LEARN Journal: Language Education and Acquisition Research Network
ISSN: 2630-0672 (Print) | ISSN: 2672-9431 (Online)
Volume: 16, No: 1, January – June 2023

Language Institute, Thammasat University
https://so04.tci-thaijo.org/index.php/LEARN/index

Corpus Linguistics and Cinematic Discourse: Lexical Bundles in Mainstream Film Scripts

Runze Xu[a,*] and Raksangob Wijitsopon[b]

[a]runzexu1@gmail.com, English as an International Language Program, Graduate School, Chulalongkorn University, Thailand
[b]raksangob.w@chula.ac.th, Department of English and Corpus Linguistics for Digital Humanities Research Unit, Faculty of Arts, Chulalongkorn University, Thailand
*Corresponding author, runzexu1@gmail.com

APA Citation:
Xu, R., & Wijitsopon, R. (2023). Corpus linguistics and Cinematic Discourse: Lexical Bundles in Mainstream Film Scripts. *LEARN Journal: Language Education and Acquisition Research Network, 16*(1), 545-574.

Received
16/06/2022

Received in revised form
17/12/2022

Accepted
01/01/2023

ABSTRACT

Hollywood blockbuster films have long attracted not only mass audiences but also scholarly attention. In line with contemporary applied linguistics interests in telecinematic discourse, the present study draws upon concepts and techniques in corpus linguistics to describe the language of American mainstream film scripts. The concept of lexical bundles was employed to identify linguistic patterns characteristic of scripts of American mainstream films produced by entertainment conglomerates, which are popular in the U.S. Results show that American mainstream film scripts are characterized mainly by spoken formulaic expressions. However, descriptive expressions, such as place-referential and action-related lexical bundles, also predominantly make up the given register. Further qualitative analysis reveals that these common multi-word expressions have functional contributions to film scripts in terms of creation of conflicts in plots, characterization, and building engagement with audiences.

Keywords: lexical bundles, film scripts, American movies, telecinematic discourse, corpus linguistics

Introduction

Over decades, Hollywood blockbuster movies, i.e., films that are produced by entertainment conglomerates and very popular in the U.S., play an influential role in the world's entertainment industry. According to the top 200 global box-office ranking (Top Lifetime Grosses, 2022), major Hollywood companies released 192 of the 200 films with the highest box-office gross earnings. Such movies, which are referred to in the present study as mainstream movies, have long attracted not only the public's attention but also academic interests in cultural studies. In recent years, language use in movies and TV series has gained an increasing interest from linguists. Piazza et al. (2011) coined the term telecinematic discourse to refer to research that "works with a linguistic study of films and television shows" (cited in Zago, 2021: 168). The reasons why telecinematic discourse has captured more and more attention in linguistics are as follows. First, it has a distinctive discourse feature: conversations are written and then delivered verbally and through acting (Zago, 2021). Studying telecinematic discourse can therefore shed light on aspects of relationship among speech, writing, non-verbal language and genre. Next, telecinematic dialogues, which are often seen as close to real-life conversations but address mass audience, display discourse complexity. While characters are talking to one another, they are also communicating something, which may not be said, with their audiences. As Bednarek (2018) illustrates, if the protagonist in a film gets an invitation to a party, it could mean more than an invitation to the audience as it may invoke a dramatic irony or strengthen the relationship between the protagonist and another character, which in turn advances the plot in some ways. Thanks to such recognized features of telecinematic discourse, a number of applied linguistic studies have been conducted on particular genres of mainstream films, often qualitatively, such as romantic comedy (Nuryani, 2016) young adult adventure films (Nursanti, 2015) and gangster dramas (Statham, 2015). The present study, therefore, seeks to investigate language use in film scripts of mainstream movies, not limited to any particular genre, in order to further an insight into the language of cinematic discourse.

To fulfill this objective, a corpus linguistic approach is adopted in this study, drawing upon the concept of lexical bundles, i.e. frequent repeated sequences of words in a corpus extracted by computer software (Biber et al., 1999). By focusing on lexical bundles, the study is intended for another aim:

to expand the scope of lexical bundle-based studies in corpus linguistics. While lexical bundles have widely been examined in corpus linguistic research, most of the studies look at their occurrences in established genres of written texts, such as academic papers (e.g. Shirazizadeh & Amirfazlian, 2021), business emails (Siricharoen & Wijitsopon, 2020) or learner English writing (Chen & Baker, 2010). Little has been done on lexical bundles in fictional texts, including telecinematic discourse, with the exceptions of Bednarek (2012), which looks at lexical bundles in fictional television series, and Freddi (2011), which examines lexical bundles in film dialogues alone. No studies have yet looked at lexical bundles in movie scripts in general.

To these ends, we first give an account of the relationship between American mainstream film scripts and lexical bundles in order to contextualize our study and research questions. Then, we describe methodological aspects of the study before moving on to the analysis and discussion of findings. Finally, a summary of findings is given together with suggestions for further studies.

Literature Review

American mainstream movies, film scripts and lexical bundles

In film studies, the concept of American *mainstream films* is often discussed in comparison with the American *independent films* (also called *arthouse films*) (Meyer, Song, & Ha, 2016). American mainstream films are usually produced and distributed by Hollywood majors, aimed at the mass market (Zuckerman & Kim, 2003) and considered "more polished, expensive and conservative films" (Tzioumakis, 2017, p. 2). On the other hand, American independent films aim at niche markets (Zuckerman & Kim, 2003) and the directors have "creative dominance over his/her work" (Sarris, 1963). Corresponding to the idea of "market identities" suggested by Zuckerman & Kim (2003), Gemser et al. (2007, p. 63) suggest that the place of film release is another criterion for defining American mainstream and independent films: "if the film is released in film theatres that show above all mainstream films, it is considered mainstream; if the film is released in film theatres that predominantly show art house films, the film is classified as art house". It should be noted that in the age of streaming services, enhanced by lockdown across the world during the Covid-19 outbreak, the above criteria still work in the context of American films. This is because, first, most American films are still released in cinemas, mainly on a streaming service. For example, Netflix's films like *Roma* (Cuarón, 2018), *The Irishmen* (Scorsese, 2019), and *Marriage Story* (Baumbach, 2019) all had a robust theatrical release in the U.S. (Brueggemann & Brueggemann, 2019; Vivarelli & Vivarelli, 2019). Also, the

film industry has gradually recovered from the market downturn caused by the Covid-19 pandemic. The widespread vaccination led audiences to return to various forms of entertainment ("'A Quiet Place Part II' Makes Serious Memorial Day Noise With A $48.4 Million Three-Day Bow; 'Cruella' Is Solid In Second With $21.3 Million," 2021). Hence, film release outlets can still be regarded as a major criterion for distinguishing mainstream and independent films these days.

Apart from consideration of producers, the market and physical outlets, Newman (2011) suggests that mainstream American movies can also be explained in terms of aesthetics and style. For example, a crucial narrative convention of American mainstream movies is goal-oriented storytelling (Thompson, 2003), i.e. movie plots and incidents in the stories are geared towards solving a problem, e.g., defeating the villain, rescuing a person, destroying the bomb, etc. Moreover, it is observed that style of mainstream film dialogues tends to be formulaic but unrealistic. For example, they tend to show a wide range of knowledge of a character that is not related to the character's profession, be revealing rather than withholding information, especially a character's heart (King, 2005, p. 64).

Film scripts

The abovementioned stylistic features of mainstream movies can be approached from linguistic perspectives, with film scripts as the source of language data. A film script is written for film shooting. It contains information about a film, which normally includes dialogues, monologues, soundtrack, voiceover, etc. It is also important for budget management, casting, and further adjustment on the story (Gregory, 1967). Almost everything that happens on screen starts from the film script. A close examination of language use in film scripts can therefore reveal several important aspects of telecinematic discourse, both as a communicative instance and as a social practice that both reflects and contributes to institutional, economic and cultural factors that surround the film industry.

The film script has its own distinctive structure, containing the following sections:

Scene Heading

A scene heading is presented in full capitalization. It tells readers the location where the scene takes place. The line must distinguish the scene as inside (INT.) or outside (EXT.), which refers to an indoor or outdoor scene. A hyphen following the location, which indicates the time of the day

(morning, evening, afternoon, etc.). Scene headings are usually presented in the following way:

2. INT. EVE'S CAR - DAY

(*Skyfall,* 2012)

In this example, the number '2' signals the scene number in the script, referring to the second scene of the film *Skyfall*, which takes place in Eve's car during the daytime.

Action Lines

An action line is normally right below the scene heading. It provides information about a character's physical actions. For example:

The warrior beats up a thousand ninjas, delivering his final blow while doing a split between two trees.

(*Kungfu Panda*, 2008)

Dialogue

A dialogue consists of utterances of characters. In a dialogue, a character's name is capitalized and its utterances are placed beneath the character's name. For example:

RYAN
Installation ninety-five percent complete. Running level one diagnostics on circuits, sensors, and power. Standby.

(*Gravity*, 2013)

Extensions

An extension is normally placed next to the character's name in parentheses. It tells readers the manner of the spoken dialogues presented, such as:

Voice over (V.O.): A narrator talks outside the scene, like she/he is directly communicating with the audience. It can also be an internal thought.

Off Screen (O.S.): A dialogue can be heard by other characters but can't be seen by audiences.

Transitions

Transitions tell the film editor how to edit between two scenes, such as:

CUT TO: the normal transition usually indicated in the end of the scene.

DISSOLVE TO: a scene dissolves to the other, usually indicating the time passing.

INTERCUT: the editor jumps back and forth between two scenes, usually used in phone conversation.

Linguistic approaches to film scripts

With the film scripts as part of cinematic discourse, several linguistic studies apply concepts in pragmatics and sociolinguistics to examine language use in film scripts. For example, Bousfield and McIntyre (2018) investigated the dialogues of *Full Metal Jacket* (1987) with the prototype-based taxonomy of impoliteness and rudeness (Bousfield, 2010). Lee (2015) investigated a non-human character Japanese anime series *From the New World* (2013), using Bucholtz and Hall's (2005) identity framework.

Crucially, a corpus-assisted approach has been increasingly adopted in the fast-growing research on telecinematic discourse (e.g., Bednarek, 2011; Jautz & Minow, 2020; Pavesi, 2020; Reichelt, 2020), most of which focus on keywords, collocations and lexical patterns observed from concordance lines. Pavesi (2020), for instance, studied demonstratives in English and American films and found that demonstrative pronouns are common in cinematic discourse, with proximal demonstratives being more frequent in film dialogues than in other spoken registers of English. It is argued that this is because the demonstrative "this" has a narrative function that can intensify the attention of the audience toward characters, objects and events. Moreover, in an investigation into one of the main characters of the series *Gilmore Girls* across 8 seasons, Bednarek (2011) extracted keywords and lexical bundles in the character's dialogue and found that there is no apparent style shift in the character's dialogue throughout the whole series. Finally, McIntyre (2012) investigated male and female dialogues in a corpus of mainstream American films by analyzing keywords, key semantic domains, and key n-grams. The study revealed that major movies tend to share a number of prototypical stylistic features which exhibit gender stereotypes. For example, the semantic domain "in power" is the most frequent one in male speech but uncommon in female speech. On the other hand, key n-grams like "get me out of here" and "help me" are common in female characters' utterances. These patterns serve to construct female identities as needing help and males as saviors and ones who fight for power. This points to the observation that gender stereotypes and the notion of ideal masculinity (e.g., men should fight for power, heroes against authority) tend to underlie Hollywood blockbusters.

Although there seems to be quite a number of recent studies on mainstream film scripts, most of them focus on the dialogue part, as illustrated above. Moreover, several previous studies tend to highlight specific aspects of language use in a particular movie or series, e.g. demonstratives and gender-related language use, rather than general linguistic patterns that occur in full texts of film scripts across movies. The present study, therefore, aims to fill this gap by describing forms and functions of linguistic patterns in mainstream film scripts as a text variety in cinematic discourse, in addition to the macro structural and organizational patterns of the film script, illustrated in Figure 1. The concept that we draw upon to connect linguistic patterns with mainstream film scripts as a text type is lexical bundles.

Lexical bundles

Lexical bundles are common uninterrupted sequences of words that repeatedly occur across several texts in certain text varieties (Biber et al., 1999). Lexical bundles are automatically extracted from a corpus via computer software, based on the criteria on length, minimum frequency and distribution set up by a researcher. As a result, they may be structurally incomplete chunks of everyday utterances and do not necessarily convey idiomatic meanings, e.g., *I don't know what*.

The concept of lexical bundles can be applied to investigation into various types of texts since they are essential elements of discourse that fulfill communicative purposes (Conrad & Biber, (2005), e.g. introducing discourse topics and expressing one's wishes, more details of which are spelled out in Section 3.4 as our analytical framework. Because lexical bundles are linked to creation, organization and communicative functions of texts, they can be used to approach text varieties. Biber *et al.* (1999) is a pioneering work on lexical bundles and differentiation of text types. They investigated different text varieties in the English language, ranging from academic writing, general prose, fiction, newspapers and conversation. Their analysis revealed that conversation and academic prose differ greatly in terms of types of lexical bundles. Biber and Barbieri (2007) demonstrated that even texts in the same setting can be differentiated through lexical bundles. They examined the use of lexical bundles in various spoken and written university registers and found that lexical bundles are more common in non-academic than in core instructional registers. Moreover, their findings demonstrated that while lexical bundles are often found to be more prevalent in spoken than written texts in previous research, they are particularly common in written course management (e.g., course syllabi) in the university setting. Furthermore, Hyland (2008) has shown that writing in different disciplines, namely

electrical engineering, biology, business studies, and applied linguistics, can also be distinguished through categories of lexical bundles.

These sample major works on lexical bundles illustrate concentration on academic registers in studies of lexical bundles. As Crosthwaite et.al. (2022) have found, lexical bundles have featured in corpus-based studies over the past two decades, especially those on academic writing. Little has been done on lexical bundles in expressive texts, such as novels and screenplays. This study therefore also seeks to extend the scope of lexical bundle-based studies by approaching language use in mainstream film scripts.

Based on these fundamentals and previous studies, the present study seeks to answer the following research questions:

1. What lexical bundles constitute American mainstream film scripts?
2. What meanings and/or functions do the retrieved lexical bundles have in mainstream film scripts?

Methodology

Four methodological aspects of the study are described in this section: corpus compilation, corpus software, lexical bundle extraction, and an analytical framework.

Corpus compilation

American film scripts have wide public accessibility on a film's official website or filmscript websites. In this study, we compiled a corpus of mainstream box-office film scripts (henceforth MFC) by downloading scripts from the website *scriptslug.com*, since it provides scripts in accordance with the United States' fair use policy, i.e. the posted scripts are available for research and educational purposes. The scripts that were formatted from scanned papers were manually checked for spelling and punctuation to ensure that there are no garbled codes caused by system error. MFC contains 100 film scripts, with 2,447,296 tokens. Two criteria were observed for the compilation of MFC:

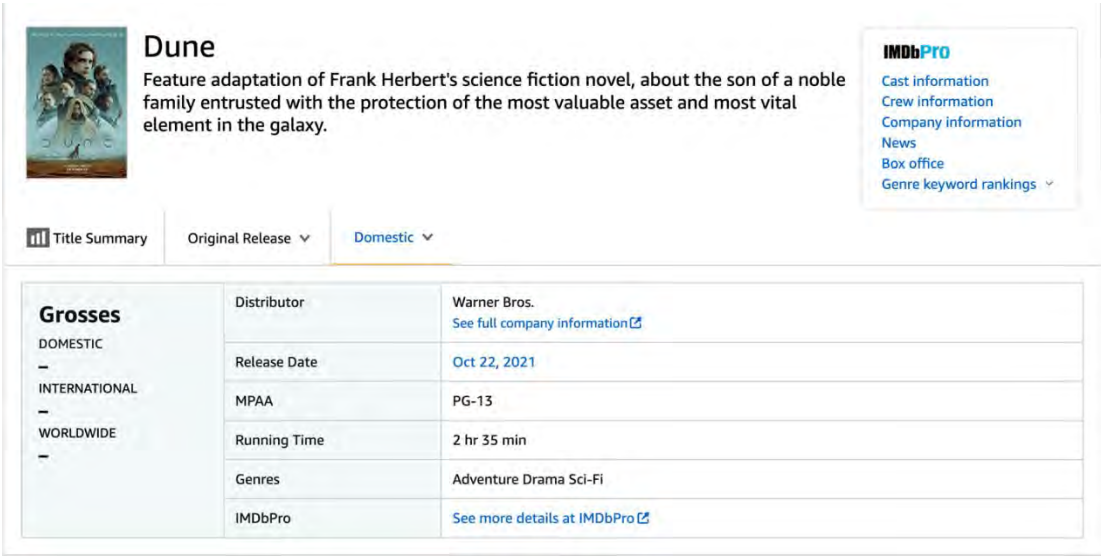
1. The scripts are of the movies ranked among top 5 in the U.S. domestic box-office annual chart during 2005 and 2019.
2. If the scripts of the top 5 films were unavailable online, the next place was selected. For example, in 2014, the film in 5th place was *Transformer: Age of Extinction*, but its script cannot be found online so the script of the 6th-place film *Maleficent* was used instead.

It must be noted that by prioritizing movie ranks in compilation of MFC, balance of movie genres in the corpus cannot be fully achieved. In other words, movie genres are not evenly distributed in MFC. This can affect

types of lexical bundles that turn up in the analysis. However, since this study is interested in linguistic patterns in mainstream movies, use of ranking as the major criterion in movie selection is inevitable. Moreover, movie genre is not a clear-cut category. Most of the films can be described as belonging to two or more genres, as indicated by the metadata 'genre' on boxofficemojo.com. As illustrated by Figure 1 below, the film *Dune* belongs to three movie genres, i.e. adventure, drama, and sci-fi. To minimize limitations that could possibly arise from a lack of evenly distribution of movie genres as suggested above, a set of criteria for extraction of lexical bundles in MFC were set up and interpretations of findings are made with caution.

Figure 1

Metadata of the film Dune on boxofficemojo.com



Lexical bundle extraction

To avoid bringing up lexical bundles likely to be more specific to particular films or genres, the following criteria were set up to extract lexical bundles in this study. First, lexical bundles in the present study must occur in at least 40 out of 100 different film scripts in the corpus. Next, lexical bundles must have a minimum frequency of 100 tokens. These quantitative criteria enabled us to retrieve lexical bundles common across different films scripts in the corpus. In terms of length of lexical bundles, a length of four words was opted for. This is because, as Biber et al. (1999, p. 992) point out, three-

word bundles are too common and five-word or longer ones are "more phrasal" and less common in a corpus. This corresponds to Stubbs and Barth (2003), who observe that longer bundles tend to be limited to specific texts. A length of four words, therefore, gives an optimal size for lexical bundles in this study.

Software

This research employs the software *Antconc* (Anthony, 2020) to extract lexical bundles in MFC. The corpus was loaded to *Antconc* and lexical bundles were extracted via the n-gram/cluster function. It must be noted that via *Antconc*, the apostrophes will not be identified in the result, and hence such expressions with apostrophes as *I'm* or *don't* were identified as two-word lexical bundles like *I am* and *do not*, respectively. This means that the bundle *I don't think* is counted as a four-word bundle (amounting to *I do not think*) like the cluster *I think that you*. We decided to follow *Antconc*'s system of automatic identification of lexical bundles to avoid the impact of spelling differences between straight and curly apostrophes in different film scripts (e.g. *don't* and *don't*). Hence, our list of lexical bundles contains those with clearly distinct four words, e.g. *I think that you*, and those with three words in an extracted form with the apostrophe, e.g. *I don't think*, which are identified by *Antconc* as four-word bundles.

Analytical Framework and Procedures

After a list of lexical bundles was generated, following the criteria spelled out in 3.2, concordance lines for each individual lexical bundle were generated to enable us to examine the use of each bundle in its context. The bundles were then grouped according to their functions, following Biber et al.'s (2004) functional taxonomy of lexical bundles, explained below. It must be noted that some lexical bundles can have more than one function. However, in order to calculate and compare lexical bundles systematically, we assigned each of the lexical bundles to only one category and sub-category, using frequency of its use as a criterion. For instance, while the bundle *I'd like to* may be put in the "Desire" or "Topic introduction" category, it was labelled as a "Desire" bundle only because the phrase is consistently used to express desire in MFC but it is not always put at the beginning of a conversation to introduce a topic. Applying combined quantitative and qualitative perspectives in our functional analysis can arguably make the results more valid and reliable.

Below is Biber et al.'s (2004) functional taxonomy, which we applied as our analytical framework for a qualitative analysis of lexical bundles in MFC.

Stance lexical bundles are those that point at attitudes or degree of certainty. There are five sub-groups of stance expressions.

- *Epistemic stance bundles* present the text-producer's (un)certainty, e.g. *I'm not sure, I don't think so, you know what I.*
- *Desire* presents the wishes and desires of the text-producer, or asks information about the other person's desire, e.g. *I just wanted to, if you want to.*
- *Obligation/directive* bundles present obligations or direct the order that the speaker wants the interlocuter to accomplish, usually ordered with second-person pronoun e.g., *you don't have to, you have to be.*
- *Intention/prediction* contains lexical bundles that describe the writer's intention to do certain future action, e.g. *I'm going to, I was going to.*
- *Ability/effort* bundles indicate ability, e.g., *to be able to, it is possible to*

Discourse-organizing lexical bundles are those that indicate a logical relation between the propositions. This category consists of three sub-groups:

- *Topic introduction/focus* serves the function of introducing a new topic. e.g., *I would like to, I'm sorry to.*
- *Topic elaboration/clarification* refers to the bundles which elaborate or clarify the topic, e.g., *know what I mean, nothing to do with.*
- *Conditions* bundles normally contains a complementizer 'if', e.g*., if you have time, if you do not*

Referential lexical bundles are those that indicate the reference to the physical/abstract units or the textual context itself.

- *Identification/focus* bundles emphasize a single important feature or identify a noteworthy part of something, e.g., *it's not a*
- *Quantity specification* e.g., *the rest of the*
- *Place* reference e.g., *in front of the, the edge of the*
- *Time* reference refers to lexical bundles that describe temporal attributes or time periods, e.g., *for the first time*

Special-function lexical bundles are those that do not go with the previous three categories. This group of lexical bundles was added in Biber et al. (2004) and Biber and Barbieri (2007). In the present study, we have also found some lexical bundles that do not belong to any of the above three categories and hence put them in this group, further dividing them into the following sub-categories:

- *Apology*: lexical bundles that express an apology, containing the word *sorry,* e.g., *I'm sorry I, I'm so sorry*
- *Inquiry*: lexical bundles that are used to ask questions, e.g., *what are you doing*, *how are you doing*
- *Command*: lexical bundles that are used to order someone to do something, e.g. *get out of here, get out of the*
- *Action*: lexical bundles that describe actions, e.g., *looks at each other*, *takes a deep breath*

Results and Discussion

Based on the above extraction criteria, a total of 77 lexical bundle types, altogether making up 8,479 lexical bundle tokens, were derived. They can be categorized functionally according to the above analytical framework, as presented in Table 1 below.

Table 1

Lexical Bundles in MFC

Functional category	Lexical bundle (Frequency)	Number of types	Percentage of types	Number of tokens	Percentage of tokens
STANCE EXPRESSIONS	**A. Epistemic stance** *I don't know* (689) *, don't know what (184), I don't think (175) , don't know I (74), you don't know(87), I didn't know (80), I can't believe (82), don't know how (75), but I don't (71), I think it's (63), and I don't (64), I'm not sure (59)*	12		1,703	
	B1. Desire *I don't want* (201)*, don't want to* (191)*, I'd like to (100), do you want to (66), you don't want (60), I don't need (60), I don't care (75)*	7		762	

	B2. Obligation *you don't have (98), I want you to (89), don't have to (86), you want me to (83), I need you to (70)*	5	41.56% (32)	426	44.47% (3,771)
	B3. Prediction/intention *I'm going to* (233), *you're going to (168), we're going to (114), I'm not going (70)*	4		585	
	B4. Ability/effort *I don't have* (107) *can't help but (69), I cant do(63), we don't have (56)*	4		295	
DISCOURSE ORGANIZERS	**A. Topic introduction/focus** *that's what I (82), you know what I (71), that's why I (55)*	3	5.19% (4)	137	2.71% (230)
	B. Conditions *if you don't (93)*	1		93	
REFERENTIAL BUNDLES	**A. Identification/focus** *it's not a (56)*	1	31.17% (24)	56	34.14% (2,895)
	B. Quantity Specification *the rest of the (169)*	1		169	
	C1. Place reference *in front of the* (196), *the edge of the (177), in the middle of (130), the end of the (165),the top of (142), the side of the (144), the middle of the(129), on the other side(119), out of the way (123), the other side of (116), at the end of (108),the center of the (101), in front of him (100), the back of the* (201), *other side of the(96), in front of a (87), at the top of (88), out of the car (82), out of the room (77), the front of the (70), in the back of(63)*	21		2,514	
	C2. Time reference *for the first time (156)*	1		156	
SPECIAL-FUNCTION BUNDLES	**A. Apology** *I'm sorry I (131), i'm so sorry (87)*	2		218	
	B. Inquiry *what are you doing* (221) *, what's going on (130), why don't you (87), are you talking about (74), what are you talking (63), what do you think (85), how do you know (61), what do you mean (79)*	8		800	
	C. Command *get out of here (62)*	2		123	

	get out of the (61)		22.28% (17)		18.67% (1,583)
	D. Actions *look at each other (140), takes a deep breath (90), turns back to the(63), looks up at the (93), looks down at the (56)*	5		442	
Total		**77**	**100%**	**8,479**	**100%**

As shown in the table, Stance bundles are the most dominant group in MFC, having the most lexical bundle types and the highest frequency of tokens, followed by Referential, Special-function and Discourse-organizing bundles, respectively. Overall, these proportions reflect the nature of film scripts as a hybrid discourse, heavily spoken plus descriptive. Specifically, mainstream film scripts are predominantly dialogic, as manifested through a large density of spoken formulaic expressions in the Stance, Special-function and Discourse-organising categories. This points to the dialogue-oriented storytelling technique employed for textual development. Nevertheless, the film script is also largely contributed by descriptive narration, as mainly illustrated by the second most frequent Referential lexical bundles in MFC. These overall characteristics correspond to Zago's (2021) observation that a distinctive discourse feature of telecinematic discourse involves conversations, which are delivered verbally and through action. Major patterns and functions of each lexical bundle category are discussed in more detail below.

Stance lexical bundles

As mentioned above, stance bundles make up the largest functional category in mainstream film scripts. This highlights the prevalence of spoken discourse features in film scripts as suggested by the predominance of lexical bundles with the first-person pronoun "I" and such contracted forms as *don't*, *I'm* and *you're*. These bundles express characters' thoughts, opinions, attitudes and evaluations, which in turn serve to initiate and advance the story. While this does not seem very surprising (given the type of texts in MFC), the corpus-based approach to film scripts here brings to attention a structural pattern shared by stance bundles across sub-categories, which has not been noted elsewhere, to the best of our knowledge: stance lexical bundles in film scripts are mainly in negative forms, with 24 out of 35 bundle types (68.57%). The predominance of negated stance bundles can be attributed to their potential for creation of conflicts in movie plots, both external and internal, e.g. characters' inabilities, disagreement or unwillingness, which constitutes the problem-solving convention in storytelling, corresponding to an

observation in film studies, e.g. Thompson (2003), noted earlier in Section 2. This is illustrated by examples below:

(1) ANDY
I need your help. ***I don't know*** what to
do. It's like I'm completely beneath her contempt.
(2006 *The Devil Wears Prada*)

(2) YOUNG CHARITY
I don't know what my future will be. Father says I
am to be a lady, but... It's all so dull when you're not
there.
(2017 *The Greatest Showman*)

The two extracts above illustrate how the negative stance bundle *I don't know,* the most frequent one, plays a role in expressing ongoing conflicts in the stories. In extract 1), the character Andy is trying to gain approval from her boss and the lexical bundle helps to express her problem that keeps the story going. In extract (2), the lexical bundle betrays the character's inner conflict, i.e. her uncertainty and anxiety.

(3) ALAN
We don't have time for this! We gotta find Doug!
(2009 *The Hangover*)

The bundle *we don't have* is a highly functional bundle for screenwriters, especially in terms of accelerating the film's plot, or creating a suspenseful atmosphere. As exemplified by example (3), the character, Alan, uses *we don't have time* to push his/her partner to take an action. It also excites the audiences for an upcoming rush moment.

Referential lexical bundles

As mentioned earlier, referential bundles constitute the second largest functional category in MFC. Its large proportion, however, comes mainly from lexical bundles that designate places (see Table 1), with 21 out of 24 referential bundle types (87.5%). While place and time are essential elements in projecting narrative text world (cf. Gavins, 2007), it is interesting that in mainstream film scripts references to place far outnumber temporal expressions. Upon examination, it is found that the large density of place-referential bundles is connected with characters' actions and movements, often described in action lines, where information about characters' physical actions is given (see Section 2). In other words, place references serve to

construct spatial context of characters' actions, which in turn becomes contextual information for storyboard design and film shooting. The fact that place referential bundles largely constitute mainstream film scripts reflects another aspect of the nature of filmscripts, i.e. it features descriptions of settings and actions, apart from the dialogue, to fulfill communicative functions of the film script by connecting places with events or actions performed by characters. It can therefore be said that spatial bundles are highly essential to the film script in meaning-making through a connection among the setting, actions and dialogues in each scene. Below are examples of place-referential lexical bundles that illustrate this function in mainstream filmscripts.

> Mary sits in a chair eating dinner from a tray.
> Miranda sits on ***the edge of the*** bed near her.
> MIRANDA
> I don't know if Steve's mentioned anything to you yet -- about us.
> Has he?
> Mary shakes her head: No -- concerned. Miranda nods.
>
> (2008 *Sex and the City*)

In the example above, the lexical bundle *the edge of the* points to the place – a private space – where Miranda is sitting before she initiates a private (and crucial) conversation with Mary concerning her separation from Mary's son (Steve).

It must be noted that while place referential bundles are associated with characters' physical actions, the only temporal lexical bundle in MFC *for the first time* is connected with characters' perceptions and emotions. Of its 156 entries, 36 (23.07 %) cases of the bundle are used in characters' dialogues and 120 (76.93%) in action lines, suggesting that, like place referential bundles, it is more characteristic of the description part than the dialogue. Analysis of its concordance lines reveals that the bundle often co-occurs with words associated with characters' vision and visual acts, such as *see, seeing, noticing, gaze* and *eyes*. Moreover, these vision-related words often co-occur with those concerning feelings and emotions, including *sympathetic gaze* and *painful looking* (See Figure 2 below). Based on these patterns, it can be observed that a primary function of the lexical bundle *for the first time* is to highlight critical emotional moments in the stories, which are reflected through characters' visual perception. Given that readers of filmscripts are primarily people involved in film production, namely directors, actors,

cinematographers, the lexical bundle also serves as a linguistic clue for them to interpret, direct and act.

Figure 2

Sample Concordance Lines of 'for the first time' in MFC

Her eyes return to Caesar -- who	***for the first time***	softens, offering a sympathetic gaze.
her little face inches from his...	***for the first time***	we see deep emotion in her
to the Colonel, and noticing	***for the first time***	painful looking BRAND SCARS on
here anyway. Mike's eye narrows.	***For the first time***	he looks determined. The light turns
on a White House tour.	***For the first time***	he looks irritated at work. INT.
that Ludendorff loosens his grip.	***for the first time***	we see that he is afraid.
alright? Ray looks up, and	***for the first time***	we see sadness. Tears trickle down
Goodbye, Bella. And here,	***for the first time***	we SEE how truly agonizing this
looks away and Dumbledore,	***for the first time***	notices his hands. They are splayed,
the tube. Ted sits on the sofa, and	***for the first time***	we see Ted in his present

In dialogues, the bundle also serves to constitute a character's emotional expression in a highlighted moment, with a similar co-occurrence pattern, as illustrated by underlined parts in the extracts below.

(1)
LORRIE
For the first time, just now, I
realized that there were 155 people on that plane.
And you were one of them.
More tears. Tears of relief and joy.

(2016 *Sully*)

(2)
EGO (CONT'D)
It was with Meredith that I
experienced love **for the first time**. I called her my
river lily. And from that love, Peter, you.

(2017 *The Guardians of Galaxy vol.2*)

In relation to emotional moments, *for the first time* is also used to describe milestones in a character's life, whether claimed by characters or narrated by screenwriters. This is suggested through co-occurrences

between *for the first time* and prepositional phrases referring to a time period, e.g. *in my life*, *in our history* and *in years*, as illustrated by Figure 3.

Figure 3

Sample Concordance Lines of 'for the first time' in MFC

theft. They're suing me because	***for the first time***	in their lives, things didn't
So, what now? MIKE You know,	***for the first time***	in my life, I don't
Gloria's eyes are clear and	***for the first time***	in a long time she looks
man in the room is surprised	***for the first time***	in a very long time by
, all weakness gone. She smiles	***for the first time***	since Stefan took her wings. But
beat, husband and wife reunited	***for the first time***	in years. In the smoke and
and reading a newspaper	***for the first time***	in 12 years, plus I just met
staring at the numbers. Stuck,	***for the first time***	in her life. Stafford stops at
her voice to meet Hippolyta's	***for the first time***	in three thousand years.
of the Nation of Wakanda and	***for the first time***	in our history we will be

Special-function lexical bundles

The significance of combination between dialogue and action in mainstream film scripts is reflected not only through the density of stance and place-referential lexical bundles but also through the Special-function lexical bundles. Except for the *Inquiry* subset, three sub-categories were newly formulated to accommodate several related lexical bundles that emerge in MFC as they were not listed in previous studies of lexical bundles, e.g. Biber et al. (2004) and Siricharoen and Wijitsopon (2020). These three new groups of special-function lexical bundles are: *Apology, Command* and *Actions*. While the *Apology, Inquiry* and *Command* sub-categories articulate important speech acts in MFC, the *Actions* sub-catgory illustrates actions common in mainstream filmscripts, e.g. *looks at each other* and *turns back to the* (see Table 1). Again, this points to the hybrid nature of American mainstream film scripts, featuring a wide range of spoken discourses in combination with description of actions. This characteristic in form of lexical bundles is linked to communicative functions of the movie script that feature written conversations that are delivered verbally and through acting (Zago, 2021).

Focusing on the *Actions* lexical bundles, the act of looking is the most common in MFC, with 289 out of 442 tokens (65.38%) of *Actions* bundles containing the verb "look". The bundle *look at each other*, the most frequent one in this group (see Table 1), is often followed by other actions, especially

those that register feelings, e.g. *laugh, shrug, smile* and *hug*, as illustrated by Figure 4 below:

Figure 4

Sample concordance lines of 'look at each other' in MFC

ENGAGEMENT RING... They	***look at each other***	ALLY (laughing) Are you kidding
starts up his bike. Doug and Bob	***look at each other***	and shrug. Dudley hurries out of
the boxers to shore. The guys all	***look at each other***	and smile. It's a nice
me, understand?! Josh and Kitty	***look at each other***	Little awkward. GABE (Josh and
right? Eddie, Richie and Stanley	***look at each other***	The mood has changed. Back to
earthquake. Ajax and Deadpool	***look at each other***	Shrug. DEADPOOL (CONT'D)
What was that? They turn back,	***look at each other***	Ginny smiles. Then: NEW
New Year! Carrie and Miranda	***look at each other***	They hug. 168L INT. LOWER
Carrie, Samantha and Miranda	***look at each other***	Carrie BURSTS out laughing.
MIRANDA appears. They	***look at each other***	MIRANDA nods, almost

This pattern points to the significance of the act of looking in mainstream films as a common preliminary action, through which two characters communicate quietly before more explicit expression of feelings or attitudes is revealed. Directors and actors have to rely on this description for film shooting so that acting can lead the audience to appropriate interpretations.

Discourse-organizing lexical bundles

According to Table 1, discourse-organizing bundles are the least frequent and varied functional category in MFC. In spite of this, some pattern can be observed from this category; they are dialogic expressions, all containing either first- or second-personal pronouns, thereby contributing to the dialogic nature of film scripts. In the *Topic introduction/focus* subcategories, the three bundles (*that's what I, you know what I, that's why I*) contain the personal pronoun *I* and are said by characters to give explanations about themselves. In fact, many of these bundles are part of longer conventionalized expressions, such as *you know what I mean/think, that's what I'm talking about,* etc. The prevalence of such formulaic expressions creates a style close to real-life conversation in the film script, also noted by scholars in film studies, such as King (2005). Moreover, discourse-organizing bundles can also provide

contextual information about characters for audiences of the movies. For example:

> BELLA
> You all know what Edward wants. And
> ***you know what I*** want. But I won't
> force myself on you...
>
> (*2009 The Twilight Saga: New Moon*)

In this excerpt, Bella emphasizes knowledge commonly shared between her and the Cullens: she desires to be a vampire like them. The bundle *you know what I want* can prompt the audience to infer what that common knowledge the characters have in order to understand meaning between the lines about the conflicts between both parties, i.e. that although both understand each other, they cannot fulfill the other's wants; Bella wants to be a vampire but Edward does not want her to be.

Conclusion

Adopting a corpus linguistic approach, the present study can make both quantitative and qualitative contributions to the study of American mainstream film scripts. While the given text type has been described at the macro level in terms of its structural components, the findings from this study have demonstrated that American mainstream film scripts are mainly characterized through common spoken lexical bundles. This may account for previous observations of mainstream films as being formulaic (e.g. King, 2005) and also for their popularity since they contain a large number of dialogic expressions relatable to real-life spoken language, which, however, are used to create a fictional text world by highlighting conflicts in stories, expressing characters' inner feelings and advancing the plots. At the same time, American mainstream film scripts are also essentially made up of formulaic expressions that describe actions, locations and movements. These groups of lexical bundles help enrich the texts with visualized contextual information and contribute to cohesive meanings from spoken words and body language in the movie script. Our study has also thrown light upon a group of lexical bundles and their lexical patterns, including time-/ place-referential and action lexical bundles, which have not been studied in detail elsewhere. This illustrates that lexical bundles in American mainstream film scripts are also accompanied by co-occurrence patterns that emerge in different movie scripts, which in turn points to their local textual functions (cf. Mahlberg, 2005) specific to the cinematic discourse.

It should be noted that while the derived lexical bundles came only from dialogues and action lines but not from other parts of the film scripts, this does not mean that expressions in those other sections are not important to mainstream film scripts. The findings presented here are lexical bundles that occur commonly enough to go through the criteria set up in the study. This in turn means that, from a quantitative perspective, dialogues and action lines are major components of the film script, since they contain various distinctive lexical sequences characteristic of the register. Other parts of the film script may contain lexical sequences too but their frequencies do not pass the criteria set up in this study. Qualitatively speaking, if we consider film scripts as discourse, looking at communicative acts and individual participants in the communication, it can be said that the common lexical bundles that emerge across different film scripts here are major bundles used to inform and facilitate film production teams, including actors, directors and administrators, allowing them to interpret cinematic texts and prepare for the production.

While the study sheds light on general linguistic patterns shared across American mainstream film scripts, several points are worth further investigation for more thorough understanding of the register. First, it would be useful to pursue a contrastive study on spoken lexical bundles in mainstream film scripts and those in real conversations. Although the present study has shown that everyday stance bundles predominate in mainstream film scripts, which are fictional, it would be interesting to examine further to what extent they are used differently, both quantitatively and qualitatively, from those in real-life conversations. Also, while it has been shown that descriptive bundles in narration, especially place-referential and action lexical bundles, have quantitative significance in mainstream film scripts, we have only slightly touched on their qualitative aspects due to space limits. It would be interesting to further investigate their patterns and functions as well as their relationship with spoken lexical bundles in film scripts as part of the multimodal cinematic discourse.

Although there are still points that need to be explored to fully describe the discourse of mainstream film scripts, this study is hoped to illustrate the value of a multidisciplinary approach, which brings together corpus linguistics, discourse studies and film studies, to the film script, a text variety central to the film industry.

About the Authors

Runze Xu: A MA student in the English as an International Language (EIL) Program, Graduate School, Chulalongkorn University. His research interests are telecinematic stylistics and corpus linguistics.

Raksangob Wijitsopon: An Associate Professor in the Department of English and Head of Corpus Linguistics for Digital Humanities Research Unit, Faculty of Arts, Chulalongkorn University. Her research interests are corpus linguistics, stylistics and discourse analysis.

References

'A Quiet Place Part II' Makes Serious Memorial Day Noise With A $48.4 Million Three-Day Bow; 'Cruella' Is Solid In Second With $21.3 Million. (2021, May 30). Retrieved June 2, 2021, from Box Office Mojo website: https://www.boxofficemojo.com/article/ed1450181636/?ref_=bo_hm_hp

Anthony, L. (2020). *AntConc* [Mac OS]. Tokyo, Japan: Waseda University. Retrieved from Available from http://www.antlab.sci.waseda.ac.jp/

Baumbach, N. (Director). (2019). *Marriage Story* [Comedy, Drama, Romance]. Heyday Films, Netflix.

Bednarek, M. (2011). The stability of the televisual character. The stability of the televisual character: A corpus stylistic case study. In R. Piazza, M. Bednarek & F. Rossi (Eds), Telecinematic discourse: Approaches to the language of films and television series (pp. 185-204). John Benjamins:

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman Grammar of Spoken and Written English*. Longman.

Bousfield, D. (2010). Researching impoliteness and rudeness: Issues and definitions. In M. A. Locher and S. L. Graham (Eds.), *Interpersonal pragmatics Vol. 6* (pp. 101–136). De Gruyter. https://doi.org/10.1515/9783110214338.1.101

Bousfield, D., & McIntyre, D. (2018). Creative linguistic impoliteness as aggression in Stanley Kubrick's Full Metal Jacket. *Journal of Literary Semantics*, *47*, 43-65.https://doi.org/10.1515/jls-2018-0003

Brueggemann, T., & Brueggemann, T. (2019, March 6). How much would 'Roma' have grossed with a traditional release? We have an answer. Retrieved June 2, 2021, from IndieWire website: https://www.indiewire.com/2019/03/netflix-roma-box-office-theatrical-release-1202041367/

Bucholtz, M., & Hall, K. (2005). Identity and interaction: A sociocultural linguistic approach. *Discourse Studies*, *7*, 585–614. https://doi.org/10.1177/1461445605054407

Conrad, S. M., & Biber, D. (2005). The frequency and use of lexical bundles in conversation and academic prose. *Lexicographica*.

Crosthwaite, P., Ningrum, S., & Schweinberger, M. (2022). Research trends in corpus linguistics: A bibliometric analysis of two decades of Scopus-indexed corpus linguistics research in arts and humanities. *International Journal of Corpus Linguistics*. Retrieved from https://www.jbe-platform.com/content/journals/10.1075/ijcl.21072.cro

Cuarón, A. (Director). (2018). *ROMA*.

Freddi, M. (2011). A phraseological approach to film dialogue: Film stylistics revisited. *Yearbook of Phraseology*, *2*(1), 137–162. https://doi.org/10.1515/9783110236200.137

Gemser, G., Van Oostrum, M., & Leenders, M. A. A. M. (2007). The impact of film reviews on the box office performance of art house versus mainstream motion pictures. *Journal of Cultural Economics*, *31*(1), 43–63. https://doi.org/10.1007/s10824-006-9025-4

Gregory, M. (1967). Aspects of varieties differentiation. *Journal of Linguistics*, *3*(2), 177–198.

Hyland, K. (2008). As can be seen: Lexical bundles and disciplinary variation. *English for Specific Purposes*, *27*(1), 4–21. https://doi.org/10.1016/j.esp.2007.06.001

Jautz, S., & Minow, V. (2020). 'Drucilla, we: The formulaic nature of problem-oriented talk in soap operas. In C. Hoffmann & M. Kirner-Ludwig (Eds.), *Telecinematic Stylistics* (1st ed., pp. 63–86). Bloomsbury Academic. Bloomsbury Collections. Retrieved from http://www.bloomsburycollections.com/book/telecinematic-stylistics/ch3-drucilla-we-need-to-talk/

King, G. (2005). *American independent cinema*. I B Tauris.

Lee, C. (2015). Digital discourse@public space: Flows of language online and offline. In R. H. Jones, A. Chik, & C. A. Hafner (Eds.), *Discourse and digital practices: Doing discourse analysis in the digital age* (pp. 175–192). Routledge.

McIntyre, D. (2012). Prototypical characteristics of blockbuster movie dialogue: A corpus stylistic analysis. *Texas Studies in Literature and Language*, *54*(3), 402–425.

Meyer, J., Song, R., & Ha, K. (2016). The effect of product placements on the evaluation of movies. *European Journal of Marketing*, *50*(3/4), 530–549. https://doi.org/10.1108/EJM-12-2014-0758

Newman, M. Z. (2011). *Indie: An American film culture*. Columbia University Press. JSTOR. https://doi.org/10.7312/newm14464

Nursanti, R. Y. (2015). *A pragmatic analysis of maxim flouting in Hunger Games Movie* [Unpublished thesis]. Yokyakarta State University.

Nuryani, E. (2016). *A pragmatic analysis of politeness features of criticism in Joseph Mcginty's This Means War.* [Unpublished thesis]. Yokyakarta State University.

Pavesi, M. (2020). 'I shouldn't have let this happen': Demonstratives in film dialogue and film representation. In C. Hoffmann & M. Kirner-Ludwig (Eds.), *Telecinematic Stylistics* (1st ed., pp. 19–38). Bloomsbury Academic. Bloomsbury Collections. Retrieved from http://www.bloomsburycollections.com/book/telecinematic-stylistics/ch1-i-shouldn-t-have-let-this-happen/

Reichelt, S. (2020). *Innovation on screen: Marked affixation as characterization cue in Buffy the Vampire Slayer*. https://doi.org/10.1075/ijcl.00038.rei

Sarris, A. (1963). The Auteur Theory and the Perils of Pauline. *Film Quarterly*, *16*(4), 26–33. JSTOR. https://doi.org/10.2307/3185951

Scorsese, M. (Director). (2019). *The Irishman* [Biography, crime, drama]. Tribeca Productions, Sikelia Productions, Winkler Films.

Shirazizadeh, M., & Amirfazlian, R. (2021). Lexical bundles in theses, articles and textbooks of applied linguistics: Investigating intradisciplinary uniformity and variation. *Journal of English for Academic Purposes*, *49*, 100946. https://doi.org/10.1016/j.jeap.2020.100946

Siricharoen, A., & Wijitsopon, R. (2020). A corpus-based comparative study of lexical bundles in authentic and textbook English business emails. *LEARN Journal: Language Education and Acquisition Research Network*, *13*(2), 41–63.

Statham, S. (2015). 'A guy in my position is a government target … You got to be extra, extra careful': Participation and strategies in crime talk in The Sopranos. *Language and Literature*, *24*(4), 322–337. https://doi.org/10.1177/0963947015605442

Thompson, K. (2003). *The Classical Hollywood Cinema: Film Style and Mode of Production to 1960*. Routledge. https://doi.org/10.4324/9780203358818

Tzioumakis, Y. (2017). Introduction: Problems of definition and the discourse of American independent cinema. In Y. Tzioumakis (Ed.), *American Independent Cinema: Second Edition* (pp. 1–14). Edinburgh University Press. Cambridge Core. Retrieved from

https://www.cambridge.org/core/books/american-independent-cinema/introduction-problems-of-definition-and-the-discourse-of-american-independent-cinema/F3DCF7D535B1314F040D96BA755CA59C

Vivarelli, N., & Vivarelli, N. (2019, November 22). 'The Irishman' is Netflix's biggest theatrical release at home and abroad. Retrieved June 2, 2021, from Variety website: https://variety.com/2019/film/news/the-irishman-martin-scorsese-netflix-biggest-theatrical-release-despite-controversy-1203412994/

Zago, R. (2020). Film discourse. In E. Friginal & J. A. Hardy *The routledge handbook of corpus approaches to discourse analysis* (pp. 168–182). Routledge.

Zuckerman, E. W., & Kim, T. (2003). The critical trade-off: Identity assignment and box-office success in the feature film industry. *Industrial and Corporate Change*, *12*(1), 27–67. https://doi.org/10.1093/icc/12.1.27

Appendix: Film scripts in MFC

Film Title	Rank of the Year	Year	Genre
Star Wars Episode III	1	2005	Action Adventure Fantasy Sci-Fi
Harry Potter and Goblet of Fire	2	2005	Adventure Family Fantasy Mystery
War of the worlds	3	2005	Adventure Sci-Fi Thriller
The Chronicles of Narnia	4	2005	Adventure Family Fantasy
Wedding Crashers	5	2005	Comedy Romance
The Pacifier	17	2005	Action Comedy Drama Family
Flight Plan	21	2005	Action Comedy Drama Family
Pirates of the Caribbean: Dead Man's Chest	1	2006	Action Adventure Fantasy
Happy Feet	7	2006	Adventure Animation Comedy

			Family Music Musical Romance
MI3	12	2006	Action Adventure Thriller
The Devil Wears Prada	14	2006	Comedy Drama
The Departed	15	2006	Crime Drama Thriller
Shrek 3	2	2007	Adventure Animation Comedy Family Fantasy
Pirates of the Caribbean: at World's End	4	2007	Action Adventure Fantasy
The Bourne Ultimatum	6	2007	Action Mystery Thriller
Ratatouille	8	2007	Adventure Animation Comedy Family Fantasy
Wild Hogs	11	2007	Action Adventure Comedy
American Gangster		2007	Biography Crime Drama
Disturbia		2007	Crime Drama Mystery Thriller
Dark Knight	1	2008	Action Crime Drama Thriller
Hancock	4	2008	Action Drama Fantasy
Wall.E	5	2008	Adventure Animation Family Romance Sci-Fi
Kung Fu Panda	6	2008	Action Adventure Animation Comedy Family
Twilight	8	2008	Drama Fantasy Romance
Sex and the City	11	2008	Comedy Drama Romance
Harry Potter and the Half-Blood Prince	2	2009	Action Adventure Family Fantasy Mystery

Up	3	2009	Adventure Animation Comedy Family
Twilight: New Moon	4	2009	Adventure Drama Fantasy Romance
Avatar	5	2009	Action Adventure Fantasy Sci-Fi
The Hangover	6	2009	Comedy
Blind Side	10	2009	Biography Drama Sport
Toy Story 3	2	2010	Adventure Animation Comedy Family Fantasy
Inception	6	2010	Action Adventure Sci-Fi Thriller
Harry Potter and the Deathly Hallows: Part 1	7	2010	Adventure Family Fantasy Mystery
Despicable Me	8	2010	Animation Comedy Crime Family Fantasy
How to Train Your Dragon	10	2010	Action Adventure Animation Family Fantasy
Social Network	31	2010	Biography Drama
The Town	32	2010	Crime Drama Thriller
Harry Potter and the Deathly Hallows: Part 2	1	2011	Adventure Drama Fantasy Mystery
Pirates of the Caribbean: On Stranger Tides	5	2011	Action Adventure Fantasy
Fast Five	6	2011	Action Adventure Crime Thriller
Cars 2	7	2011	Adventure Animation Comedy Family Sci-Fi Sport
Thor	8	2011	Action Adventure Fantasy
The Help	11	2011	Drama
The Avengers	1	2012	Action Adventure Sci-Fi
The Dark Knight Rises	2	2012	Action Adventure

The Hunger Games	3	2012	Action Adventure Sci-Fi Thriller
Skyfall	4	2012	Action Adventure Thriller
Ted	9	2012	Comedy
Lincoln	19	2012	Biography Drama History War
Despicable Me 2	3	2013	Adventure Animation Comedy Crime Family Fantasy Sci-Fi
Monster University	5	2013	Adventure Animation Comedy Family Fantasy
Frozen	6	2013	Adventure Animation Comedy Family Fantasy Musical
Gravity	7	2013	Drama Sci-Fi Thriller
The Croods	14	2013	Adventure Animation Comedy Family Fantasy
The Butler	25	2013	Adventure Biography Crime Drama Thriller
Captain Philips	28	2013	Biography Drama
Maleficent	6	2014	Action Adventure Family Fantasy Romance
X-Men: Days Of Future Past	7	2014	Action Adventure Sci-Fi Thriller
War For The Planet Of The Apes	8	2014	Action Adventure Sci-Fi Thriller
Godzilla	11	2014	Adventure Drama Sci-Fi
Interstellar	15	2014	Action Biography Drama War
Lone Survivor	24	2014	Action Adventure Family Fantasy Romance
The Fault in Our Stars	25	2014	Drama Romance

Star Wars: The Force Awakens	2	2015	Action Adventure Sci-Fi
Inside Out	4	2015	Adventure Animation Comedy Drama Family Fantasy
American Sniper	6	2015	Action Biography Drama War
The Martian	9	2015	Adventure Drama Sci-Fi
Pitch Perfect 2	13	2015	Comedy Music
Straight Outta Compton	19	2015	Biography Drama History Music
Trainwreck	27	2015	Comedy Drama Romance
Creed	28	2015	Drama Sport
Deadpool	6	2016	Action Adventure Comedy Sci-Fi
The Revenant	14	2016	Action Adventure Drama History Western
Sing	18	2016	Animation Comedy Family Musical
Sully	24	2016	Biography Drama
Bad Mom	25	2016	Comedy
Beauty And The Beast	2	2017	Family Fantasy Musical Romance
Wonder Woman	3	2017	Action Adventure Fantasy Sci-Fi War
The Guardians Of Galaxy Vol.2	4	2017	Action Adventure Comedy Sci-Fi
It	6	2017	Horror
Thor: Ragnarok	7	2017	Action Adventure Comedy Fantasy Sci-Fi
Hidden Figure	20	2017	Biography Drama History
War For the Planet of the Apes	22	2017	Action Adventure Drama Sci-Fi Thriller
Black Panther	1	2018	Action Adventure Sci-Fi

A Star Is Born	12	2018	Drama Music Romance
Bohemian Rhapsody	14	2018	Biography Drama Music
A Quiet Place	15	2018	Drama Horror Sci-Fi
Crazy Rich Asians	17	2018	Biography Drama Musical
The Greatest Snowman on Earth	24	2018	Drama Sport
Creed 2	28	2018	Action Adventure Sci-Fi
The Avengers IV: Endgame	1	2019	Action Adventure Drama Sci-Fi
Toy Story 4	3	2019	Adventure Animation Comedy Family Fantasy
Joker	9	2019	Crime Drama Thriller
Us	12	2019	Horror Mystery Thriller
Knives Out	21	2019	Comedy Crime Drama Mystery Thriller
Ford VS Ferrari	27	2019	Action Biography Drama Sport