Teaching English for Real-Life Communication with Movie Reviews: A Corpus Approach

*Romina Castagno**
Hawai'i Pacific University

Milang Shin
Hawai'i Pacific University

Yuki Tanaka
Intercultural Communications College, Hawai'i

Abstract

Corpus analysis can inform and enrich language teaching on a practical level. This paper aims to show how to effectively incorporate corpora, which are large electronic collections of authentic language, into vocabulary and grammar teaching while utilizing authentic tasks. By using the concordancer AntConc, we explored the word frequencies of nouns, adjectives, and verbs used in online reviews of fantasy, romance, and action movies written between 2012-2017. Based on the analysis, we suggest corpus-based teaching activities that can help ESL/EFL learners to make use of mini-corpora and to understand frequently used expressions in movie reviews. Such understanding may lead to increased language proficiency in this content domain.

Introduction

The purpose of this paper is to show how corpus analysis can inform and enrich language teaching on a practical level. One benefit that corpus analysis provides is the opportunity for learners to inductively learn features of the target language through their own analysis. Another benefit is the potential to increase student abilities to use corpora, which expands their language learning toolbox and helps them become independent, data-driven learners. In particular, this paper focuses on the teaching applications of a corpus analysis of movie reviews.

Movie reviews have been the target of previous genre analysis (e.g., Taboada, 2011) and corpus analysis (e.g., Bloom & Argamon, 2010; Liang, 2006; Pang & Lee, 2004; Pang, Lee, & Vaithyanathan, 2002; Sorostinean, Sana, Mohamed, & Targhi, 2017) and teachers have introduced movie reviews as data for corpus analysis by learners (e.g., Gavioli, 2009). However, to our knowledge, no previous study has analyzed movie reviews with a focus on the genre of the movies being reviewed and the intention of creating language teaching applications based on this analysis. In this paper, we hope to fill this gap. Introducing movies reviews to language learners is

important for a number of reasons. First, movies are a common topic that learners of English are likely to encounter in real-life communication, and they should be exposed to the linguistic forms related to this topic. Second, in writing movie reviews, students practice a number of language functions, including expressing opinions, making recommendations, and engaging in descriptions. Third, sharing personal opinions by posting movie reviews may provide learners with a sense of belonging in the target language community.

Using corpora for and in a language class involves not only materials which are different from traditional textbook-based teaching but also a different instructional approach. In the next section, we will provide an overview of this instructional approach, namely, Data-Driven Learning.

Data Driven Learning (DDL)

"Data-Driven Learning" (DDL) was coined by Tim Johns (1991), and it proposes an approach to language teaching and learning whereby learners are exposed to large amounts of authentic data and encouraged to discover language patterns on their own. Within DDL, the teacher plays the role of a "research director," who guides students rather than giving them information directly and explicitly. DDL emphasizes the active role of students in learning and the idea that effective learning can take place when learners are in charge of their own learning. Talail and Fotovotnia (2012, p. 1526) pointed out three advantages of DDL learning: developing learner autonomy, enhancing language awareness, and learning to cope with authentic language (see also Gavioli, 2009). Thus, applying DDL to movie reviews will give students the opportunity to become independent learners with an increased awareness of target language forms and an increased capacity to use the language of this genre. Further, the ability to identify important components of this genre, such as its vocabulary, grammar, collocations, and typical formulaic phrases, will help students recognize the importance of these language components in any genre, which may help them adapt their language use better according to context. In DDL, a heavy amount of input is needed to understand how language works. Specific or general purpose corpora can be used to provide this input.

Specialized Corpora

According to Connor and Upton (2014, p. 2), a general purpose corpus is beneficial because it supports language learning by providing language samples in a wide range of contexts. However, to conduct language analyses for specific academic or professional purposes, one needs specialized corpora, or corpora that focus on a specific genre or domain. Unlike general purpose corpora, which are collections of language from a variety of situations, specialized corpora are collections of language from a specific domain. A corpus of a given domain has its own characteristics and functionality (see also Warren, 2010). Connor and Upton (2014, p. 2) explained that greater attention has been paid to specialized corpora for language learning as a means to help learners develop the ability to use specific language forms in particular genres or situations more effectively. Our research draws on a specialized corpus for the following pedagogical advantages. The primary advantage is that a specialized corpus allows learners to focus on language in its full context. In addition, specialized corpora afford the opportunity to help learners investigate linguistic features of a familiar genre. Finally, its smaller size and greater

context can facilitate learners' access to target words and grammatical points in a particular domain. Timmis (2015, p. 14) listed the following useful functions that a specialized corpus built by teachers or learners may have:

1. To inform students about the language used in a specific *teaching genre*, such as hotel reviews, business reports, recommendations, etc.
2. To indicate the *features* that will be taught
3. To provide a direct source of *materials*, i.e. DDL
4. To provide a *resource* for learners to use autonomously which is appropriate in terms of topic, level, and accessibility
5. To investigate *learners'* language

These purposes are not mutually exclusive, and a well-designed corpus may maintain one or more of these functions. Our corpus encompasses four of these functions (1, 2, 3, 4). The primary focus of our corpus is to inform learners of the commonly used real-world language in movie reviews and to improve students' independence in learning by showing them how to create their own corpus of language for any genre or domain.

Movie Reviews As Teaching Materials

Richards and Rodgers (2001) stated that the primary purpose of teaching materials is to provide students with the appropriate amount of input and instruction as well as to ensure students receive enough practice. Further, teaching materials need to be adequate as materials-in-action, meaning that the language and content should be as they are used in real-life situations. Reviews of films can be a valuable teaching tool for developing the four language skills, critical thinking skills, and cultural understanding. For example, a documentary like *The Corporation* poses profound questions about social and political problems. This type of movie can help learners by engaging students in oral discussion and/or the written expression of their opinions to develop their critical thinking skills. Moreover, movie reviews can contribute to speaking and listening skills in various ways, as students have the opportunity to read movie reviews, watch and listen to movie trailers, and discuss movies and their reviews in the language classroom. Movie reviews can be powerful language teaching materials, as they typically contain language found in daily speech. Further, each movie genre also has its own specific configuration of vocabulary and grammatical features. Since movie reviews describe movie features unique to that genre, using them as teaching materials helps students recognize vocabulary that is often used in describing movies of each genre. Further, in doing their own corpus analysis of movie reviews, learners can develop the ability to inductively recognize language patterns common to a particular domain, while reading varied opinions of movie reviews may prompt learners to analyze films from different perspectives.

Research and Practical Questions

With an interest in developing teaching materials and activities grounded in the ideas explored above, we posed the following questions:

1. What are the frequent words commonly used across reviews of three movie genres—action, fantasy, and romance—and what are their common collocates?

2. What are the keywords that distinguish each genre and what are their collocates?
3. What teaching materials and activities can be created from the corpus findings?

Method

Data

Limiting our focus to reviews of romance, action, and fantasy films, we collected 79 real-life movie reviews in the past five years (2014-2018) from the website Rotten Tomatoes ((https://www.rottentomatoes.com). These reviews include 18 action movie reviews, 24 romance movie reviews, and 37 fantasy movie reviews. The number of reviews for each genre is not the same due to variation in review length. Reviews for romance movies were shorter than those for action movies, and fantasy movie reviews were even shorter than romance movie reviews. Therefore, we collected a higher number of fantasy and romance reviews to reach a similar word count across the different categories of movie reviews. The final corpus included a total of 22,038 tokens (4,029 types). For this paper, each genre was also compared against the other two.

Analytical Procedure

The texts we chose for analysis were mainly movie reviews that contain more than two sentences with a positive, negative, or neutral rating. The movie reviews chosen did not contain only information about the movie itself; they also referred to personal opinions, feelings, and related movies. All movie reviews were written by ordinary people, as we chose not to gather those written by expert critics. In order to analyze the texts of this specialized corpus, we used the AntConc concordancer (version 3.5.0, available at http://www.laurenceanthony.net/) designed by Laurence Anthony (2011). AntConc is a concordancer computer program which presents systematically organized data of a user-generated corpus through its program tools, allowing users to identify language features and their usage from data that they upload. Once the texts for analysis were chosen, the corpora were uploaded onto the program. Reviews of fantasy, romance, and action movies were collected and saved as separate sub-corpora and then loaded onto Antconc. The basic tools of AntConc we used for our analysis were *word list*, which presents the most frequent words in the corpus along with their frequencies, *concordance*, which allows users to view concordance lines of selected words and their context; and *keyword list*, which yields the words which make the corpus unique compared to a reference corpus and a measure of how unique they are to the corpus. In our analysis, we focused on the following:

1. Top ten most frequent function words in all movie reviews
2. Top ten most frequent content words in all movie reviews
3. Collocations of the most frequent content word
4. Top ten determiners, pronouns, and verbs
5. Keyness of nouns across the three movie review sub-categories
6. Keyness of adjectives across the three movie review sub-categories

Findings
Top Ten Most Frequent Function Words in the Movie Reviews

First, we used the word list function to generate a list of the most frequently used words across all

collected movie reviews. As is typically the case for many corpora, the top ten words in our corpus were all function words. The definite article *the* was the most frequent word in the corpus, and the indefinite article *a* was the second most frequent. The rest of the top ten words are all conjunctions, prepositions, and pronouns (see Figure 1).
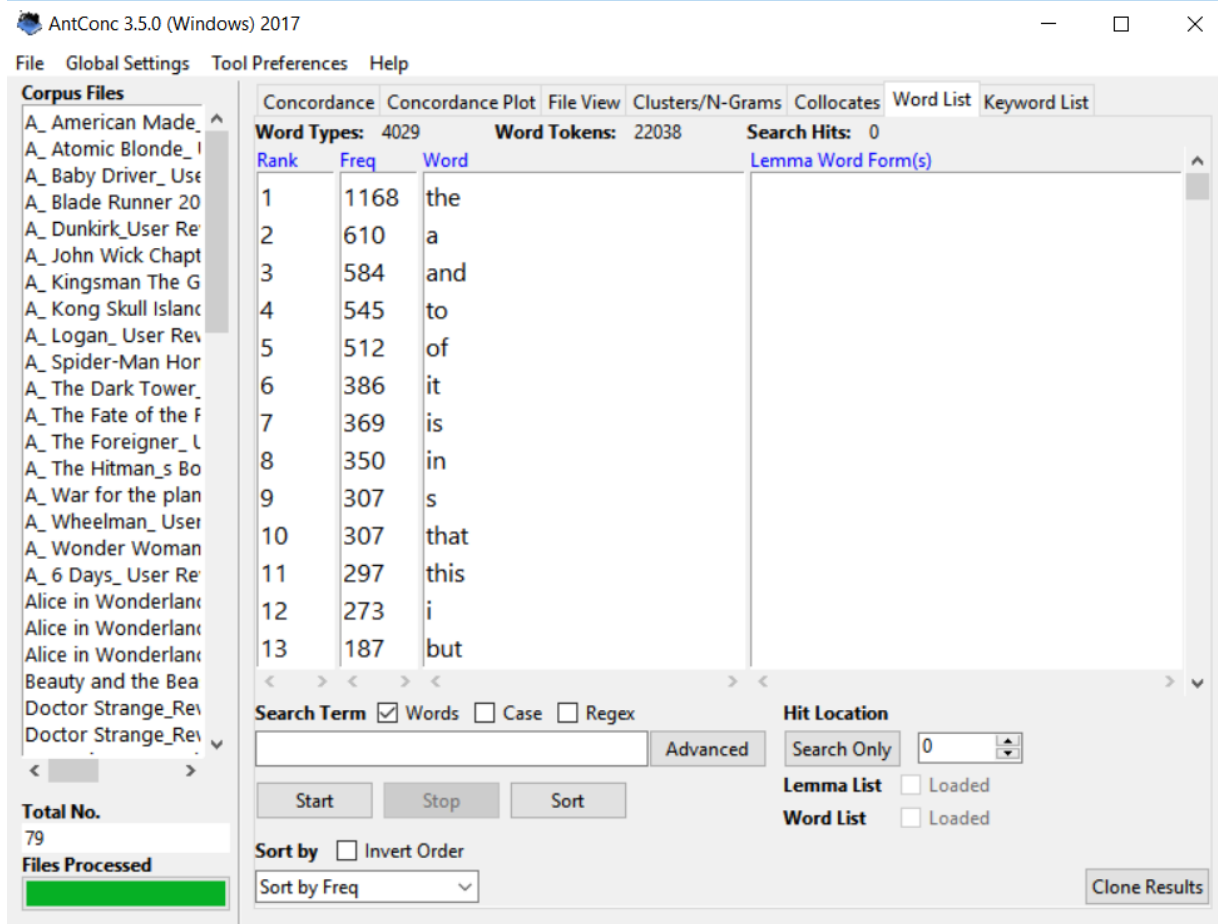

Figure 1. Top Ten Words Used in Movie Reviews

Top Ten Most Frequent Content Words in the Movie Reviews

We manually collected the ten most common content words from the overall frequency list (Table 1). The words *movie* and *film* are often used for identification as in "this *movie* is good." Other words in Table 1, such as *like* and *good*, suggest that movie reviewers often express opinions and evaluate in writing their reviews. The adverbs *just, really,* and *very* are also commonly used for evaluation. Knowing the most frequent words can be useful for learners because they have a high chance of encountering them in this context in real-life situations.

Table 1
Top Ten Content Words

Top Ten Content words
1. *movie*
2. *film*
3. *have*
4. *like*
5. *just*
6. *has**
7. *really*
8. *time*
9. *very*
10. *good*

Note. **The total frequency of *has* may be due to its usage as both a content word and a function word.

Collocations of the Most Frequent Content Word

As has been found in previous corpus analyses of movie reviews (e.g., Sorostinean, Sana, Mohamed, & Targhi, 2017), the most frequent word in the corpus is *movie*. We will analyze its collocates to ascertain its semantic prosody. For this analysis, we used the concordance function in AntConc and executed a KWIC search with the parameters 1L, 1R, or 2L, 2R. Figure 2 shows how the word *movie* is used in movie reviews.

In terms of collocates, the word *movie* is frequently associated with evaluative adjectives, such as *bad, best, fun, good, stupid,* etc. *Movie* is also often collocated with words used to describe or categorize movies, for example, *family, fantasy,* and *superhero,* etc. Additionally, the word *movie* is commonly used in describing movies in a sequence such as *first, final,* and *new.*

By analyzing the collocates of *movie*, we sought to identify its semantic prosody. The word *movie* frequently co-occurs with positive adjectives, such as *amazing, best, clever, excellent, entertaining, good, great, greatest, gorgeous, loving,* etc. At the same time, the word *movie* frequently co-occurs with negative adjectives, such as *cheesy, cringe-inducing, disliking, poorly-done, stupid, sick, terrible, unnecessary,* etc. This evidence shows that *movie* does not have a positive or negative semantic prosody in and of itself; it can be used with both positive and negative modifiers.

Overall, drawing from this analysis, learners can be led to identify the most frequent adjectives and fixed expressions in order to use them both in everyday conversations about movies and in crafting their own movie reviews.
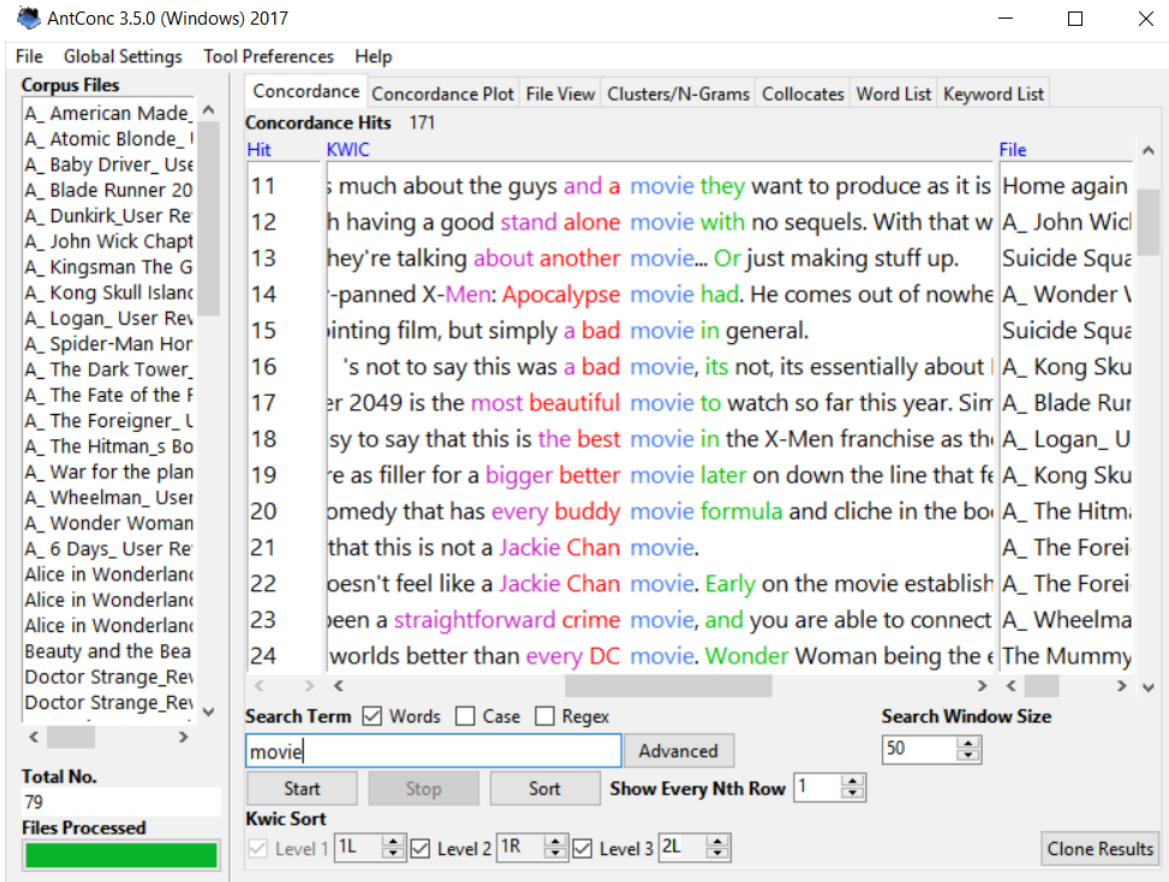
Figure 2. Collocations of *movie*, the Most Frequent Content Word

Grammatical Features of Movie Reviews

Table 2

Top Ten Determiners, Pronouns, and Verbs

Determiners	Pronouns	Verbs
1. *the*	1. *it*	1. *is*
2. *a*	2. *I*	2. *was*
3. *that*	3. *you*	3. *are*
4. *this*	4. *he*	4. *have*
5. *one*	5. *his*	5. *like*
6. *an*	6. *they*	6. *has*
7. *there*	7. *their*	7. *can*
8. *all*	8. *we*	8. *been*
9. *some*	9. *her*	9. *will*
10. *more*	10. *my*	10. *get*

Definite Article **the**

Definite article *the* is the most common determiner, which usually refers back to a referent that reviewers and readers are familiar with in general or from the preceding context. In movie reviews, it is used most frequently in one of two ways.

The word *the* in Example Set 1 is used as part of a superlative. For instance, *the* belongs to the superlative form in cases such as "*the* most impressive" and "*the* best".

Examples Set 1

1.1 Easily, *the* most impressive aspects throughout this film were *the* ways the camera was used throughout *the* intense moments. (Action)

1.2 She provides most of the humour in this movie, and executes her role very well, easily *the* best performance in the film. (Romance)

1.3 A Silent Voice is one of *the* best movies I've seen so far this year. (Fantasy)

1.4 As long as you can buckle up and be ready for *the* most ridiculous action yet, I believe you will have a blast with this storyline, while also being treated with quite a few heartfelt moments that this series has based its core story around (family) [SIC]. (Action)

Through these superlative constructions, reviewers express their own opinions or evaluate movies. The high frequency of *the* in superlatives may be due to the fact that people tend to use the superlative structure to emphasize their own reactions, which may be positive or negative towards one particular movie and not another. This construction is mainly used to describe and the movie's overall quality.

The word *the* in Example Set 2 is used to indicate a specific feature. Here, *the* is used to refer to specific entities, such as movies or characters, and to indicate the same referent more than once.

Example Set 2

2.1 Burton's more Gothic look at Alice in Wonderland just doesn't make me a fan, but that's not saying that others won't enjoy *the* film. (Fantasy)

2.2 While nothing really happened to get me extremely into *the* movie, I definitely remained entertained throughout the entire thing (Romance)

2.3 Jared Leto is granted limited screen time but still manages to leave his mark as *the* sinister Niander Wallace. (Action)

2.4 "*The* Hitman's Bodyguard" is a buddy, action-comedy that has every buddy movie formula and cliche in *the* book. (Action)

2.5 If you enjoyed previous night at the museum [SIC] movies you enjoy this one as well. It's not as good as *the* first one but it's better than *the* second one. The movie is cute with a great cast even if a lot of the jokes come off pretty cheesy. (Fantasy)

Examples 2.1 and 2.2 in this set indicate a particular movie previously mentioned by the reviewers. Example 2.3 describes a character who showed up and how that character was related to others in the film. Example 2.4 illustrates the use of *the* in *The* Hitman's Bodyguard" as part of

a movie title. Example 2.5 illustrates how *the* may be used to indicate movie sequels. Generally, in reviewing one movie in a series, reviewers will mention other movies for comparison, such as was the case for reviews of *Harry Potter, Bridget Jones's Diary,* and *Kingsman*.

Indefinite Article a

Figure 3 displays the top ten collocates of the indefinite article *a*.
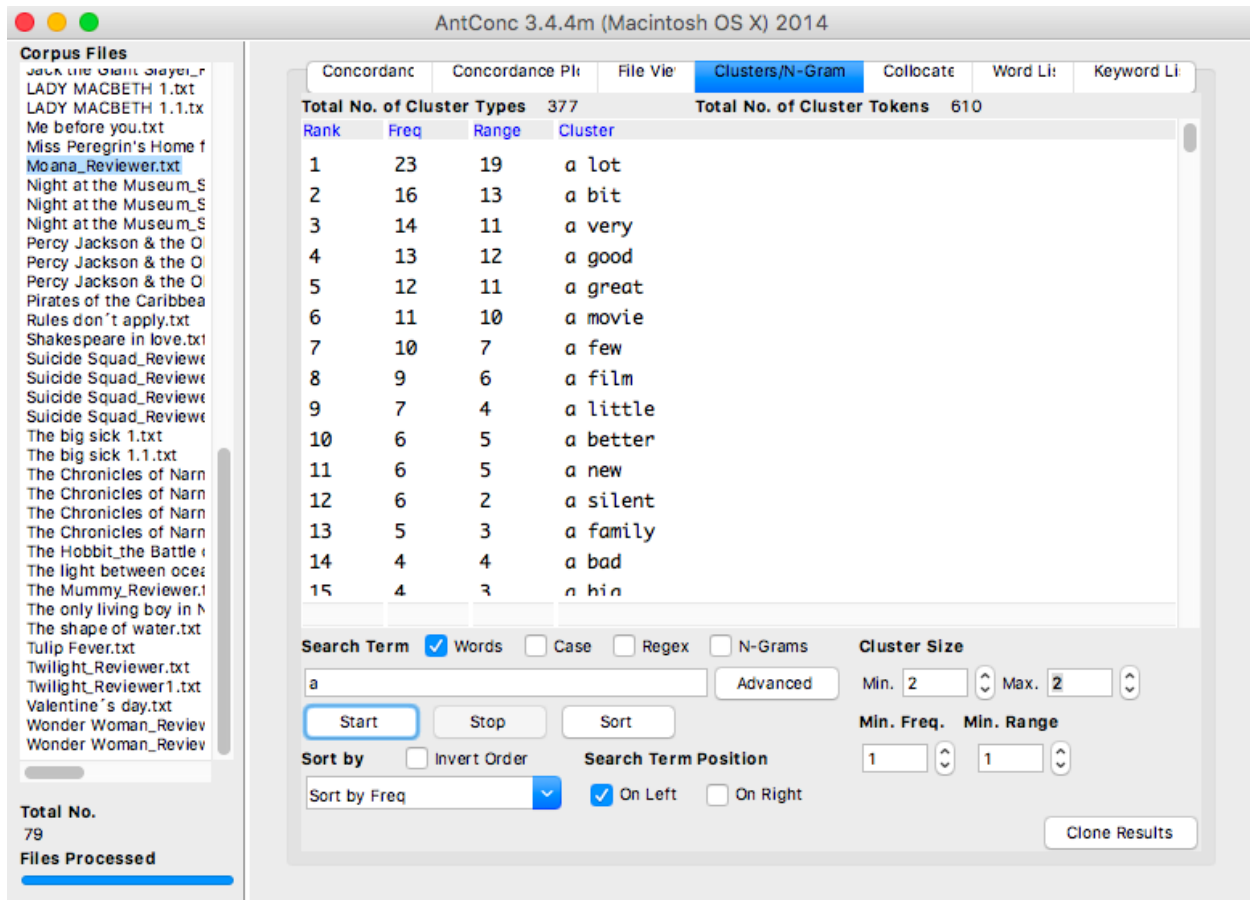

Figure 3. Top Ten Collocates with the Indefinite Article *a*

The indefinite article *a* is found in several common expressions which are mostly made up of expressions of quantity (Example Set 3). It is also often used in the corpus to comment on unspecified details of the movie being reviewed (Example Set 4). Here, movie reviews typically describe a key point, a signature scene, or a character to provide the reader with a type of summary and the reviewer's opinion without giving away too much about the movie. In addition, indefinite article *a* is usually used with adjectives or adverbs that modify nouns. (Example Set 5). They tend to evaluate the entire movie in these cases.

Example Set 3
>3.1 *A bit of* a predictability with the overall story I will admit and even felt like a typical Disney flick in some cases, but this still stands out pretty well on it's own much like previous films by Disney Animation. (Fantasy)
>3.2 He's abley surrounded with an impressive and colourful collection of support as well; Jon Bernthal, Jon Hamm and Jamie Foxx trying to outdo each other in the menacing stakes is *a lot of* fun in itself and it's great to see them spearheaded by the reliable and infinitely watchable Kevin Spacey. (Action)
>3.3 Hellboy II: The Golden Army still has overwhelming action and pretty weak character development, but it is ultimately an improvement over the original with a much more entertaining plot with a good fantasy angle to it and *a lot of* great set pieces, interesting creatures and great visuals. (Fantasy)

Example Set 4
>4.1 At the same time, he cares for Professor Xavier, one of the most powerful mutants alive, who is now old and suffering from *a* brain disease and seizures which, when they hit, cause people within *a* surrounding area to go into *a* comatose state. (Action)
>4.2 Even a more interesting subtext, punishing his father for putting their family dynamic at risk, is only kept at *a* distance. (Romance)

Example Set 5
>5.1 This is a spectacular film from start to finish. Keep in mind, this can be *a very* tough film to sit through for multiple reasons, and that is exactly what I'm going to dive into throughout this review, because I feel that a few precautions should be taken when going in. (Fantasy)
>5.2 Narnia Prince Caspian was *a good* Fantasy movie I like all of this movie from start to finish. (Fantasy)
>5.3 Renee Zellweger (Bridget Jones) does *a great job* in this movie, she provides most of the humour in this movie, and executes her role very well, easily the best performance in the film. (Romance)

Movie reviews include many descriptions about films to supply useful information for users. The indefinite article *a* is frequently used with evaluative and descriptive adjectives. In short, movie reviews frequently contain detailed evaluations and descriptions of films, as people are likely to voice opinions about films through evaluative language and to recount movies through descriptive language.

Complementizer, Relative Pronoun, Demonstrative Pronoun, and Determiner **That**

That often functions as a complementizer to introduce a noun clause (Example Set 6) or as a relative pronoun to introduce a relative clause (Example Set 7).

Example Set 6
>6.1 It's not *that* I hated watching this film, but unlike a John Wick that gets straight to the

point, this film felt very bogged down, trying to complicate itself in twists and turns. (Action)

6.2 I have to admit *that* I did find the movie slow at times. (Romance)

6.3 Burton's more Gothic look at Alice in Wonderland just doesn't make me a fan, but that's not saying *that* others won't enjoy the film. (Fantasy)

Example Set 7

7.1 He's a director *that* can seemingly do no wrong in many people's eyes but this enthusiasm is one *that* I've often questioned. (Action)

7.2 However, this film wastes a lot of time with unnecessary storylines *that* don't always end up going anywhere significant. (Fantasy)

7.3 In the end, this is a truly remarkable piece of dramatic filmmaking and there are some very memorable scenes *that* keep replaying [SIC] themselves in my head. (Romance)

Finally, *that* is also often used as a demonstrative pronoun in the movie reviews we have examined. In this function, it sums up a point made about an aspect of the movie, such as characters, scenes, directors, music, and so on (Example set 8).

Example Set 8

8.1 It's nearly impossible to not compare Atomic Blonde to John Wick, as the styles are very similar and the trailers make it out to be in the same vein, but they are very different films, so please don't be fooled by *that*. (Action)

8.2 The Columbus architecture was nice to see, but *that*'s it. (Romance)

8.3 I found myself disliking the movie throughout the majority of it, until my eventual realization, which is my one warning, so I feel I won't be alone on *that*. A Silent Voice is one of the best movies I've seen so far this year. (Romance).

In short, the top three function words *the*, *a*, and *that* have various functions. When people express their own opinions or describe characters, scenes, and movies, the determiner *a* can be used to express new information. The determiner *the* is used for detailed information. Finally, *that* is used to introduce the writer's opinions and provide further descriptions of movies, characters, scenes, etc.

Pronouns

The top three pronouns in movie reviews are *it*, *I*, and *you*. The frequent use of *it*, which often refers to movies as a subject or object, indicates that the reviews are often about one movie at a time. The frequent use of *I* suggests that the reviewer's perspective is salient in movie reviews. Finally, the frequent use of *you* indicates that the reviews address the audience directly. The frequent use of first and second person pronouns suggests that movie reviews are personal and interactive.

In the movie reviews we examined, *it* is used anaphorically to refer to something previously mentioned in the review (Example Set 9).

Example Set 9
>   9.1 Everybody needs that one movie in their collection that they know is bland and not very well made, but they love watching *it* because *it* makes them smile and laugh. (Romance)
>   9.2 The pacing is off and *it* definitely could've benefitted from a shorter run time. (Action)
>   9.3 The visual effects looked amazing, and *it* honestly looked better than some films today. (Fantasy)
>   9.4 So with that the live-Action adaptation of Beauty and the Beast it's not only able to stay true to the original but is able to add to *it* to make like a new movie. (Fantasy)
>   9.5 This is her movie, and she owns *it* from start to finish. (Romance)

The pronoun *it* is also used as a dummy pronoun in general statements made in movie reviews (Example Set10).

Example Set10
>   10.1 *It* is easy to say that this is the best movie in the X-Men franchise as the majority of their films are pretty easy to top. (Action)
>   10.2 To add to that, *it* was clear that certain scenes or lines of dialogue were added to the film in order to give it a sense of realism. (Romance)
>   10.3 There are some definite forced aspects about this film, but when everything is filmed and presented so wonderfully, *it*'s hard to let that clash with your enjoyment of the movie as a whole. (Romance)
>   10.4 It's honestly hard to watch all these beloved characters fight out a situation so dark and brutal. (Fantasy)

The frequent usage of the pronoun *I* comes from the subjective and opinion-based aspects of movie reviews. In this domain, people mainly express opinions or evaluate films based on their own thought or experiences. Movie reviews are subjective; hence, the pronoun *I* is very common (Example Set 11).

Example Set 11
>   11.1 The themes presented throughout this film are taken very seriously, so *I* don't think it will appeal to a younger audience, but *I* think older teenagers will be the perfect demographic here, for the lesson it's giving. (Romance)
>   11.2 *I* felt epic while doing so. *I* guess what my point is, the music was well composed, it fit perfectly with the tone of the film, and it was memorable. (Fantasy)
>   11.3 Personally, *I* left this film not quite understanding where some of the praise is coming from, because *I* found myself incredibly bored throughout the majority of the moments that people are finding interesting. (Action)

The pronoun *you* is frequently used in movie reviews when reviewers comment on what they believe readers might do or how they expect or imagine readers will react, think or feel (see

Example Set 12). Moreover, the pronoun *you* represents people in general (Example Set 13).

Example Set 12
>   12.1 Without a doubt, *you* will walk out of this film remembering the action sequences, because they are incredibly choreographed, but that's not quite what I will remember this film for. (Action)
>   12.2 *You* might find yourself another guilty pleasure just like I have. (Romance)
>   12.3 In the end, A GHOST STORY will have *you* thinking, have *you* discussing about life & time. (Fantasy)

Example Set 13
>   13.1 I think what makes this movie great is that it could also work as a silent film and *you* would still understand what's going on. (Fantasy)
>   13.2 When a movie makes *you* wait for it to be really good, it can be a pretty frustrating experience, but I see something different when reflecting on this film. (Romance)
>   13.3 I found myself absolutely hating his character, which was very frustrating, due to the fact that the film asks *you* to sympathize with him on multiple occasions. (Romance)
>   13.4 If *you* watch the trailers for this, they promise *you* Jeff Bridges and Channing Tatum as American cowboy spies. (Action)

Function Verbs

Focusing on the ten most frequent verbs (Table 2), we can see that present simple, past simple, and present perfect tenses are frequently used in the corpus.

First, the verb *be* is dominantly used among the top five verbs, appearing in different forms (*is*, *are*, *be*, *was*). In most concordance lines, these *be* verbs were used to to make suggestions or to describe the plot, effects, main character, feelings, sequels, or book which a movie represents. The third person singular *is* occurs with high frequency in the corpus. It tends to co-occur with adjectives, such as *good, amazing,* and *pretty,* which are also on the list of top 20 content words by frequency. The third person singular *is/was* frequently occurs in sentences as the linking verb of evaluative subject complements.(Example Set 14). The following construction is often used in movie reviews: *is/was* + a film, a story, or a character's name (Example Set 15). Interestingly, the types of adjectives and nouns used are influenced by movie genres, which can be seen in the section on keyness which follows.

Example Set 14
>   14.1 The casting for the movie *is* amazing with having Emma Watson as belle with how much she really looks like her and with all of the other very talented cast members doing an amazing job at their roles. (Fantasy)
>   14.2 Starring Dakota Johnson, Rebel Wilson, Alison Brie, and Leslie Mann, the casting *is* pretty good. (Romance)
>   14.3 The final twist (and villain) *was* so goddamn ridiculous I nearly burst out laughing in the theater. (Action)

14.4 All in all, Tulip Fever" *is* [a] disappointment. (Romance)

Example Set 15
      15.1 But it *is* Harry Potter who may be called upon to make the ultimate sacrifice as he draws closer to the climactic showdown with Lord Voldemort.(Fantasy)
      15.2 For the villains, the standout *is* Boyd Holbrook as Donald Pierce, the leader of the militant Reavers group who is sent to retrieve Laura. (Action)
      15.3 This *is* a film about time, love, loss, and legacy (there's a deep and honest monologue midway discussing legacy, which is brilliantly delivered by Will Oldham). (Romance)

A second relevant point about Table 2 is that the verb *have* is one of the most frequent verbs across the movie genres represented in the corpus. The reason why it is frequent is that the verb *have* performs various grammatical functions, such as expressing aspect (Example Set 16) or obligation (Example Set 17), and various semantic functions, such as expressing causality (Example Set 18) and possession (Example Set 19).

Auxiliary verbs marking aspect are one of the most frequent in our data. In particular, use of present perfect often allows reviewers to draw on their past moviegoing experiences or opinions formed through such experiences when writing a new review (Examples 16.1 and 16.2).

Example Set 16
      16.1 For years, we *have* seen Wolverine use his claws in various films but never to the level as seen in "Logan". This violence stresses the brutality this man *has* experienced and *has* carried with him for the hundreds of years that he has been alive. (Action)
      16.2 Not to single out this particular premise, but films revolving around gay couples *have* been terrific over the last decade or so. (Romance)
      16.3 In the epic finale, the battle between the good and evil forces of the wizarding world escalates into an all-out war. The stakes *have* never been higher and no one is safe. (Fantasy)

Have is included in the modal expression for obligation (*have to* + verb) in the corpus (Example 17).

Example Set 17
      17.1 Fans definitely receive all of the above and more throughout The Fate of the Furious, but to be perfectly honest, you *have* to watch these films as if you are sitting in the theatre watching a Marvel film. (Action)
      17.2 There are very long sequences of silence, making you relax and take in either a romantic moment or the scenery around them, which I also found poetic at the same time, so I actually have trouble coming up with issues, but I *have to* admit that I did find the movie slow at times. (Romance)

Have is also frequently used as main verb with a causative meaning to express the effects of the review on the readers or the movie itself on the viewers (Example 18).

Example Set 18

    18.1 Now I'm certain that I will *have* people upset with me because I didn't like the movie or the book and don't consider it worthy of being a big budget motion picture… (Romance)

    18.2 In the end, A GHOST STORY will *have* you thinking, *have* you discussing about life & time. (Romance)

Example Set 19

    19.1 Domhnall Gleeson is also really good as the CIA guy who recruits Barry Seal and he and Tom Cruise *have* nice chemistry together. (Action)

    19.2 I *have* a couple of nits with Kumail's acting at certain parts (which is odd considering he plays himself), but this is a genuine heartfelt movie with a lot of original hysterical moments. (Romance)

    19.3 While it may *have* a few occasionally uneven narratives, Night at the Museum: Secret of the Tomb still serves as a fast-paced, adventurous, funny an [and] reasonably satisfying conclusion to the trilogy… (Fantasy)

In brief, the grammatical and semantic properties of *have* can differ by context. In this corpus, *have* functions to mark aspect, show obligation, indicate causal relations, and express other meanings. Many of these functions have their own characteristics in the context of the corpus. For example, it is clear that reviewers tend to select the modal verb *have to* to indirectly express positive evaluations by asserting obligation (that readers must see the movie).

Keyness in Reviews of the Three Movie Genres

Words with high keyness occur at higher frequencies than expected in a target corpus relative to a reference corpus, thus indicating that such words or phrases make up the defining features of the corpus. We compared the three sets of reviews grouped according to movie genre to see whether the top keyness words would differ. In our analysis, we set up the reviews of each movie genre as target sub-corpus. Then, for each sub-corpus, we used the combination of the two other sub-corpora as a reference corpus in order to identify the words which make the target sub-corpus unique. For example, we used romance movie reviews as a target sub-corpus and action and fantasy reviews as the reference corpus.

By using the keyness function in Antconc, we identified nouns, pronouns, and adjectives with high keyness in the movie reviews for each genre. The unique patterning of nouns, pronouns, and adjectives in each genre will be discussed in the following.

Keyness of Nouns & Pronouns in Reviews of the Three Movie Genres

Table 3 lists the top keyness words in each movie review sub-category. Since the list contains pronouns and proper nouns, such as movie characters' names, we decided to look further for more content words to see whether there were any patterns that characterized a certain genre.

Table 3
Keyness of Nouns in Reviews of Each Genre

Action	Romance	Fantasy
he/his	*I/me/he/him/her*	*Disney/Potter*
Logan Jackman	*Anastatia/Christian*	*fantasy*
action	*love*	*effects*
war	*relationship*	*cgi*
military	*romance*	*adventure*
CIA	*marriage*	*prince*
guns	*affair*	*trilogy*
sequence	*sex*	*superhero*
blood	*couple*	*monster*
terrorists	*boyfriend*	*war*

In the action movie reviews, the nouns with high keyness are frequently associated with words whose meanings are related to conflict, such as *war, guns,* and *blood*. The pronouns with high keyness in the action reviews were the words *he* and *his*, and the proper nouns with high keyness are *Logan,* and *Jackman*, both of which are associated with men and masculinity.

In romance movies reviews, the nouns with high keyness are commonly associated with emotions and relationships, such as *boyfriend, love, romance,* and *relationship*. The pronouns *I* and *me*, which are used to convey aspects of the self, are frequently used in romantic movie reviews to express feelings and emotions connected to the movies or its characters. Words such as *relationship, marriage,* and *couple* are used to describe romantic relationships between characters in the movies.

In fantasy movies reviews, words with high keyness have meanings associated with the surreal or magical, such as *cgi, superhero,* and *monster*. Movie reviews in fantasy genre evidence less frequent pronoun usage compared to the other two genres. The high keyness index of the noun *trilogy* may indicate that fantasy movies are often parts of trilogies.

The findings above can help students become aware of the unique features of each movie review type and inform them about which words to use when they write a review of a movie in one of these genres.

Keyness of Adjectives in Reviews of the Three Movie Genres

Table 4 shows the adjectives with the highest keyness in each sub-corpus. In this list, we categorized the adjectives used in movie reviews into two types, Evaluative (E) and Descriptive (D). Some adjectives have been coded as both E and D. We further classified the adjectives as Positive (red color), Negative (blue color), or Neutral (black color).

Table 4
Top Keyness of Adjectives in Reviews of Each Genre

Action	Romance	Fantasy
ridiculous (E)	*romantic* (D)	*amazing* (E)
criminal (D)	*single* (D)	*great* (E)
dark (D)	*silent* (D)	*good* (E)
cool (E)	*sick* (E & D)	*entertaining* (E)
killed (D)	*clever* (E)	*favorite* (E)
contrived (E)	*stupid* (E)	*greatest* (E)
alive (D)	*wrong* (E)	*animated* (D)
black (D)	*best* (E)	*boring* (E)
cinematic (D)	*common* (D)	*fantastic* (E)
main (D)	*new* (D)	*beloved* (E)

Notes. E = evaluative, D = descriptive, red text = positive, blue text = negative, black text = neutral

According to Table 4, descriptive adjectives are used more frequently than evaluative adjectives in action movie reviews, with seven out of ten adjectives on the list being descriptive. This finding is in line with previous corpus analysis of sentiments in movie reviews (e.g., Liang, 2006). In regards to connotation, the adjectives with high keyness in the action movie review sub-corpus have negative connotations such as *ridiculous, criminal, dark,* and *killed*. Adjectives with negative connotations are used to describe movie contents as well as the quality.

In the romance movie reviews, descriptive adjectives are frequently used. Seven out of ten adjectives on the list are descriptive, while *sick* can be used either as an evaluative or a descriptive adjective. In terms of connotation, four out of ten adjectives with high keyness have neither a positive nor a negative connotation. The rest of the adjectives with high keyness reveal a balance of negative and positive connotations in the romance movie reviews, with three positive and three negative adjectives on the list. The adjective with the highest keyness, *romantic*, is often used to describe a story or scene in a movie or the movie itself. Consider this concordance line, in which *romantic* describes a scene: "*There are very long sequences of silence, making you relax and take in either a romantic moment or the scenery around them, which I also found poetic at the same time…*" In addition, the following concordance line serves as an example of *romantic* being used to describe a movie: "*It was nice to see a romantic comedy!*"

In fantasy movie reviews, Table 4 indicates that evaluative adjectives are dominantly represented in the top words with high keyness. Adjectives used in fantasy films tended to have positive connotations such a*s amazing, great, good* and *entertaining*. The evaluative adjectives, such as *amazing*, are usually deployed to express reviewers' opinions on movie characters, casting, acting, soundtrack, visual effects, and the movie scenes (see also Bloom & Argamon, 2010, pp. 256-257).

To sum up, each movie review sub-corpus displays distinct adjective usage. Action movie reviewers tend to use descriptive adjectives and adjectives with more negative connotations. Adjectives with high keyness in romance genre include both descriptive and evaluative adjectives with varying connotations. In fantasy movie reviews, adjectives with high keyness tend to be

evaluative and to have a positive connotation.

It appears that adjective choices are not related to the movie ratings in the reviews. In the corpus, the action movie reviews had an average rating of 6.85 out of 10; the romance movie reviews had an average rating of 6.34 out of 10; and the fantasy movie reviews had an average rating of 6.17 out of 10. This means that regardless of the reviews' rating of the movies, the action movie reviews contained more negative adjectives whereas the fantasy movie reviews had more positive adjectives.

In many reviews of all three movie genres, descriptive adjectives tend to be used to illustrate a movie's content, such as the plot, characters, and scenes, and evaluative adjectives are frequently used to offer opinions or evaluations on those features. With this in mind, teachers should inform their students about the relationships between reviews of different movie genres and the linguistic patterns they contain.

Discussion

In the analysis above, we have shown the top ten most frequent function and content words in a corpus of movie reviews. Then, we focused on four features of the corpus: collocates of the most frequent content words, contexts of use, grammatical features, and keyness of nouns and adjectives in three genres (action, romance, and fantasy). Regarding grammatical features, we explored the usage of high frequency determiners, pronouns, and verbs. The analysis suggests frequent use of *the* to specify, *it* to restate, and *have* to encode a number of grammatical and semantic properties. Collectively, reviewers drew on these resources to state their opinions, express their feelings, and share their experiences. Each type of movie review has its own pattern of noun usage: conflict-orientated nouns for action movies, relationship-oriented nouns for romance movies, and surreal-oriented nouns for fantasy movies. Each type of movie review also has a unique pattern of pronoun usage; in action movie reviews, masculine pronouns and proper nouns are most frequent, whereas personal pronouns were most frequent in romance movie reviews. An analysis of fantasy movie reviews reveals less pronoun usage compared with the other types of reviews. As for adjective usage in each type of movie review, action movie reviews show a strong tendency to contain descriptive adjectives and adjectives with negative connotations, while fantasy movie reviews tend to contain more evaluative and positive adjectives. Romance movies tend to contain both descriptive and evaluative adjectives, many of which are neutral.

Based on these findings, we will present three teaching ideas that make use of authentic language from the website Rotten Tomatoes® and the corpus tool Antconc.

Pedagogical Applications

The following teaching activities can be adapted according to students' proficiency, age, and interest.

Teaching Activity 1: Keyness (50 minutes)

Before the main activity:

1. Ask the students to pick a movie genre.
2. Group the students according to the genre they chose.
3. Ask each group to choose a movie in the genre they chose to watch on their own.

4. Ask the students to watch the selected movie on their own time.
5. Ask students to read and bring to class at least 5 reviews of the movie they just watched.

Activity:
1. Have students underline all the nouns and adjectives that are unique to the movie genre in the reviews they bring to class.
2. Provide students with the list of nouns and adjectives ranked by keyness from the three movie review categories (Tables 3 and 4 above).
3. Ask students to mark the nouns and adjectives from the tables that do appear on the 5 reviews they have collected.
4. For words that do not appear in reviews collected by students, give examples of how they are used in movie reviews.
5. Ask students to write their own review of the movie that they watched by incorporating words with high keyness and language patterns noticed through the corpus data.

Teaching Activity 2: Common Verb Constructions (20 minutes)
When there is not enough time to have students perform corpus analysis themselves, teachers can prepare concordance lines of various types of verb, such as *get, make*, and *have*, organized by movie review type.
1. Put students into pairs as A and B.
2. Give each learner a collection of concordance lines for one verb.
3. Ask students to look for the constructions that the verb is used in, e.g., *get + adj, make somebody do something, have someone do something*
4. Ask the students to compare similarities and differences of the found constructions with each verb.
5. Ask for volunteers to share their findings in class.

Teaching Activity 3: Verb Tenses (about 150 minutes, in a computer lab)
1. Ask the students to form groups of 4.
1. Have each learner find at least five movie reviews from the website https://www.rottentomatoes.com and share them with their group partners. (In total 15-20 reviews per group).
2. Have each group create their own mini corpus of reviews.
3. Show students how to upload their reviews into AntConc.
4. Show students how to use the frequency word list function. Ask them to pick out the most frequent verbs and write them down in a list.
5. Have the students write next to each entry the tense of the verb.
6. Show the students how to use the concordance function to see the most common verbs in context.
7. Ask the students to look at concordance lines with the verbs they have written down to see how these verbs are used in context and what reviewers usually use these verbs to accomplish.

8. Give students time to discuss in their groups and write down the common patterns of usage for each verb.
9. For homework, assign students to write a review of their favorite movie. Encourage students to use the verb tenses they focused on and remind them of points they can write about, such as the story (plot) of movie, main characters, how they felt about the movie, effect of movies, sequels to movies, recommendations for movies, and if the movie was adapted from a book.
10. Have students read and provide peer feedback on each other's review.
11. Ask students to revise their reviews based on their peer's comments.
12. Encourage students to post their revised reviews on the Rotten Tomatoes website.

Conclusion

The activities above guide students through awareness-raising tasks and production tasks to practice the genre of movie reviews. Together with other educators (e.g., Crawford & Csomay, 2016; Breyer, 2011; Johns, 1991; Gavioli, 2005, 2009; Lombardo, 2009; Timmis, 2015), we believe that corpora and corpus tools are effective teaching tools for increasing students' awareness of features of the target language and helping students to develop proficiency and confidence as independent learners.

References

Anthony, L. (2011). *AntConc Concordancer.* Retrieved from:
http://www.laurenceanthony.net/software/antconc/
Bloom, K. & Argamon, S. (2010). Automated learning of appraisal extraction patterns. In S. Gries, S. Wulff, & M. Davies (Eds.) *Corpus linguistic applications: Current studies, new directions* (pp. 249-260). Amsterdam: Rodopi.
Breyer, Y. A. (2011). *Corpora in language teaching and learning: Potential, evaluation, challenges*. Frankfurt: Peter Lang.
Connor, U. & Upton, T., A. (2014). *Discourse in the professions: Perspectives from corpus linguistics.* Philadelphia: John Benjamins Publishing Company.
Crawford, W. J. & Csomay, E. (2016). *Doing corpus linguistics*. New York: Routledge.
Johns, T. (1991). Should you be persuaded – two samples of data-driven learning materials. In T. Johns & P. King (Eds.) *ELR Journal 4: Classroom Concordancing* (pp. 1-16). Birmingham: CELS, The University of Birmingham. Available at https://lexically.net/wordsmith/corpus_linguistics_links/Tim%20Johns%20and%20DDL.pdf.
Gavioli, L. (2005). *Exploring corpora for ESP learning*. Amsterdam: John Benjamins.
Gavioli, L. (2009). Corpus analysis and the achievement of learner autonomy in action. In L. Lombardo (Ed.) *Using corpora to learn about language and discourse* (pp. 39-71). Bern: Peter Lang.
Liang, A. (2006). Rotten Tomatoes: Sentiment classification in movie reviews [Report]. Standford University. Retrieved from http://cs229.stanford.edu/proj2006/Liang-SentimentClassification.pdf.
Lombarddo, L. (2009). *Using corpora to learn about language and discourse* (pp. 39-71). Bern: Peter Lang.

Pang, B. & Lee, L. (2004). Seeing stars: Exploiting class relationships for sentiment categorization with respect to rating scales. *Proceedings of ACL*, 115-124. Retrieved from http://www.cs.cornell.edu/home/llee/papers/pang-lee-stars.home.html.

Pang, B., Lee, L. & Vaithyanathan, S. (2002) Thumbs up? Sentiment classification using machine learning techniques. *Proceedings of EMNLP*, 79–86. Retrieved from http://www.cs.cornell.edu/home/llee/papers/sentiment.home.html.

Sorostinean, M., Sana, K, Mohamed M., & Targhi A. (2017). Sentiment analysis on movie reviews [Report]. Retrieved from http://www2.agroparistech.fr/ufr-info/membres/cornuejols/Teaching/Master-AIC/PROJETS-M2-AIC/PROJETS-2016-2017/main(Amal%20Targhi-%20Mihaela%20SOROSTINEAN-%20Katia%20Sana-Mohamed%20Mohamed).pdf

Richards, J. C., Rodgers. T. S. (2001). *Approaches and methods in language teaching*. Cambridge: Cambridge University Press.

Taboada, M. (2011). Stages in an online review genre. *Text & Talk 31*(2), 247–269.

Timmis, I. (2015). *Corpus linguistics for ELT: Research and practice*. London, New York: Routledge.

Warren, M. (2010) *Online corpora for specific purposes*. *ICAME Journal 34*, 169-188. Retrieved from http://clu.uni.no/icame/ij34/warren.pdf

Talai, T. & Fotovatnia, Z. (2012). *Data-driven learning: A student-centered technique for language learning Theory and Practice in Language Studies 2* (7), 1526-1531. Retrieved from https://pdfs.semanticscholar.org/75dc/5428719618d8afc99c70bdb57e2bcc27785d.pdf

About the Authors:

Romina Castagno is an English/Spanish translator and an MA TESOL candidate at Hawaii Pacific University. She is interested in analyzing the patterns of written and spoken language use in natural texts and contexts and how this analysis can enhance language learning and teaching.

Milang Shin is an MA TESOL candidate at Hawaii Pacific University. Her research interests are second language pedagogy, sociolinguistics, discourse analysis, and corpus linguistics.

Yuki Tanaka holds an MA in TESOL from Hawaii Pacific University. Her research and teaching interests are communicative language teaching, corpus linguistics, and the feasibility of an immersion experience in an EFL setting.