,175,965 tokens from 500 published texts. It was chosen for the purpose of the present study because it contains contemporary written American English texts, which have been categorized into four different text

genres: (a) general prose (478,402 words), such as articles on hobbies, letters, memoirs, reports, and biography; (b) academic writing (185,506 words); (c) fiction (308,486 words); and (d) media texts (203,571 words), including reportage, reviews and editorials. The categorization of text genres is already made by corpus's developers (Potts & Baker, 2012).

There are some other corpora representing American English. A notable one is the Corpus of Contemporary American English (COCA), which contains approximately 560 million tokens. However, it was not chosen for the present study mainly because COCA is a monitor corpus. Its size keeps growing and the frequency of a search term is tied to differing periods of time. This would bring time periods into the study as another central factor when interpreting data from different text genres. The AmE06, though much smaller, can give adequate information on distribution patterns in relations to text genres, fitting the scope of this research.

### **3.2 LA identification and data extraction**

Based on Liu's (2008) list of LAs, a total of 117 LAs were identified by the criteria explained in Section 2.1. All LAs were categorized into 5 main categories: (a) *additive*



(38 items), (b) *adversative* (26 items), (c) *causal/resultative* (16 items), (d) *sequential* (34 items), and (e) *multi-meaning* (3 items), based on Liu's suggestion.

Both raw and relative frequency information of each LA was derived from an online web-based corpus processor—the CQPWeb (Hardie, 2012). Every occurrence of LAs in the list was checked manually to prevent inclusion of cases wherein a search item was not used as an LA in a certain context, such as *still*, which is sometimes used as an adverb modifying a verb, rather than as an LA, or *besides*, which, in some cases, is used as a preposition, not as an LA.

#### **4. Analysis and findings**

The analysis of LAs in the present study focuses on an interface between text genres and LA categories. An overall distribution pattern of LAs across four text genres will be first reported. Then, examinations of LA distribution by category will be presented and discussed upon their specific uses in relation to text genres.

##### **4.1 Overall distribution**

Table 1 below shows distribution patterns of LAs across text genres. According to Table 1, in the AmE06, LAs

were found most frequently in academic prose. As hypothesized, this particularly high frequency can be linked to the nature of academic texts in that they involve systematic explanations and argumentations by the authors such as explaining concepts or presenting research findings to a number of readers, who are likely to be members of academic disciplines both professionals or students (Biber & Conrad, 2009). Hence, explicit uses of LAs can indicate a kind of relationship among propositions and facilitate readers’ comprehension of the academic contents.

**Table 1** Distribution of LAs in the AmE06 (instances per one million word)

LA Category	Overall	Text genre			
		General	Academic	Fiction	Media
		Prose	prose		
Additive	2,133.61	2,242.88	3,439.23	1,024.34	2,004.19
Adversative	1,181.99	1,164.28	1,724.99	959.51	1,065.96
Causal/resultative	420.08	476.08	1,067.34	100.48	181.75
Sequential	367.36	386.69	506.71	291.71	309.51
Multi-meaning	1,462.62	1,427.91	1,045.79	2,466.89	825.26
<b>Total</b>	<b>5,565.66</b>	<b>5,697.84</b>	<b>7,784.06</b>	<b>4,842.93</b>	<b>4,386.67</b>

This predominance of LAs in academic prose in the AmE06 corresponds to the findings in previous corpus-based studies of LAs. Such correspondence gives a strong indicator

of functional characteristics of LAs and their relations to texts in which they occur. At the same time, this seems to hint at a tendency in which occurrences of LAs largely depend on text genres, rather than on language varieties in the way Liu (2008) hypothesized. However, it should be noted that the predominance of additive LAs may be linked to a relatively broad definition of this category, that is, those that introduce additional information related to prior proposition, which can occur in different ways, such as exemplification, elaboration, or comparison, while the other groups are more specific in functions. As a result, this LA category contains more members than the other groups and putting together the frequency of each additive LA can result in the highest frequency of this LA group.

Not only the predominance but also a rare occurrence of LAs in some text genre, particularly in media texts, corresponds to previous studies (e.g., Biber et al., 1999; Liu, 2008) as it was also found here that the overall LA frequency is the lowest in this text genre. This may be attributable to a journalistic style of writing, wherein different pieces of information are put in different paragraphs, rather

than joined via LAs in a single paragraph. In some cases, only one sentence can make up a paragraph in news reports.

That said, when we look more closely at the frequencies of each LA category, it can be seen that the relationship between LA distribution and text genre is not as straightforward. For example, the predominance of LAs in academic prose is not evenly distributed across the four LA categories, with multi-meaning LAs being found less common in academic prose than in fictional texts. At the same time, media texts, which were found to contain the least LAs, display more additive LAs than fiction, and in fact fiction falls short in all kinds of LAs but the multi-meaning group, whose frequency excels their counterparts in the other three text genres. All these give a strong indication that frequencies of LAs in each genre are associated with LA category. This issue will be investigated in detail the next section.

#### **4.2 Distribution and usage patterns by LA category**

In this section, apart from LA categories, some individual LAs in each category will also be focused on.

4.2.1 Additive LAs

Table 2 below illustrates distribution patterns of individual additive LAs found in the AmE06. Overall, additive LAs occurred most frequently in academic prose. Among 31 additive LAs found in the corpus, *also* is the most common in all text genres. Compared with other additive LAs on the list, the frequency of *also* shows a considerable disparity from that of the second most frequent additive LA *too*. This suggests that *also* tends to be used far more frequently and in a wider range of texts than other additive LAs. Such distribution patterns of *also* are in line with those findings in the literature (e.g., Liu, 2008), indicating a strong dominance of *also* in both American and British English academic texts.

**Table 2** Distribution of additive LAs in the AmE06  
(instances per one million word)

LA	Overall	Text genre			
		General	Academic		
		Prose	prose	Fiction	Media
1. also	1,008.53	1,116.22	1,552.51	317.68	1,306.67
2. too	236.40	200.67	107.81	379.27	221.05
3. for example	197.28	252.93	463.60	25.93	83.51
4. as well	110.55	100.33	107.81	132.91	103.16
5. besides	81.71	12.54	0	45.38	9.82

LA	Overall	Text genre			
		General	Academic	Fiction	Media
		Prose	prose		
6. that is	56.97	56.44	194.06	9.72	4.91
7. moreover	56.12	94.06	75.47	0	34.39
8. for instance	47.62	58.53	80.86	6.48	54.04
9. in addition to NP	47.62	66.89	75.47	12.97	29.47
10. in addition	41.67	37.63	161.72	0	4.91
11. again	32.31	43.90	5.39	35.66	24.56
12. i.e.	31.46	12.54	167.11	0	0
13. in other words	28.91	39.72	32.34	12.97	24.56
14. likewise	28.91	27.17	75.47	16.21	9.82
15. furthermore	25.51	8.36	129.38	3.24	4.91
16. similarly	20.41	27.17	43.13	0	14.74
17. namely	15.31	12.54	53.91	0	9.82
18. above all	14.46	16.72	21.56	9.72	9.82
19. further	13.61	10.45	37.73	0	19.65
20. additionally	7.65	12.54	10.78	0	4.91
21. alternatively	5.10	8.36	10.78	0	0
22. that is to say	5.10	4.18	16.17	3.24	0
23. for one thing	4.25	2.09	10.78	3.24	4.91
24. what's more	4.25	2.09	0	3.24	14.74
25. as they say	2.55	4.18	0	0	4.91
26. which is to say	2.55	6.27	0	0	0
27. by the same token	1.70	4.18	0	0	0
28. what I mean is	1.70	2.09	0	3.24	0
29. what is more	1.70	0	5.39	0	4.91
30. correspondingly	0.85	2.09	0	0	0
31. to put it mildly	0.85	0	0	3.24	0
<b>Total</b>	<b>2,133.61</b>	<b>2,242.88</b>	<b>3,439.23</b>	<b>1,024.34</b>	<b>2,004.19</b>

*Note.* The following LAs were not found in the AmE06 (frequency = 0): *as I say*, *as you say*, *to cap it all*, *to crown it all*, *to put it another way*, *to put it bluntly*, and *what I'm saying is*.

In addition to *also*, there are some LAs that are more common in academic prose than in other text genres, including *for example*, *that is*, *i.e.*, *in addition*, *furthermore*, *likewise* and *namely*. Based on their semantic properties, they are used for illustrating, elaborating or introducing new points in discussion, which can possibly in turn explain why they are relative frequent in academic texts.

A more interesting point is perhaps the fact that not all additive LAs are most frequent in academic prose. This is the case with as many as 14 out of 31 additive LAs (45.16%): *too*, *as well*, *besides*, *moreover*, *again*, *in other words*, *additionally*, *what's more*, *as they say*, *which is to say*, *by the same token*, *what I mean is*, *correspondingly*, and *to put it mildly*. While it is not very surprising that some of these were not found in the academic prose component of AmE06 because their forms involve personal pronouns or contracted forms, usually characteristic of spoken texts, such as *what's more*, *as they say*, *what I mean is*. The rarity of some additive items, including *moreover* and *besides*, is relatively of a surprise, given that they are often taught to L2 learners of English in academic writing classes

(Ha, 2016; Ishikawa, 2010; Ro & Na, 2012). In fact, in AmE06, *besides* was not found at all in academic prose but most frequently in fiction, a text genre in stark contrast with the academic prose. This can raise an issue of correspondence between ELT lessons and authentic use of language, especially in terms of the degree of emphasis on some LAs over the others in ESL or EFL writing course materials.

At the same time, the findings have shown that LAs generally treated as interchangeable like *furthermore* and *moreover* have different distribution patterns, the former being most common in academic writing while the latter in general prose and occurring less than 10 times per million words in the other three genres. This case illustrates the complexity of synonymous words often taught to L2 learners of English as different alternatives in writing, which does not seem to be adequately dealt with in the pedagogical context.

#### **4.2.2 Adversative LAs**

Table 3 below illustrates the distribution patterns of adversative LAs in the corpus. Overall, adversative is another LA category that dominates academic prose. As can be seen, *however* is the most frequent adversative LA



overall and also in academic prose. Other LAs that are noticeably more frequent in academic prose than other text genres are *rather*, *nevertheless*, *nonetheless*, *in contrast*, *on the other hand*, and *conversely*.

**Table 3** Distribution of adversative LAs in the AmE06 (instances per one million word)

LA	Overall	Text genre			
		General	Academic	Fiction	Media
		Prose	prose		
1. however	376.71	405.52	900.24	110.22	235.79
2. in fact	140.31	156.77	134.77	119.94	137.54
3. instead	110.55	94.06	97.03	119.94	147.37
4. still	78.23	52.26	37.73	94.01	152.28
5. though	74.83	58.53	21.56	110.22	108.07
6. after all	73.13	77.34	16.17	106.97	63.86
7. rather	56.97	85.70	102.42	3.24	29.47
8. anyway	56.12	31.35	10.78	126.42	49.12
9. nevertheless	43.37	43.90	97.03	16.21	34.39
10. on the other hand	37.42	41.81	59.30	19.45	34.39
11. nonetheless	32.31	20.90	86.25	16.21	34.39
12. actually	20.41	8.36	5.39	48.62	19.65
13. by contrast	14.46	27.17	21.56	0	0
14. in contrast	13.61	8.36	59.30	0	4.91
15. then again	11.91	2.09	5.39	32.42	9.82
16. conversely	10.20	8.36	32.34	6.48	0
17. in any case	10.20	12.54	10.78	9.72	4.91
18. at any rate	5.10	4.18	16.17	3.24	0
19. on the contrary	4.25	10.45	0	0	0
20. by comparison	3.40	4.18	10.78	0	0
21. in reality	3.40	6.27	0	3.24	0
22. anyhow	2.55	0	0	9.72	0
23. as a matter of fact	1.70	4.18	0	0	0
24. admittedly	0.85	0	0	3.24	0
<b>Total</b>	<b>1,181.99</b>	<b>1,164.28</b>	<b>1,724.99</b>	<b>959.51</b>	<b>1,065.96</b>

*Note.* The following LAs were not found in the AmE06 (frequency = 0): *in comparison* and *all the same*.

It is interesting that among all 24 adversative LAs types found in AmE06, as many as 14 adversative LAs (58.33%) occur less frequently in academic prose. In a way, this suggests that the overall predominance of adversative LAs in academic prose is associated with a particularly high frequency of a certain individual adversatives in academic texts, such as *however*, *in fact*, *rather*, *nevertheless*, and *nonetheless*. However, two adversative LAs that appear to be remarkably characteristic of non-academic texts and deserve particular attention here are *anyway* and *still*. This is because *anyway*, whose frequency is the highest in fiction, is an LA that shows the strongest contrast between its occurrence in academic texts, with a relative frequency of more than 125 tokens per million words, and in other text genres, with less than 50 tokens per million words (see Table 3). This is likely to be due to the conversational and narrative characteristics of fictional texts. Eighteen instances of *anyway* from its 39 occurrence (46.16%) were found to occur within quotation marks. While this predominance tendency of *anyway* in fiction was also noted in British English (Liu 2008), the finding based on an American English corpus here suggests that this LA seems to be

characteristic of the fictional genre, rather than language varieties.

In the case of *still*, it is the second most frequent adversative LA in media texts, following *however*—the most common adversative in academic prose, but occurs only approximately 38 times per million words in academic texts. Based on an examination of its concordance lines, it is found that 70.97% of *still* in media texts co-occur with words that have negative meanings associated with crime, violence, death, or illness. Some concordance lines are illustrated in Figure 1 below.

Figure 1. Concordance lines of *still* in media section of the AmE06

1	A08	<b>poor</b> and the <b>sick</b> wherever they are. They are also,	<b>Still,</b>	American exceptionalists, meaning they believe
2	A30	I have zero empathy for sex <b>offenders</b> .	<b>Still,</b>	there are reasons to believe <b>banishment</b> laws
3	A33	the rise among those between the ages of 16 and 24.	<b>Still,</b>	we find it <b>alarming</b> that a country that is so
4	A37	But the mission remains <b>excruciatingly difficult</b> .	<b>Still,</b>	those who make it their life's work know that
5	A37	nearly enough material to make a nuclear <b>weapon</b> .	<b>Still,</b>	though rolling up the AQ Khan network was
6	B01	economy would <b>buckle</b> if they were all <b>deported</b> .	<b>Still,</b>	their request for legal relief is often met with
7	B04	can and will do it for itself cannot be high.	<b>Still,</b>	the Ethiopian <b>offensive</b> against Islamic Courts
8	B08	with the <b>inflammatory</b> words from 600 years ago.	<b>Still,</b>	the <b>fallout</b> doesn't mean that the speech was a
9	B20	of about \$44 a share--and I don't think it should.	<b>Still,</b>	<b>pretending</b> to nationalize Wal-Mart
10	B23	<b>murders</b> and other <b>carnage</b> in and around Baghdad.	<b>Still,</b>	Zarqawi's <b>death</b> represents the most spectacular

On the basis of the AmE06 corpus, it can be said that the use of *still* as an LA not only serves the cohesive purpose but can also be evaluative, as a result of its collocational

patterns. In other words, the LA *still* may have a negative discourse prosody—a word’s association with evaluative meanings through its repeated co-occurrences with a group of words (Stubbs, 2007). This observation deserves to be investigated in further studies of LAs with a larger corpus. It should be noted, however, that this potential property of *still* is more or less linked with its distribution profile—its high density in the press.

#### **4.2.3 Causal/resultative LAs**

Causal/resultative LAs are also most frequently found in academic prose while noticeably uncommon in the other three text genres, especially in fiction, as illustrated in Table 4 below. According to Table 4, 10 out of 13 causal/resultative LAs (76.92%) predominate academic prose. This suggests that causal/resultative LAs are another strong stylistic feature of academic writing. The density of these LAs is largely due to communicative purposes of academic texts, which aim to report, explain, interpret findings and to make arguments (Biber & Conrad, 2009; Mackey & Gass, 2016).

It should be noted, however, that the strong association of causal/resultative LAs with academic prose is largely due to the frequencies of *thus* and *therefore*. They are the

only two causal/resultative LAs whose frequencies are over 250 tokens per million words (see Table 4), while among the rest, each LA occurs less than 100 tokens per million words and few are just below 10. The predominance of *thus* and *therefore* has also been noted from the historical perspective, as Biber and Gray (2016) reported that these two LAs have been used greatly in academic texts since 1800s although their frequencies have been decreasing.

**Table 4** Distribution of causal/resultative LAs in the AmE06 (instances per one million word)

LA	Overall	Text genre			
		General	Academic	Fiction	Media
		Prose	prose		
1. thus	175.18	183.95	495.94	58.35	39.30
2. therefore	102.89	127.00	253.36	16.21	39.30
3. hence	37.42	43.90	97.03	9.72	9.82
4. as a result	29.76	31.35	64.69	3.24	34.39
5. as a result of NP	28.06	41.81	32.34	3.24	29.47
6. otherwise	14.46	20.90	10.78	3.24	19.65
7. consequently	12.76	10.45	53.91	0	0
8. naturally	5.10	2.09	10.78	6.48	4.91
9. accordingly	4.25	8.36	5.39	0	0
10. as a consequence	3.40	4.18	10.78	0	0
11. as a consequence of NP	3.40	0	21.56	0	0
12. in that case	2.55	2.09	5.39	0	4.91
13. in such cases	0.85	0	5.39	0	0
Total	420.08	476.08	1,067.34	100.48	181.75

Note. The following LAs were not found in the AmE06 (frequency = 0): in consequence, in such a case, and all things considered.

Of 13 causal/resultative LAs found in the corpus (see Table 4), it is particularly interesting that while *as a result* is common in academic prose, its variant *as a result of NP (noun phrase)* is less frequent by half. A possible account for the difference between the frequencies of *as a result* and *as a result of NP* deals with the progression of information in discourse. Based on the concordance lines of *as a result of NP* in general prose, only 10% of it occurs in the initial position. This means that *as a result of NP* is more common in non-initial positions—medial and final. This leads to the “effect-before-cause” rhetorical pattern, as illustrated in (6). In academic texts, information tends to follow a clear and straightforward logical sequence of actions (Mackey & Gass, 2016; Swales & Feak, 2012), that is, in the “cause-before-effect” pattern. It is therefore possible that *as a result* is preferred in academic prose to maintain this more logical propositional relation, as illustrated in (7).

- (6) [.....*effect*.....  
We have revised our operating performance expectations for the full year

.....] LA [.....*cause*.....]  
2007 **as a result** of significant increases in dairy input costs.

(AmE06, H27, emphasis added)

- (7) [.....*cause*.....  
In many cases, containers are being stacked rather than stored on wheeled  
.....] LA [.....*effect*.....]  
chassis. **As a result**, more labor is needed to move and stack the cargo.

(AmE06, J72, emphasis added)

This qualitative observation is supported by the fact that eight of 12 concordance lines of *as a result* in the AmE06 academic prose component occur in the initial position (66.67%) in academic texts, pointing to the cause-to-effect informational progression pattern and this LA’s usual status as a sentence theme marker to highlight results explicitly. This is illustrated in Figure 2 below.

**Figure 2.** Concordance lines of as a result in the academic prose section of the AmE06

1	J25	authority (Kessner 1977; Zelizer 1985).	As a result	, immigrant children's adaptation to life in America
2	J10	Most aspects of the study will improve	as a result	, including reliability, resource utilization, quality
3	J33	of issues and problems into academics;	as a result	, instructors and their institutions have had to
4	J72	rather than stored on wheeled chassis.	As a result	, more labor is needed to move and stack the cargo.
5	J35	, so [demonstratives] turn into articles."	As a result	, Seiler notes, "new and still stronger
6	J42	by" the Federal Circuit's remand.	As a result	, the CIT affirmed the Commission's application of
7	J78	a similar manner at the single-cell level .	As a result	, the creation of a promoter library based on a
8	J25	through age 15 (Jacobs and Greene 1994).	As a result	, the shares of young adults who were enrolled in
9	J30	of emotion that can not be expressed .	As a result	, the unexpressed emotional experiences find
10	J42	incentive to provide the data needed."	As a result	, they reasoned that this presumption may well
11	J28	will be liked more than others, and, partly	as a result	, will be seen has having better leadership qualities
12	J70	slightly depressed from last years average	as a result	. This is thought to be due to a much wetter summer

#### 4.2.4 Sequential LAs

Table 5 below illustrates the distribution patterns of sequential LAs in the corpus. Overall, the total frequency of sequential LAs is lower than other LA categories, and yet it still dominates academic prose like the other groups.

**Table 5** Distribution of sequential LAs in the AmE06 (instances per one million word)

LA	Overall	Text genre			
		General Prose	Academic prose	Fiction	Media
1. finally	65.48	81.52	86.25	48.62	34.39
2. eventually	64.63	64.80	91.64	42.14	73.68
3. first	58.68	58.53	107.81	45.38	34.39
4. meanwhile	43.37	37.63	10.78	42.14	88.42
5. second	27.21	29.26	70.08	6.48	14.74
6. in the first place	17.86	12.54	16.17	25.93	19.65
7. by the way	17.01	8.36	0	48.62	4.91



LA	Overall	Text genre			
		General Prose	Academic prose	Fiction	Media
8. in short	15.31	22.99	16.17	9.72	4.91
9. third	14.46	12.54	43.13	0	14.74
10. in sum	7.65	16.72	5.39	0	0
11. in the	6.80	8.36	5.39	6.48	4.91
12. to begin with	5.10	8.36	5.39	3.24	0
13. incidentally	4.25	4.18	0	6.48	4.95
14. fourth	3.40	4.18	5.39	0	4.91
15. afterwards	2.55	2.09	10.78	0	0
16. lastly	2.55	6.27	0	0	0
17. next	2.55	2.09	5.39	3.24	0
18. all in all	2.55	0	10.78	0	4.91
19. secondly	1.70	0	5.39	3.24	0
20. in summary	1.70	4.18	0	0	0
21. first of all	0.85	0	5.39	0	0
22. in conclusion	0.85	2.09	0	0	0
23. to summarize	0.85	0	5.39	0	0
<b>Total</b>	<b>367.36</b>	<b>386.69</b>	<b>506.71</b>	<b>291.71</b>	<b>309.51</b>

*Note.* The following LAs were not found in the AmE06 (frequency = 0): *firstly*, *first and foremost*, *thirdly*, *fourthly*, *last*, *last but not least*, *last of all*, *in a word*, *to conclude*, *to sum up*, and *by the by*.

Despite the overall predominance of sequential LAs in academic prose, a closer look at distribution patterns of individual LAs reveals that 11 of 23 sequential LAs that were found in AmE06 (47.83%) occur more frequently in other text genres. These include *meanwhile*, *in the first place*, *by the way*, *in short*, *in sum*, *in the meantime*, *to begin with*, *incidentally*, *lastly*, *in summary*, and *in conclusion*. Moreover, it is also found that as many as 11 of 34 searched sequential LAs do not occur at all in AmE06 (see Table 5).

Surprisingly, some of these sequential LAs have received extensive attention in English language teaching (see, e.g., Leedham & Cai, 2013). Again, while it is possible that the corpus used here is relatively small and hence some words cannot be found, a high frequency of such sequential LA items mentioned earlier, can raise a question about the weight given to some LAs in the pedagogical context of ESL or EFL academic writing, which may be linked to L2 learner overuse of some of these LA items in their written output (see, e.g., Ishikawa, 2010; Leedham & Cai, 2013; Lei, 2012).

#### **4.2.5 Multi-meaning LAs**

The last LA category is the multi-meaning LAs—those that can signal multiple relationships among connected propositions, which makes it hardly possible to put them in a single category. These LAs include *then*, *of course*, and *at the same time*.

Table 6 below illustrates the distribution pattern of multi-meaning LAs in the corpus. Overall, this group of LAs occurs most frequently in fictional texts and least frequently in media texts. Among these, *then* was found to be the most frequent multi-meaning LA. An analysis of *then* shows that it can indicate at least two types of

relationships: sequential and causal/resultative (by means of inference). As illustrated in (8) below, the use of *then* is sequential, showing that the latter proposition—fronted by *then*—is subsequent to the preceding sentence.

**Table 6** Distribution of multi-meaning LAs in the AmE06 (instances per one million word)

LA	Overall	Text genre			
		General	Academic	Fiction	Media
		Prose	prose		
1. then	1,173.50	976.17	797.82	2,090.86	589.47
2. of course	218.54	196.49	161.72	327.41	157.19
3. at the same time	70.58	75.25	86.25	48.62	78.60
<b>Total</b>	<b>1,462.62</b>	<b>1,247.91</b>	<b>1,045.79</b>	<b>2,466.89</b>	<b>825.26</b>

- (8) He waved gallantly as he turned his horse. **Then** he turned his horse back around toward the girls once more.  
(AmE06, N15, emphasis added)

In some cases, *then* indicates both sequential and causal/resultative relationships. Consider the following extract from the fiction component in (9).

- (9) “I’ve got to stay. I’m working.”  
“But I really must speak with you, Miss Mintz.”  
“I had my break already,” Frieda says.  
“**Then** I guess we’ll just have to talk here.”  
(AmE06, K14, emphasis added)

In (9), the use of *then* indicate sequential relationship between what has been introduced in the discourse and the sentence containing *then*, but it can also be interpreted through inferencing as indicating that the prior information leads to the latter-causal/resultative.

The second most frequent multi-meaning LA is *of course*, which is also an LA that occurs less frequently in academic prose, but most predominantly in fiction. In terms of the multiple meanings of *of course*, it can signal either additive or adversative relationship. This observation is in line with Liu's (2008) analysis but we have also found that *of course* in fiction appears most frequently in the initial position of a clause, 57 out of 101 instances (56.44%), while in the other three text genres it tends to occur in the non-initial positions. The high density of *of course* in the initial position is associated with the fact that it often occurs in fictional dialogues, beginning a response to the previous turn, as illustrated in (10).

(10) "That girl can drum."

"**Of course** she can," said Darva.

(AmE06, K08, emphasis added)

In (10), the use of *of course* indicates an emphasis on the girl's ability, following the first speaker's opinion.

In some cases, *of course* indicates adversative relationship, with an ironic and emphatic tint in the speaker's disagreement, as illustrated in (11).

(11) "Time can't stop!"

"**Of course** it can," she said, pretending to be offended.

(AmE06, K26, emphasis added)

Apart from that, it was also found that the use of *of course* is likely to trigger additional meaning varying according to contexts. Consider the following example in (12).

(12) For me, the models were Mike and Elaine, Bob and Ray, and, **of course**, Benny.

(AmE06, K16, emphasis added)

In (12), *of course* generally signals an emphasis on an item following it in relation to the other items on the list; however, it can also express modality in terms of necessity in that it suggests the meaning: *it is necessary for Benny to be included in the list [of people]*. An absence of *of course* thus removes the reading of such modality.

The last multi-meaning LA is *at the same time*. While the other two multi-meaning LAs predominate in fictional texts, *at the same time* is particularly common in academic texts. Interestingly, the case of *at the same time* can illustrate differences in preferred textual positions of LAs when they are used in different text genres. Of 16 instances of *at the same time* in academic prose, 11 cases occur in the initial position (68.75%). In contrast, in fictional texts, only two of 15 instances occur in the initial position (13.33%). This difference in surface patterns is linked with textual meanings the LA expresses.

To illustrate, the use of *at the same time* in the initial position of a sentence in academic prose tends to signal additional pieces of information related to what has been stated in discourse, which simultaneously shows adversative and temporal relationship between propositions, as shown in (13). In this case, *at the same time* in (13) can be said to serve simultaneously as an additive, adversative and temporal LA.

- (13) Bike messengers work in a dirty and dangerous occupation with low pay and no benefits. **At the same time**, many

messengers consider their occupation to be their primary source of identity.

(AmE06, J24, emphasis added)

On the other hand, the use of *at the same time* in fiction tends to occur in the final position, indicating that two states of affairs occur simultaneously, especially those cases with the conjunction *and*, as shown in (14) and (15), respectively.

- (14) Her heart seemed to weep and rejoice **at the same time**.

(AmE06, P03, emphasis added)

- (15) Charlie liked looking at her smile, always had; it felt like approval and permission **at the same time**.

(AmE06, R02, emphasis added)

In (14), the phrase *at the same time* indicates a temporal relationship of the two opposite emotional expressions while simultaneously creating a paradox of reactions, *weep* and *rejoice*. Unlike (14), the use of *at the same time* in (15) links two words that both show positive values, thereby serving to provide temporal and additive links between the propositions involved.

## 6. Conclusion

This study explores distribution patterns and usage profiles of LAs in present-day written American English, in response to a noted lack of empirical study on LAs in American English. The findings from the present study correspond to an observation that LAs are associated with text genres in which they occur. The major tendency found in this study, also in line with previous corpus-based studies (e.g., Biber et al., 1999; Liu, 2008), is that LAs are most common in academic prose. This suggests that language varieties have little role in uses of LAs in the English language, at least in American and British English according to the corpora. Rather, the present study argues that the major factors related to LA uses and patterns are likely to be text genres and LA categories.

The present study not only yields support to the prior observations on a strong association between LAs and academic prose through a look at an American English corpus, but also argues that such an association tends to occur only at a general overall level. This is because actually quite a number of individual LAs are used less frequently in academic prose than in other text genres. This point has rarely



been discussed in the literature on LAs. The present study therefore makes contributions in throwing light on the general picture of LA distribution patterns in written American English and in directing attention to usage profiles of some specific LAs common in text genres other than academic prose.

Since the primary aims of the study are distribution patterns of LAs, thorough qualitative explanations about some individual LAs cannot be offered fully here. The analysis of individual LAs in each group, nevertheless, has raised a very important point to pursue in future research; that is, there may be contextual factors other than text genres that play an essential role in the frequencies, uses, and interpretations of LAs. These contextual factors include textual positions, information progression, and co-occurrence patterns. Further studies can focus on these features of LAs to explore various under-researched aspects of LA uses, ranging from evaluative meanings of LAs, usage differences between synonymous LAs or LA variants, the relationship between textual positions of LAs and text genres or meanings, as well as the relationship between multi-meaning LAs and contexts.

Moreover, although we have chosen a very comprehensive list of LAs, the corpus adopted in this

study, the AmE06, which constitutes approximately a million words, can be considered as relatively small and hence does not show some less common LAs. Those that do not turn up in our corpus can also be studied in future studies using a larger corpus. Another possibility for future research is to develop a more complete list of English LAs to complement the findings on LAs that may occur in the corpora but have not been included in the list.

Finally, if we take an applied linguistic perspective, the study has pedagogical implications. The results from this study can be utilized to develop guidelines for teaching of English LAs. Specifically, distribution patterns of English LAs by text genres can serve as a basis for developing materials on common LAs in authentic texts, which would help learners of English in terms of reading and writing various text genres. Also, findings from this study can be integrated in classrooms for English for Academic Purposes (EAP) or English for Specific Purposes (ESP), for example. It is recommended that the quantitative findings here be calibrated with other factors such as learner needs, learning motivation (Webb & Nation, 2017), and goals of language

learning (Saville-Troike, 2012). Last but not least, the discrepancies observed between frequencies of LA in authentic and taught academic writing can be furthered in ELT research.

## References

- Biber, D., & Conrad, S. (2009). *Register, genre, and style*. Cambridge, UK: Cambridge University Press.
- Biber, D., & Gray, B. (2016). *Grammatical complexity in academic English: Linguistic change in writing*. Cambridge, UK: Cambridge University Press.
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Harlow, UK: Longman.
- Brown, G., & Yule, G. (1984). *Discourse analysis*. Cambridge, UK: Cambridge University Press.
- Celce-Murcia, M., & Larsen-Freeman, D. (1999). *The grammar book: An ESL/EFL teacher's course* (2nd ed.). Boston, MA: Heinle & Heinle.
- Conrad, S. M. (1999). The importance of corpus-based research for language teachers. *System*, 27(1), 1–18. [https://doi.org/10.1016/S0346-251X\(98\)00046-3](https://doi.org/10.1016/S0346-251X(98)00046-3)

- Gilquin, G., Granger, S., & Paquot, M. (2007). Learner corpora: The missing link in EAP pedagogy. *Journal of English for Academic Purposes*, 6, 319–335. <https://doi.org/10.1016/j.jeap.2007.09.007>
- Granger, S., & Paquot, M. (2009). Lexical verbs in academic discourse: A corpus-driven study of learner use. In M. Charles, D. Pecorari, & S. Hunston (Eds.), *Academic writing: At the interface of corpus and discourse* (pp. 193–214). London, UK: Continuum.
- Ha, M. J. (2016). Linking adverbials in first-year Korean university EFL learners' writing: A corpus-informed analysis. *Journal Computer Assisted Language Learning*, 29(6), 1090–1101. <https://doi.org/10.1080/09588221.2015.1068814>
- Hardie, A. (2012). CQPweb — combining power, flexibility and usability in a corpus analysis tool. *International Journal of Corpus Linguistics*, 17(3), 380–409.

- Ishikawa, S. (2010). A corpus-based study on Asian learners' use of English linking adverbials. *Themes in Science and Technology Education*, 3, 139–157.
- Leedham, M., & Cai, G. (2013). *Besides ... on the other hand*: Using a corpus approach to explore the influence of teaching materials on Chinese students' use of linking adverbials. *Journal of Second Language Writing*, 22(4), 374–389. <https://doi.org/10.1016/j.jslw.2013.07.002>
- Lei, L. (2012). Linking adverbials in academic writing on applied linguistics by Chinese doctoral students. *Journal of English for Academic Purposes*, 11, 267–275. <https://doi.org/10.1016/j.jeap.2012.05.003>
- Liu, D. (2008). Linking adverbials: An across-register corpus study and its implications. *International Journal of Corpus Linguistics*, 13(4), 491–518. <https://doi.org/10.1075/ijcl.13.4.05liu>
- Mackey, A., & Gass, S. M. (2016). *Second language research: Methodology and design* (2nd ed.). New York, NY: Routledge.

- Potts, A., & Baker, P. (2012). Does semantic tagging identify cultural change in British and American English?. *International Journal of Corpus Linguistics*, 17(3), 295–324. <https://doi.org/10.1075/ijcl.17.3.01pot>
- Ro, S. H., & Na, Y. H. (2012). Use and misuse of linking adverbials: A corpus-based study on NS and NNS writings. *English21*, 25(3), 291–317. Retrieved from <http://www.earticle.net/Article.aspx?sn=196792>
- Saville-Troike, M. (2012). *Introducing second language acquisition* (2nd ed.). Cambridge, UK: Cambridge University Press.
- Sinclair, J. (1991). *Corpus, concordance, collocations*. Oxford, UK: Oxford University Press.
- Shaw, P. (2009). Linking adverbials in student and professional writing in literary studies: What makes writing mature. In M. Charles, D. Pecorari, & S. Hunston (Eds.), *Academic writing: At the interface of corpus and discourse* (pp. 215–235). London, UK: Continuum.

- Stubbs, M. (2007). Quantitative data on multi-word sequences in English: The case of the word world. In M. Hoey, M. Mahlberg, M. Stubbs, & W. Teubert (Eds.), *Text, Discourse and Corpora: Theory and Analysis* (pp. 163–190). London, UK: Continuum.
- Swales, J. M., & Feak, C. B. (2012). *Academic writing for graduate students* (3rd ed.). Ann Arbor, Michigan: University of Michigan Press.
- Webb, S., & Nation, P. (2017). *How vocabulary is learned*. Oxford, UK: Oxford University Press.