



Starling David Hunter III¹
Susan Smith²

Received 9 January 2016
Reviewed 14 February 2016
Accepted 24 February 2016
UDC: 81'42:791]:004.4

CENTER OF ATTENTION: A NETWORK TEXT ANALYSIS OF *AMERICAN SNIPER*

Abstract: *Network Text Analysis (NTA) is a term used to describe a variety of software-supported methods for modeling texts as networks of concepts. In this study we apply NTA to the screenplay of American Sniper, an Academy Award nominee for Best Adapted Screenplay in 2014. Specifically, we establish prior expectations as to the key themes associated with war films. We then empirically test whether words associated with the most influentially-positioned nodes in the network signify themes common to the war-film genre. As predicted, we find that words and concepts associated with the least constrained nodes in the text network were significantly more likely to be associated with the war genre and significantly less likely to be associated with genres to which the film did not belong.*

Keywords: *network analysis, network text analysis, film studies, screenplay, war movie, war film, genre analysis, film genre*

Introduction

Network text analysis (NTA) is a term used to describe a broad set of software-supported solutions for modeling linguistic data as interconnected networks “of words and the relations between them” (Diesner & Carley, 2005, p. 83). The creation of these text networks involves four distinct steps, the first of which is *inclusion*, i.e. the determination of which words to include in the analysis and, by extension, which to exclude. The second step is *generalization* and it involves the decision of whether to assign the included words to higher-order conceptual categories. The third step is *inter-relation* and it concerns the choice of the

¹ Associate Teaching Professor, Carnegie Mellon University's Tepper School of Business (Doha).

² Professor, American University of Sharjah.

relationship used to link pairs of words or concepts. The fourth and final step is *extraction* and it involves the identification of the key themes encoded in the text. While several approaches to each of the four steps have been identified in the extant literature (Nerghes, Lee, Groenewegen, & Hellsten, 2015), there is very little variation concerning the fourth step. The most common approach to theme extraction is inductive. That is to say, researchers typically extract themes or key meaning from the network without any prior expectation concerning the content of these themes. A notable exception is Hunter & Singh's (2015) analysis of the screenplay for *Fight Club*.

That study is the only one with which we are familiar where key themes were identified *prior* to the text analysis. That said, we are aware of no prior research where falsifiable hypotheses were formulated concerning the *position* of themes in the text network. The purpose of the current study is to address the gap in the literature concerning the extraction of themes. In particular we rely upon the literature on the war genre (Eberwein, 2009) to establish in advance what the important themes in *American Sniper* should be and by extension, where those themes should be found in the text network constructed from the film's screenplay. Unlike prior research, we empirically test our hypotheses through the application of survey-based methods. As predicted, we find that words and concepts associated with the least constrained nodes in the text network were significantly more likely to be associated with the war and action genres—two to which the films belongs—and significantly less likely to be associated with the mystery, science-fiction, fantasy, film noir and other genres to which the film did not belong.

Literature review and hypotheses

Several recent studies have affirmed the widely held assumption that thematically-relevant words are associated with the most influentially positioned nodes in a text network. These include analyses of *nuclear energy policy frames* (Shim, Park, & Wilding, 2015), the abstracts of *medical journals* (Beam, et al, , 2014), *medical school mission statements* (Grbic Hafferty, & Hafferty, 2013), *presidential inaugural addresses* (Light, 2014), *press coverage* of the global financial crisis (Nerghes, Lee, Groenewegen, & Hellsten, 2014) and mad cow disease (Lim,

Berry & Lee, 2015), *propaganda* from violent extremists (Morris, 2014), as well as *novellas* (Hunter & Smith, 2014) and *movie scripts* (Hunter & Singh, 2015). The gap in this literature concerns the research methods that have been employed to date. Specifically, in each case it was assumed—rather than hypothesized *a priori*—that the most thematically-relevant words would be those associated with the most influential positions in the text networks. No other study of which we are aware has tested this important proposition directly. Film scripts offer an intriguing opportunity to do so. That’s because American feature films belong to long-standing, well-defined genres, a term that the *Wordsmyth English Dictionary* defines as “a particular type, sort, or category” and more specifically as “a category of artistic work marked by a particular specified form, technique or content.” The analysis of specific genres, particularly their essential codes and conventions, has a long history within film theory (Bignell, 2002; Bordwell & Thompson, 2013).

Aspects of a film that may convey its genre include, but are not limited to, its iconography, setting, locations and backdrops, props, the narrative style, dialog, emphasized camera shots, story structure, the musical score and soundscape, and lighting (Bignell, 2002; Altman, 1984). In the introduction of his book entitled *The Hollywood War Film*, Eberwein (2009) places the conventions that define the war genre into two categories—*stock characters* and *basic narrative elements*. As summarized in the first column of Figure 1, below, the stock characters are of three types—*males*, *females*, and *youth/children/pets*. The former are further subdivided into (a) the *older, seasoned leader* (b) the *young recruits* (c) the *camp/platoon clown* (d) the *ladies’ man* [e] the *newly married or recent father* (f) the *regional, ethnic, and racial types* and (g) *examples of different social classes*. The basic narrative elements are of several kinds—(1) the *basic training* that characterizes the preparation for combat (2) activities characterizing the *specific branches of the armed services* at war (3) activities or *elements common to all branches* of the armed services and where appropriate (4) the *aftermath of war*.

MALE CHARACTERS

- Older, seasoned leader
- Young recruits
- Camp/platoon clown
- Ladies' man
- Newly married or recent father
- Regional, ethnic, & racial types
- Different social classes

FEMALE CHARACTERS

- Loyal wife, girlfriend, nurse
- Prostitute, floozie
- Wise, sustaining mother

YOUTH, CHILDREN, & PETS

- Eager brothers, boys
- Younger sisters
- Endangered or killed child
- Animals (dogs, cats)

PRE-COMBAT: BASIC TRAINING

- Tyrannical squad leader
- Demanding exercises, drills
- Bonding, pranks
- Weekend passes
- Sexual initiation
- Successful graduation, completion of training

COMBAT: Army/Infantry/Marines

- Water landings, patrols, ambushes, raids, digging in
- Combat in jungles, deserts, mountains
- Tanks, grenades, flamethrowers
- Dealing with heat/cold (or elements)

ELEMENTS COMMON TO ALL BRANCHES

- Writing letters; receiving mail from home
- Sharing and observing photographs
- Listening to the radio
- Spontaneous & improvised play to alleviate tension and boredom
- Singing; prayers/church service; communion
- Burials with short, moving eulogies & tributes
- Leaves and R&R
- Reflections on the nature of the enemy

POST-COMBAT: AFTERMATH OF WAR

- Recovery/rehab for physical/psychological injuries
- Difficulty adjusting to civilian life
- Reunion with wife, girl, family, friends

Figure 1: Codes and Conventions that Define the War Film Genre—Adapted from Eberwein, (2009)

The seventh chapter in Eberwein's book is entitled *The Iraq Wars on Film* and it details the specific ways in which these conventions define the sub-genre of films about the first (Gulf) second Iraq wars. These include "cramped doorways and narrow, almost impassable streets", "endless checkpoint confrontations"—indicating the inability of soldiers to determine friend from foe—"suicide bombers' cars that explode" and "the gunfire that rains down from snipers above" (p. 134). In addition, "atrocities appear with frightening regularity" and crucial information about them, including images, is frequently found on "cell phones" (ibid., p. 135) In general, Eberwein continues, the tone of the films is "despairing" and "veterans and their families and survivors find little if any solace." Often emphasized are soldiers (1) returning home with drinking, substance abuse, or anxiety problems, e.g. post-traumatic stress disorder (2) having difficulty adjusting to prostheses or war-inflicted infirmities and (3) struggling with survivors' guilt.

Given that the war genre has very well-known conventions, it is possible to directly test whether words that signify the genre appear in influential positions within a text network. Specifically, we hypothesize that

H1: In a text network constructed from the screenplay of a war film, words associated with the most influentially-positioned will signify the codes and conventions of the war genre.

H2: In a text network constructed from the screenplay of a war film, words associated with the least influentially-positioned will not signify the codes and conventions of the war genre.

Methods and data

The war film that we chose to analyze is *American Sniper*, an autobiographical drama based on the book *American Sniper: The Autobiography of the Most Lethal Sniper in US Military History* by Chris Kyle. The screenplay was written by Jason Hall whose prior screenwriting credits include the 2009 dramedy *Spread* and the 2007 dramatic thriller *Paranoia*. The film itself was directed by Clint Eastwood and starred Bradley Cooper as Chris Kyle. The film premiered on November 11th, 2014—Veteran's Day—at the

American Film Festival Institute. A limited release followed on Christmas Day while the wide theatrical release was on January 16th, 2015. The film was an enormous commercial success. According to *Box Office Mojo*, as of June 25th, 2015—the end of the film’s run in theaters—it earned \$350.1 million domestically and another \$197.3 million internationally, making it the highest-grossing film released in 2014.

The copy of the screenplay used in this paper was the latest (second) of two drafts available for purchase from online screenplay seller *Scriptfly*. Dated July 17th, 2013, the second draft is 141 pages, about 20 pages longer than the industry standard. Interestingly, a little more than two weeks after the completion of the second draft, the originally-intended director, Stephen Spielberg, withdrew from the project—supposedly due to budgetary constraints—and was replaced by Eastwood within days. Critical response to the film has been largely positive. Thirty-three of forty reviews by “top critics” collected by *Rotten Tomatoes* are classified as “fresh” with an average rating of 7.2 out of 10. The film also received six Award nominations—Best Motion Picture, Best Performance by an Actor in a Leading Role, Best Writing (Adapted Screenplay), Best Achievement in Film Editing, Best Achievement in Sound Mixing, and Best Achievement in Sound Editing—but won only one, the latter. The screenplay itself won the *British Academy of Film & Television Arts* (BAFTA) award for Best Adapted Screenplay and was nominated for best screenplay by the *Denver Society of Film Critics*, the *Phoenix Film Critics Society*, the *Satellite Awards*, and the *Writers Guild of America*.

Building the text network

Well over a dozen distinct families of methods for constructing networks from texts have been developed and applied in the last four decades (Nerghes, et al, 2015). They can be distinguished from one another on the basis of a variety of characteristics including the degree of automation involved, whether words are abstracted to higher order categories, and the nature of the relationship used to construct the network. In this study we opted for Hunter’s (2014) morpho-etymological approach, one that is semi-automated, that abstracts words into higher-order conceptual categories defined by common

etymology, and that relates those categories based upon their co-occurrence within the same multi-morphemic compound (MMC).

MMCs may include, but are not necessarily limited to, *closed compounds* (battlefield, gunshot, Blackhawk, cowboy), *copulative compounds* (attorney-client, actor/model), *hyphenated compounds* (blast-hole, blood-clot, panic-stricken), *hyphenated multiword expressions* (follow-on-target, hand-to-hand), *infixes* (un-bloody-believable, fan-blooming-tastic), *abbreviations* and *acronyms* (GPS, HUMVEE), *blend words* (rebar= reinforced + bar), *clipped words* (internet[work], e[lectronic]-mail), *open compounds* (real estate, living room, full moon), and *pseudo-compound words* (understand, overcompensate). The text of the screenplay of *American Sniper* contained 453 such multi-morphemic compounds, about 7.8% of the number of unique words.

Our first step in creation of the text network entailed identifying the MMCs in each screenplay. To accomplish this we used the *Generate Concept List* and the *Identify Possible Acronyms* commands in the CASOS Institute's *Automap* software (Carley & Diesner, 2005). This involved two steps, the first of which was eliminating from further consideration all words in the screenplay that were not MMCs. This was accomplished through the use of a "stop list", i.e. a self-generated list of words that were previously determined to not be MMCs, e.g. *town, around, stone, cooperate, boat, sand, house, cable*, etc. The next step was to determine which of the remaining words were MMCs. We accomplished this by comparing the remaining words for each screenplay to Hunter's (2014) proprietary, Excel database which contains over 30,000 unique MMCs extracted from over 500 contemporary screenplays and teleplays. Approximately 65% of the MMCs in the screenplay of *American Sniper* were already contained in that database. All that were not contained therein were manually identified by the two authors.

The next step involved decomposing every MMC in each screenplay into its constituent morphemes. For example, the closed compound *gunshot* is comprised of the morphemes *gun* and *shot*. Next, each morpheme was assigned to a conceptual category defined by its most remote etymological root. Typically, the most remote root was Indo-European, as defined in the 3rd edition of the *American Heritage Dictionary of Indo-European Roots* (Watkins, 2010). That source assigns over 13,000 English

words to over 1,300 Indo-European (IE) roots. Over 85% of the individual morphemes in our sample were assigned to IE roots. For the example above, the two morphemes—*gun* and *shot*—we assigned to the IE roots **gwhen-**, which means “to shoot, chase, throw” (Watkins, 2011, p. 81) and **skeud-**, which means “to strike, kill” (ibid., p. 36), respectively. Where IE roots of constituent morphemes could not be identified, then etymological roots provided in the *American Heritage Dictionary of the English Language* were used. Most typically these were Latin, Greek, Germanic, or Old English. Because no software exists to etymologically stem words in this fashion, the mapping had to be performed manually before adding them to the database. At the conclusion of this process the 453 MMCs were traced back to 402 unique roots, 302 of which were Indo-European.

The next step was to create a symmetrical matrix for each screenplay where the rows and column labels were the etymological roots associated with all MMCs in the screenplay. Once the matrix was created for each screenplay, the size of the resulting network was calculated using the UCINET software program (Borgatti, Everett, & Freeman, 2002). In social network analysis, the largest cluster of mutually-reachable nodes in a network is referred to as the “main component” (Borgatti, 2006). Our indicator of positional influence—which is detailed in the next section—concerned only the network’s main component and not the many minor components.

The 402 nodes in the semantic network for the screenplay of *American Sniper* were connected by 485 links. The main component of the network, shown in Figure 2, below, contained 309 nodes connected by 404 linkages.

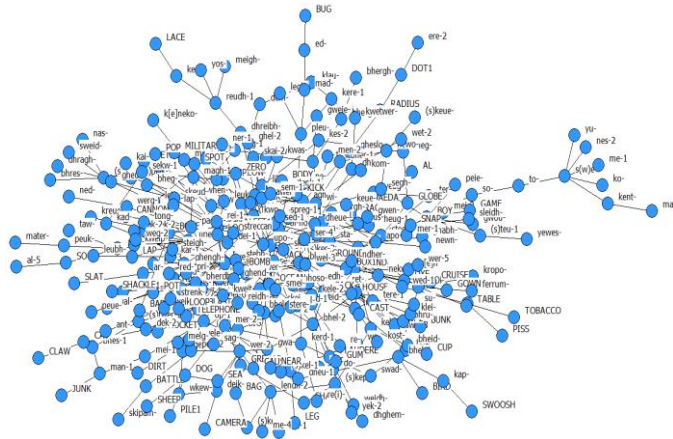


Figure 2. Main component of the text network of the screenplay of *American Sniper*

Identifying Key Themes

Following other semantic analyses of screenplays, we rely upon the network-level measure of constraint (Burt, 2000; Hunter, 2014) which captures the degree to which a node serves to link otherwise disconnected segments of a network. In this study we define the most influentially-positioned MMCs as being those that link two or more of the least constrained nodes in the text network’s main component. We defined “least constrained” as scores in the bottom 10% of the sample. In practice, this amounted to 54 nodes with constraint values of less than or equal to 0.25. As shown in Figure 3, below, the sub-network that we focused on had 43 nodes. Excluding prepositions and pronouns, these nodes were linked by the following 62 multi-morphemic compounds:

air-raid, asshole, back-and-forth, backseat, back-up, bonfire, bull’s-eye, bullshit, daylight, dog-ass, downrange, eyeball, firelight, football, forehead, GPS (Global Positioning System), gunfire, handstand, hand-to-hand, hard-headed, headlights, head-on, headshots, Hellfire, HUMVEE, JDAM (Joint Direct Attack Munition), JTAC (Joint Terminal Attack Controller), off-eye, on-board, outer-hallway, outpost, outrank, overhead, overweight, poster-boy, ranch-hand, Ranger-One, ringside, roadside,

low-constraint MMCs, the terms *cornhusker* and *ranch-hand* were referenced with regard to the “regional, ethnic, racial type” of “male character” while *downrange* appeared in the screenplay in the context of “demanding exercises, drills” under the *Pre-Combat: Basic Training* category. The hyphenated compound *outer-hallway* is an example of the “cramped doorways and narrow, almost impassable streets” that Eberwein said uniquely distinguish the Iraq-era war films.

Another low-constraint MMC was *uprange*, which occurs in the context of combat and in *Post-Combat: Aftermath of War* categories. In the latter case, the term is employed in scenes where Kyle spends time with injured fellow veterans at a shooting range, thus typifying the *recovery/rehab for physical (or) psychological injuries* convention. Finally, the terms *Ranger-One* (the name for a detachment of Army Rangers), *JDAM* (Joint Direct Attack Munition), *JTAC* (Joint Terminal Attack Controller), *Hellfire* (missile), *GPS*, *headshot*, *HUMVEE*, *ROEs* (rules of engagement), *settling in*, *off-eye* (the eye of the sniper that is not looking through the rifle scope), *gunfire*, *stand-down*, *hand-to-hand*, and *on-board* are all low-constraint MMCs that typify or signify one or more first three of Eberwein’s combat conventions—*water landings*, *patrols*, *ambushes*, *raids*, *digging in* and other maneuvering; *combat in the desert*; and *tanks*, *grenades*, and *flamethrowers* and other vehicles and weapons of war.

Although there are many fewer of them, there were also a number of highly-constrained MMCs that—when taken together—did convey some sense of combat, armed conflict, and/or the armed forces more generally. One of these was the closed compound *warfare* which appeared in the screenplay as part of the proper noun *Naval Special Warfare Center*. Notably, had that name’s acronym been used instead, the resulting MMC would have been much less constrained than was *warfare* itself. Other highly-constrained words were *chest-full* (which appeared in the phrase “a chest-full of medals”), *concertina-wire* (a type of barbed or razor wire formed in large coils and commonly used around prisons and military installations), *duct-taped*, *flack-jacket*, *middle-east*, *ill-at-ease*, *rifle-barrel*, *trigger-slack*, and *hash marks* (a service stripe on the sleeve of an enlisted person’s uniform).

Finally, two other highly-constrained MMCs were *baby-crib* and *breast-feeding* which were relevant in the context of the

“newly-married or recent father” (male character) and the “loyal wife” (female character).

In addition to this qualitative analysis—analysis which is typical of the semantic network analyses earlier reviewed—we opted to further validate our coding with an approach not previously undertaken in any study of which we are aware. Specifically, we developed a survey that allowed us to directly compare how well the two sets of MMCs conveyed not just the codes and conventions of the war genre, but also the other genres to which the film belongs and does not belong.

Empirical results

To empirically test our hypotheses, our first step was to divide the 93 words into three groups—one comprised of 31 randomly-selected low-constraint MMCs, another consisting of the remaining 31 low-constraint MMCs, and a third group consisting of the highly-constrained MMCs.³ We next recruited a sample of respondents from Amazon.com’s *mTurk* e-worker service (mTurk.com). All survey respondents had previously completed at least 5000 human intelligence tasks (HITs) with a 98% or better approval rates from other employers and were resident in the USA. Respondents were told in the introduction to the survey that they would be matching keywords extracted from the screenplay of a film to types or genres of films.

After viewing just one of the three groups of keywords, respondents were asked to answer a series of 20 questions, each of which provided a definition of a genre and which required the respondent to rate on a 1-10 scale the likelihood that the resulting film belonged to that genre. The definitions embedded in all questions were taken directly from the *International Movie Database* (IMDb.com, 2015). The question specific to the war genre read as follows: “How likely is it that this list of words was taken from a WAR film, i.e. one that contains numerous scenes and/or a narrative that pertains to a real war—past or current.”

The question regarding the action genre was worded similarly: “How likely is it that this list of words was taken from an ACTION film, i.e. one that contains “numerous scenes where action is spectacular and usually destructive.” The eighteen other genres about which the respondents provided opinions were

³ The survey may be viewed at www.surveymonkey.com/s/7PCJPVG. To access it, use “12345678901234” (without quotes) as the “mTurk Worker ID.”

Adventure, Animation, Biography, Comedy, Crime, Drama, Family, Fantasy, Film-Noir, History, Horror, Music/Musical, Mystery, Romance, Science-Fiction, Sport, Thriller, and Western. Table 1 contains a summary of the results of survey respondents.

Genre	Average	Range	St. Dev.
WAR	7.45	1 - 10	2.71
ACTION	7.74	2 - 10	1.86
BIOGRAPHY	4.38	1 - 9	2.02
DRAMA	5.51	1 - 10	2.24
Mystery	3.53	1 - 9	2.18
Sports	2.57	1 - 8	1.97
Romance	2.00	1 - 6	1.34
Crime	5.14	1 - 10	2.47
Western	3.33	1 - 10	2.49
Film noir	3.92	1 - 10	2.53
Family	5.27	1 - 10	2.53
Sci-fi	3.69	1 - 10	2.32
Fantasy	2.73	1 - 10	1.92
Horror	3.88	1 - 10	2.27
Thriller	5.63	1 - 10	2.36
Comedy	3.02	1 - 8	1.98
Animated	2.14	1 - 10	1.50
History	5.27	1 - 10	2.53
Music	1.70	1 - 6	1.07
Adventure	6.02	1 - 10	2.36

Table 1: Descriptive statistics

Table 2, below, contains the results of the regression analyses conducted to test the two hypotheses. Columns 2-5 contain the average scores given by respondents to the question concerning the film's membership in a given genre. The asterisks indicate the statistical significance of the β -coefficients of ordinary least squares (OLS) regressions where the dependent variable, $SCORE_{Genre}$ is the score given (on a 1-10 scale) by respondents to the question of whether the film belongs to a given genre. The variable $LOWCON1$ was a categorical variable equal to 1 if the respondent viewed the first set of low-constraint keywords and 0 otherwise while $LOWCON2$ was a dummy

variable whose value is 1 if the respondent viewed the second set of low-constraint keywords and 0 otherwise. Specifically, our statistical model was as follows: $SCORE_{Genre} = \alpha + \beta_1 * LOWCON1 + \beta_2 * LOWCON2 + \epsilon$

Genre	High Constraint	Low Constraint (1)	Low Constraint (2)	Model F-statistic	Model adj-R ²
WAR	6.11	7.37*	9.00****	7.74***	16.7%
ACTION	6.80	8.06**	8.41***	6.15***	13.3%
BIOGRAPHY	3.51	5.14***	4.50*	4.37**	9.1%
DRAMA	5.94	5.06	5.53	1.01	0.0%
Mystery	4.91	2.89****	2.72****	11.69****	24.1%
Sports	2.66	3.26	1.72	4.18**	8.6%
Romance	2.54	1.74**	1.69**	3.35*	6.5%
Crime	6.00	4.31**	5.09	3.72*	7.5%
Western	4.14	2.66*	3.19	2.34	3.8%
Film noir	5.03	3.03***	3.69*	4.56**	9.6%
Family	2.43	1.63	1.38*	2.28	3.7%
Sci-fi	3.51	2.94	4.69**	3.96**	8.1%
Fantasy	3.51	2.03***	2.63*	3.96**	8.1%
Horror	4.63	3.69	3.28*	2.31	3.7%
Thriller	6.51	5.17*	5.16*	3.44*	6.8%
Comedy	3.49	3.08	2.44*	2.05	3.0%
Animated	2.51	1.94	1.94	1.91	2.6%
History	4.80	4.89	6.19*	2.28	3.7%
Music	1.97	1.69	1.41*	1.60	1.7%
Adventure	6.31	6.06	5.66	1.17	0.5%

Table 2: Results of survey & Regression Model

In the first row of Table 2 we see that when asked if the film belonged to the WAR genre, responses from those who viewed the “high-constraint” words averaged 6.11 points (on a scale of 1-10). Responses of those who viewed the first and second groups “low-constraint” words averaged 7.37 and 9.00, respectively. The former score was significantly higher—at the $p < 0.05$ level, 2-tailed—than 6.11 while the latter score was even more significantly higher ($p < 0.0001$, 2-tailed). Similarly, the second row of the table indicates that the average score given to the ACTION genre by those reviewing the high-constraint words

was 6.80 while averages by those viewing these low-constraint words were 8.06 and 8.41, both of which were significantly higher. But whereas respondents who saw either group of low-constraint MMCs were far better able to identify the film as belonging to the WAR and ACTION genres, they were not able to better identify the other two genres that IMDb assigned to the film—BIOGRAPHY and DRAMA.

Respondents viewing either low-constraint group were also better able to determine the genres to which the film did *not* belong. Specifically, they were much better able to tell that the film did not belong to the FAMILY, WESTERN, ROMANCE, MYSTERY, FANTASY, NOIR, HORROR, MUSIC, COMEDY, ANIMATION, or SPORTS genres. They were slightly better able to tell that the film did not belong to SCI-FI, CRIME, or HISTORY genres. Finally, they were no better able to determine that the film did not belong to THRILLER or ADVENTURE genres.

Conclusion

We began this study by noting that the extraction of meaning from text networks requires consideration of the network's most influentially-positioned nodes. We further noted that in the prior literature, this relationship between meaning and position has rarely, if ever, been tested directly. The present study was undertaken to address this gap in the literature. To that end we developed and tested two hypotheses concerning that relationship. Specifically we hypothesized that words signifying a war film would be associated with the most influentially positioned nodes and while words associated with the least influentially-positioned ones would not. Both hypotheses were strongly supported. Specifically, we found that the words associated with the more influential positions—words like *air-raid*, *bull's-eye*, *gunfire*, *hand-to-hand*, *Hellfire*, *HUMVEE*, *JDAM (Joint Direct Attack Munition)*, *JTAC (Joint Terminal Attack Controller)*, *outrank*, *ROEs (rules of engagement)*, *shotgun*, and *stand-down* were more strongly linked by our survey respondents to the war genre than were the words associated with the less influentially-positioned nodes.

While our hypotheses were strongly supported, it is important to note the factors that limit our ability to generalize our findings beyond this study. First of all, this is an in-depth study of just one case, of one single screenplay. Secondly, the

screenplay that we analyzed was not only highly-acclaimed critically, it was also an enormous box office success. Third, the film belongs to one of the most popular and long-standing film genres, one which American film audiences know well and for which they have well-developed expectations. An excellent follow-up study would assess whether our results extend to other genres, to a larger sample, and to less well-known and less successful films. It would also be interesting to determine whether the presence of genre-signifying words in influential positions in the screenplay's network is associated with box office performance and/or critical acclaim.

References

- Altman, Rick. "A semantic/syntactic approach to film genre." *Cinema Journal* (1984): 6-18.
- Beam, Elizabeth, et al. "Mapping the semantic structure of cognitive neuroscience." *Journal of cognitive neuroscience* 26.9 (2014): 1949-1965.
- Bignell, Jonathan. *Media semiotics: An introduction*. Manchester University Press, 2002.
- Burt, Ronald S. "The network structure of social capital." *Research in organizational behavior* 22 (2000): 345-423.
- Diesner, Jana, and Kathleen M. Carley. "Revealing social structure from texts." *Causal mapping for research in information technology* (2004): 81.
- Nerghes, Adina, et al. "Mapping discursive dynamics of the financial crisis: a structural perspective of concept roles in semantic networks." *Computational Social Networks* 2.1 (2015): 1-29.
- Eberwein, Robert. *The Hollywood war film*. Vol. 13. John Wiley & Sons, 2009.
- Grbic, Douglas, Frederic W. Hafferty, and Phillip K. Hafferty. "Medical school mission statements as reflections of institutional identity and educational purpose: A network text analysis." *Academic Medicine* 88.6 (2013): 852-860.
- Hunter, Starling. "A Novel Method of Network Text Analysis." *Open Journal of Modern Linguistics* 4.02 (2014): 350.
- Hunter, Starling, and Saba Singh. "A Network Text Analysis of Fight Club." *Theory and Practice in Language Studies* 5.4 (2015): 737-749.
- Hunter, Starling, and Susan Smith. "Thematic and Lexical Repetition in a Contemporary Screenplay." *Open Journal of Modern Linguistics* 3.01 (2013): 9-19
- . "A Network Text Analysis of Conrad's Heart of Darkness." *English Linguistics Research* 3.2 (2014): 39-53.
- International Movie Database. Genre Definitions.
http://www.imdb.com/help/search?domain=helpdesk_faq&index=2&file=genres (accessed 10 January 2016).

- Light, Ryan. "From Words to Networks and Back: Digital Text, Computational Social Science, and the Case of Presidential Inaugural Addresses." *Social Currents* 1.2 (2014): 111-29.
- Lim, Seunghoo, Frances S. Berry, and Keon-Hyung Lee. "Stakeholders in the Same Bed with Different Dreams: Semantic Network Analysis of Issue Interpretation in Risk Policy Related to Mad Cow Disease." *Journal of Public Administration Research and Theory* 26.1 (2015): 79-93.
- Morris, Travis. "Networking vehement frames: neo-Nazi and violent jihadi demagoguery." *Behavioral Sciences of Terrorism and Political Aggression* 6.3 (2014): 163-182.
- Nerghes, Adina, et al. "The shifting discourse of the European Central Bank: Exploring structural space in semantic networks." *Signal-Image Technology and Internet-Based Systems (SITIS), 2014 Tenth International Conference on*. IEEE, 2014.
- Shim, Junseop, Chisung Park, and Mark Wilding. "Identifying policy frames through semantic network analysis: an examination of nuclear energy policy across six countries." *Policy Sciences* 48.1 (2015): 51-83.
- Watkins, Calvert. *The American heritage dictionary of Indo-European roots*. Houghton Mifflin Harcourt, 2000.