Title: Pledging to harm: A linguistic appraisal analysis of judgement comparing

realized and non-realized violent fantasies

Authors: Marlon Hurt and Tim Grant

Aston University, Birmingham, UK

Version: 2 (September 9, 2018)

Emails: Marlon Hurt: **hurtm@aston.ac.uk**

Tim Grant: t.d.grant@aston.ac.uk

Phone: +

Short Title: Pledging to harm

Word Count: 7,085

Please send proofs to:

Abstract

Intent is a psychological quality which threat assessors view as a required step on a threatener's pathway to action (Meloy & Hoffmann, 2013). Recognizing the presence of intent in threatening language is therefore crucial to determining whether a threat is credible. Nevertheless, a "lack of empirical guidance" (Borum et al., 1999: 326) is available concerning how violent intent is expressed linguistically. Using the subsystem of judgement in Appraisal analysis (Martin & White, 2005), this study compares realized with non-realized 'pledges to harm' (Harmon, 2008), revealing occasionally counterintuitive patterns of stancetaking by both author types—e.g., that the non-realized texts are both prosodically more violent and more threatening while the realized pledges are more ethically nuanced—which may begin to shed light on which attitudinal markers reliably correlate with an author's intention to do future harm.

1. Introduction

In 2014, Elliot Rodger emailed a 137-page manifesto to his family, his friends, and one of his therapists before killing six and wounding thirteen people outside the UC Santa Barbara campus in California. Towards the end of the document, Rodger details "every single fantasy I had about how I would punish my enemies" (Rodger, 2014: 132) on what he called his Day of Retribution. Rodger followed through on much of this imagined violence, but not all—several of the depicted assaults were left conspicuously unattempted. The current research began by asking whether the language Rodger used to express the realized fantasies differs in some systematic way from the language he used to detail their non-realized counterparts. In other words, are there linguistic features which correlate with descriptions of realized and/or non-realized episodes of violence which are imagined to take place in an author's personal future?

The answer to this question has practical implications. In 2015, for instance, the Los Angeles Unified School District closed all 900 schools, at a potential cost of $29 million, after Board of Education members received an email threatening bombings and physical assaults on the schools. The FBI later determined the threat was not credible (Branson-Potts et al., 2015). Not all such writings remain fantasies, however. That same year a New Yorker named Tyrelle Shaw published a blog post explaining his plan to "hit over a million Asian Women in the face with a stick" (Shaw, 2015). Although this threat is superficially hyperbolic, and thus potentially discreditable, Shaw nevertheless followed through on his stated ideations, managing to injure

four women, and terrify the larger community, before he committed suicide. Determining whether an author does or does not intend to act can thus have real world benefits, both in terms of where resources are allocated and how the personal safety of potential targets is secured.

However, threat assessors and other professionals tasked with navigating the types of situations presented by Shaw, Rodger, and others are not analyzing a threat as it is commonly understood in the taxonomy of Speech Act Theory (Searle, 1976), i.e., as a commissive whose proposed outcome is 1) under the speaker's control, 2) to the speaker's benefit, and, most crucially, 3) to the hearer's detriment (Shuy, 1993). Gales's (2011) significant work on this kind of *direct threat*—where the hearer is also the individual being threatened—characterizes such a text as containing an interpersonal stance of 'disalignment' between a grammatical 1st and 2nd person ('I' vs. 'you'). In other words, "the threatener is naturally poised against his or her intended audience" (Gales, 2010: 214). By contrast, writers such as Rodger and Shaw are not threatening to harm their intended readers, but rather a third, outside party. To distinguish these utterances from direct threats, then, Harmon's (2008) term *pledges to harm* is applied instead.

In a pledge to harm, the audience and the victim of the threatening language are not conflated in the same way as within a direct threat, but are rather separate, featuring disalignment between a grammatical 1st and 3rd person ('I' vs 'him/her/them'), with the 2nd person of the reader in the role of audience. Bell (2014: 298) notes that "[s]peakers design their style primarily for and in response to their audience." Where a direct threat construes an addressee who is likely "resistant," or opposed to the "position naturalised by the co-selection of meanings" in the threat (Martin & White, 2005: 62), an addressee other than the threatened party is far more likely to be

construed by the writer as non-resistant, neutral, or even as potentially sympathetic to the author's message. And so this type of threatener's linguistic style cannot automatically be assumed to follow the patterns discovered in more general studies of threatening communications (e.g., Gales, 2010, 2011).

This article explores such pledges to harm using the discourse semantic method of Appraisal analysis (Martin & White, 2005)—an approach similar to that of Gales. To do so, a small corpus of authentic pledges has been collected, and then divided into two categories: on the one hand, pledges to harm where there was some real-world attempt to realize the violent ideations outlined in the text, e.g., by the author trying or succeeding to harm the people he or she mentions; and on the other, pledges where there was no such attempt. The former are referred to as realized pledges; the latter are called non-realized.

2. Data

The dataset under examination contains fourteen pledges. Eight of these are classified as non-realized, or infelicitous, pledges, where the authors took no known steps to enact their ideations. Six are classified as realized, or felicitous, pledges, where some action was taken by the authors, up to and including the attack itself. Henceforth, the realized texts are referred to as 'R' texts and non-realized texts as 'NR,' and these designations are appended to author names to clarify which corpus they fall within, e.g., Rodger NR.

Table 1. Pledge to Harm Dataset

Realization Type	Author	Description	Word Count
Non-realized	'Archangel Michael'	Email threatening to bomb official buildings across Wyoming	104

	Brahm, Jake	Chat room post describing the simultaneous bombing of several NFL stadiums	209
	Dickens, Ebony	Facebook post threatening to shoot policemen	120
	'LA Schools'	Email threatening bomb and gun assaults at Los Angeles area high schools	354
	McKelvey, Kayla	Consecutive Twitter posts threatening Kean University students	104
	Rodger, Elliott	Excerpt of emailed 'autobiography' threatening strangers around his Santa Barbara apartment	230
	'Skyline HS'	Chat room post threatening a gun assault at Skyline High School	248
	Valle, Gilberto	Chat messages threatening to abduct and cook women	183
		Total	1,552
Realized	Hribal, Alex	Handwritten essay—subsequent knife attack wounded 21 at Franklin Regional High School	979
	Kinkel, Kip	Handwritten note—subsequent gun attack killed 2 and wounded 25 at Thurston High School	189
	Long, Gavin	Email—subsequent gun attack killed 3 policemen	653
	Rodger, Elliott	Excerpt of self-published 'autobiography' threatening a sorority house near UC Santa Barbara campus—subsequent gun attack killed 6 and wounded 13	183
	Roof, Dylann	Blog post—subsequent gun attack killed 9 church parishioners in Charleston, NC	200
	Shaw, Tyrelle	Blog post—subsequent blunt object attack wounded 4 in New York City	434
		Total	2,638

This dataset is heterogeneous in several ways. First, the texts appear in a range of registers—though fully three quarters are CMC—including blog posts, emails, message boards (e.g., 4Chan), and social media posts (e.g., Facebook). Second, ten of the fourteen are excerpts of a longer piece of writing. In these cases, the excerpt captures the beginning and ending of the

episode, i.e., the 'story' of a particular imagined violent event. Finally, the production of the texts spans a period of eighteen years, with the earliest written in 1998, and the last in 2016, although ten of the fourteen were authored between 2014 and 2016.

However, the dataset is homogeneous in certain key ways, beyond the important generic qualities each text contains as a pledge to harm, e.g., positing that the author will be personally responsible for future harm to a grammatical 3rd Person. First, and most practically, each text is a minimum of 100 words. Second, each is an authentic linguistic production, and thus a legitimate forensic text. Third, evidence indicates that each text was composed before—or in the case of Kinkel R, between—any attacks, meaning each somehow discusses violent action(s) which the author has not yet attempted. Fourth, no matter the author's ethnic background (where such information is known), all are written in a relatively informal register of Standard American English. Finally, each was intended to be received and read by a particular audience, however narrow (one individual) or broad (loved ones, legal authorities, etc.), meaning no text in the dataset was a private production later unearthed against the author's will, like a journal entry.

3. Approach to the data: Judgement in Appraisal

In the psychological literature, a fantasy of committing violence is an example of what is called episodic future thinking, or EFT (Regis, 2013; Schacter et al., 2008), a neurological process which "involves the construction of possible future personal episodes or scenarios" (Schacter et al., 2015: 14). Fantasy is primarily understood as a response to some kind of emotional stimulus. For instance, "we feel hungry, a physical emotion, and have the fantasy of eating" (Freeman & Kupfermann, 1988: 4). Whether an intent to realize the imagined episode is subsequently formed

is largely "based on the intensity of the emotional impact of an experience" (Mueller & Dyer, 1985: 5), i.e., of the stimulus that spurred the EFT. In the threat assessment literature, intent is also recognized as an added element, but one whose presence necessarily precedes action. Meloy and Hoffmann (2013: 3), for instance, formulate the basic sequence a threatener progresses through on the way to physical violence as "goal → intent → behavior."

However, because both fantasy and intent are private psychological processes, accessing either in their pre-behavioral forms is impossible (Singer, 2014). And even in cases where intention is strong, actions may not be realized due to practical constraints. For example, a review of suicide prevention literature suggests that even for those with a strong intent to realize suicidal ideations, restricting access to lethal methods reduces successful attempts (Mann et al., 2005). Whether or not this extrapolates to pledges to harm, this invisibility of intent means that the absence of action on the part of an author cannot automatically be conflated with a true absence of intent when a pledge was written. All that is available for analysis is the language used to express the violent fantasy and any correlation with the subsequent behaviors of the author.

Nonetheless, if violent fantasy is a response to emotion then linguistic tools which assess an author's affective stance towards committing imagined future violence are best suited to reveal the strength of this response, and to possibly identify markers encoding any intent to act. Because Appraisal provides a framework for, among other things, identifying "the linguistic mechanisms for the sharing of emotions" and the "subjective presence" of authors in the texts they generate (Martin & White, 2005: 1), this method is ideally suited to uncovering any such markers of affective strength.

Appraisal is divided into three separate but interlocking systems, each aimed at uncovering a different set of prosodic meanings within a text. The first of these, called attitude, addresses the "kinds of feelings" a writer may seek to communicate. These feelings are divided into "three semantic regions covering what is traditionally referred to as emotion, ethics and aesthetics" (Martin & White, 2005: 42), concepts which correspond to what the authors call affect, judgement, and appreciation, respectively. Where affect concerns the inner experience of an author, and appreciation speaks to a writer's evaluations of 'things,' the subsystem of judgement, which is the focus of this article, encompasses the writer's "attitudes to people and the way they behave" (Martin & White, 2005: 52). These feelings are directed outward by the authorial voice, to address the social and moral behavior of external human actors. Judgement itself is subdivided into five variables, situated at two different strata. The first stratum, called social esteem, is a collection of meanings central to social networks, and includes questions of *normality* (how special is a person), *capacity* (how capable) and *tenacity* (how dependable). Social sanction, by contrast, captures values more central to our roles as citizens, including *veracity* (how truthful) and *propriety* (how ethical).

The dataset was first coded for these five variables of judgement using a specialized concordancer, called UAM CorpusTool (O'Donnell, 2017). A sample illustrating the coding of these five variables is excerpted from Hribal R's text, with attitudinally laden clauses and lexemes in italics, e.g.:

> all this was caused by was *dehumanization* [-prop] of public school. When I go there, it reminds me why I am doing this. All public school is is *trash* [-norm] teaching *trash* [-norm]. *Laziness* [-ten] teaching *ecstasy* [+norm]. *Selfishness* [-prop] teaching *addiction* [-

cap]. Dozens of teachers teaching 1200 students and almost all just *want* [-ten] drugs, alcohol, sex, and/or money.

Chi-square was then applied to the resulting token counts, revealing a statistically significant difference in usage between the realized and non-realized corpora across three of the coding features—capacity, tenacity, and propriety.

Table 2. Statistical Significance in Judgement

Judgement Type		Token Frequency per 1000 words		Probability (p)
		Non-realized	Realized	
Social Esteem	Normality	9.74	17.26	< .1
	Capacity	36.08	13.49	< .01
	Tenacity	1.15	6.90	< .05
Social Sanction	Propriety	15.46	29.49	< .05
	Veracity	1.15	2.20	< .5

Once this quantitative effort pinpointed these variables as warranting further investigation, various grammatical and semantic investigations were undertaken on the specific language of all three areas.

4. Analysis

Of capacity, tenacity, and propriety, judgements of tenacity account for very little of the data. However, capacity and propriety prove to be rich semantic resources for pledge authors, and so are the focal points of this analysis. Of the areas of potential divergence between the two corpora, it is interesting to find a statistical difference in usage with these two variables particularly. In the threat assessment literature, attention is paid not just to how a threatened event *may* occur but also to *why* the threatener feels the action is necessary. As Geurts et al. (2016: 55) say, "one's wish to live in a just world reflects desirability whereas all actions taken

to create a just world reflect feasibility." In many ways, propriety and capacity map to this delineation of desirability and feasibility—the why and the how of a threat, respectively. A pledge author's concern with the ethical meanings found within propriety is, *ipso facto*, a concern with how just the world is, and thus speaks to the author's possible motivations. Meanwhile, capacity is especially suited to coding meanings related to actions, and violent actions especially, and thus speaks to how an author may act to set the world right.

4.1. Capacity

One would expect that texts whose primary focus is the harming of others would be deeply concerned with the meanings of incapacity, since the desire to injure or kill someone represents an explicit assault on that person's capabilities, i.e., his or her physical capacity. For instance, Valle NR is threatening to curtail the personal agency of his imagined victim when he writes *I can knock her out*. Similarly, instilling fear in a targeted audience, or behaving so far beyond the bounds of normalcy that an audience is unable to comprehend the actions in question, constitutes an attack on that target's mental capacity, something evident in Kinkel R's assertion that *the embarrassment would be too much for them*. As a resource for expressing the kinds of physical and mental injury a pledge author imagines inflicting on others, lexical and phrasal tokens encoding negative capacity provide a means of quantifying a pledge's level of violent ideation, thus opening a window onto the emotional temperature of a text.

Incapacity covers a wide range of semantic and discursive meanings in the two corpora, from denoting violent acts like *massacring*, to derogating or infantilizing various classes of people—e.g., *faggots*, *cunts*, *girls* (for adult women), *kids* (for teenaged peers)—down to more

straightforward, anodyne judgements of impotence such as *your security will not be able to stop*

us (LA Schools NR). Instances of violent physical incapacity, such as when Kinkel R writes *I*

have to kill people, explicitly represent the 'harm' in a pledge to harm. Such instantiations are

therefore central to pledges to harm as a genre, and are prevalent in both realization categories.

However, the relative strength of this desire to harm others is distributed unevenly across the two

corpora. A logical hypothesis would be that people who felt strongly enough to proceed from

language to violence—i.e., the authors of the realized pledges—would devote more energy to

describing their imagined actions. As Figure 1 shows, though, this is not the case.

Figure 1. Violent vs Non-Violent Lexemes of Incapacity (Frequency per 1000 words)
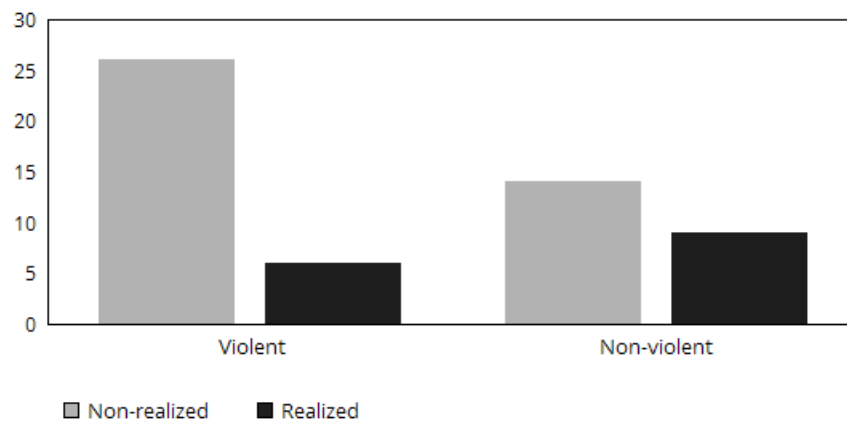
Instead, the non-realized texts are more preoccupied with violent ideation, in two quantitative

aspects. First, the sheer number of tokens denoting violent acts (*kill*, *behead*, *open fire*) and the

results of violence (*massacre*, *pain*, *suffering*) is far higher in the non-realized texts, at a normed

frequency of 26 per 1000 words versus just 6 in the realized corpus. This means that a token of

violent incapacity is four times more likely to appear in the writing of an author who

theoretically *lacks* real-world intent. Second, the overall ratio of violent to non-violent negative

tokens (*faggots*, *luring*, *asleep*) is also higher in the non-realized corpus, appearing at a rate of

almost 2:1. By contrast, the realized authors prefer nonviolent judgments of incapacity at a rate

of 1.5:1. Taken globally, the non-realized texts are thus far more violent than their realized

counterparts.

When the violent lexemes are broken into their parts of speech, an additional, somewhat

counterintuitive phenomenon emerges, captured by Figure 2. Three content word types are used

to communicate violence in the two corpora: nouns, verbs, and adjectives (violent adverbs do not

appear). Of these three, verbs naturally spur great analytical interest, since they most directly

address a basic question every assessor faces when analyzing a communication of this type: what

action is the author threatening to perform?

Figure 2. Violent Lexemes Parts of Speech (Frequency per 1000 words)
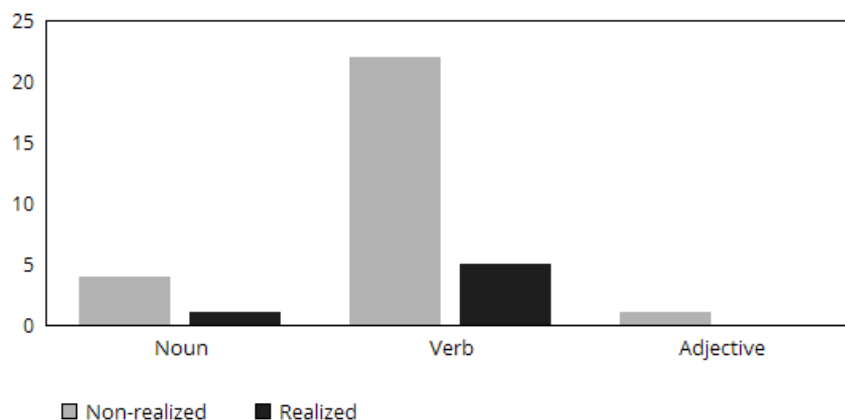
While the difference presented in Figure 2 is striking, it should be noted that a substantial

proportion of the verbs in the non-realized corpus are accounted for by just two authors—Rodger

NR and Valle NR. Without these texts, the normed count of violent verbs falls from a frequency

of 22 per 1000 words to 12 per 1000. However, even controlling for these potential outliers,

which will be excluded from the remainder of the discussion of incapacity, a non-realized text is

still over twice as likely to employ a verb of violence.

Despite the difference in frequency, however, the lexical diversity of the verbs employed is essentially equal in both realization types. Thus, though the non-realized texts contain more mentions of violent actions, the types of violence described are not more 'imaginative' in one realization category versus the other. Both share the relatively unmarked troponyms *kill*, *die*, and *destroy*, for instance. (And arguably, the senses of *slaughter* and *massacre* have extensive overlap.) Interestingly, the two areas of divergence—lexemes which appear in one corpus and not the other—coalesce around particular semantic fields. Meanings related to hand-to-hand assault (*hit*, *punch*) only appear in the realized texts, while those related to firearms (*shoot*, *open fire*) only appear in the non-realized. This remains true even if Rodger NR and Valle NR are readmitted. This is despite the fact that four of the six realized authors would use guns in their subsequent attacks.

Table 3. Violent Verb Lemmas

Verb	Frequency per 1000 words	
	Non-realized Corpus	Realized Corpus
KILL	5	1
DIE	2	1
DESTROY	1	1
HIT	-	2
PUNCH	-	1
ATTACK	-	1
SLAUGHTER	-	1
SHOOT	2	-
MASSACRE	1	-
OPEN FIRE	1	-

While these verbs explicitly represent the menace of a pledge to harm, not all of the violent actions are located by the authors in a future time. To understand how much of each text type is, in fact, commissive, it is necessary to group these lexemes according to their temporal frame.

Grammatically, 'future' includes verbs that are modified by prediction modals (*will*, *shall*), or governed by circumstantial (*tomorrow*, *soon*) or conjunctive adjuncts (*then*) which locate actions in a future time (Lock, 1996). The verbal construction *be going to* is also considered (see Lock, 1996: 150). Examples of these various resources of future framing in the dataset include *i <u>will</u> kill all the blacks tonight* (McKelvey NR), *i <u>am going to</u> open fire* (Skyline NR), and *Might kill at least fifteen <u>tomorrow</u>* (Dickens NR).

Figure 3. Violent Verbs by Tense (Frequency per 1000 words)
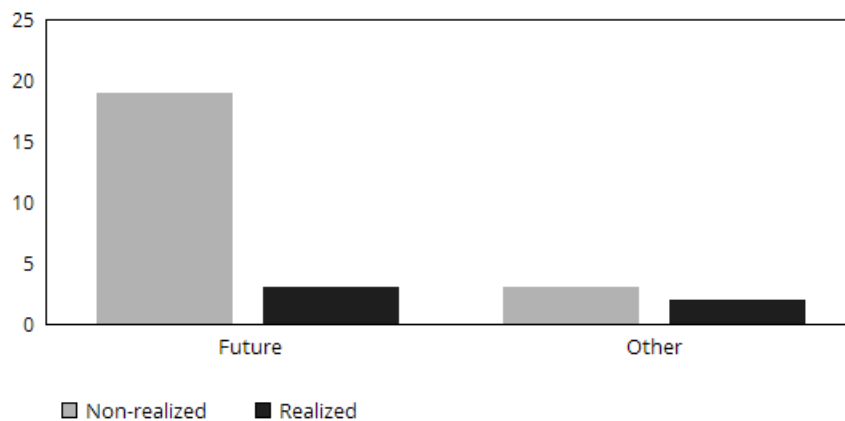

The result is perhaps counterintuitive: a non-realized text is over six times as likely to locate a violent event in a future time. With Rodger NR and Valle NR reintegrated, verbs of future violence spike from 11 tokens per 1000 to 19, nearly ten times the frequency of the realized corpus. In general, then, a non-realized text contains more instances of violent ideation (though this ideation is not more semantically varied than in a realized text), while also locating this violence in a future time frame. Thus, those pledges which are not just more violent, but also more threatening, are ones the authors would historically make no attempt to enact.

4.2. Propriety

Where Martin and White (2005: 52) discuss capacity in terms of 'social esteem,' an area that is "critical to the formation of social networks," they consider propriety a type of 'social sanction,' an area which encompasses ideas of "civic duty and religious observances." Moving from the former to the latter, the authors analogize, is comparable to shifting from venial to mortal sins. Transgressing against social expectations at this level, then, is thought to have much more dire repercussions: "too much negative esteem, and we may need to visit a therapist; too much negative sanction, and a lawyer may need to be called in" (Martin & White, 2005: 53). Propriety thus serves as a gauge for how well or poorly an author believes a third party has upheld a given ethical value. As such, it covers a range of meanings from the approving—e.g., 'good,' 'kind,' 'charitable'—to the disapproving—e.g., 'bad,' 'cruel,' 'selfish' (Martin & White, 2005: 53). Examples of positive and negative propriety from the two corpora include: *we are all selfish* (Hribal R); *They are all spoiled, heartless, wicked bitches* (Rodger R); *Hell they condone crimes against us* (Dickens NR).

It is no surprise to find an abundance of such judgements in writings which contemplate ethically-charged topics like assault and murder. Nor is it surprising to see that a healthy majority of the tokens present across the corpora, or roughly 73% from the entire dataset, are negative.

Figure 4. Tokens of Propriety by Realization Category (Frequency per 1000 words)
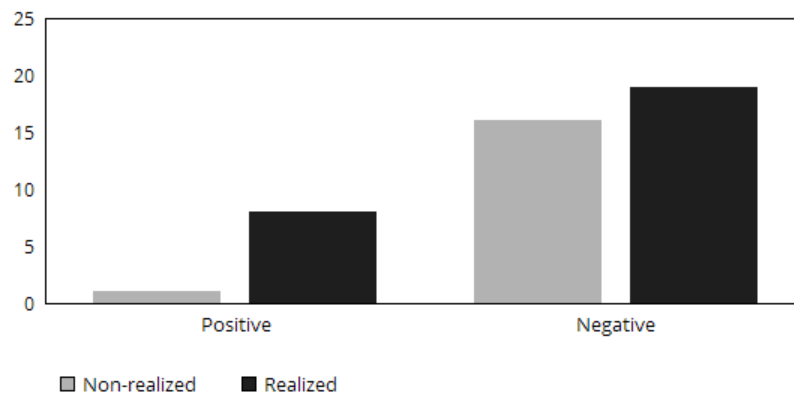
Of the fourteen authors in the dataset, all but one employ judgements of propriety, and negative propriety in particular. Interestingly, the one author who does not make use of this resource, Archangel Michael NR, still voices ethical concerns, just through the matrix of social valuation, a variable within the separate Appraisal subsystem of appreciation which nevertheless has direct correlations to propriety (O'Donnel, 2007). Thus, for all intents and purposes, some preoccupation with morality is universal across the dataset.

However, while the two corpora share a preponderance of negativity, the distribution of negative to positive tokens is not uniform, as Figure 4 shows. The non-realized authors judge a third party as being somehow unethical in nearly 93% of the propriety tokens which appear in the corpus. Somewhat surprisingly, the negativity of the realized texts is lower rather than higher. For these writings, people's perceived moral failings are the content of this class of judgements just 70% of the time. In other words, although the realized corpus is quantitatively more negative, the non-realized writers are nearly 33% more likely to use the semantic resources of propriety to accuse people of having failed their more solemn social obligations (be they legal, religious, etc.) than those authors who would later follow through on their expressed ideations. So, although the

realized corpus makes greater use of propriety—with 27 tokens per 1000 compared to the NR

corpus's 17—the realized corpus is also more evenhanded in the judgements meted out.

This result is, on the one hand, predictable, and, on the other, somewhat unexpected. Returning

to Martin and White's (2005) analogy, if negative sanction is comparable to mortal sin—or, in

secular terms, some class of felony—then it is the potential rationale for the harshest punitive

actions a society may deliver, including a death sentence. Propriety is thus the strongest resource

a fantasist may call on to justify violent ideation, something like *Person X is bad, therefore*

Person X deserves harm, therefore I will harm Person X. And so it makes a certain amount of

sense that those authors who proceeded to action would include more of this kind of meaning

than their inactive counterparts. Less expected, however, is the heightened proportion of

impropriety in the non-realized corpus. The non-realized authors overwhelmingly devote their

energy to damning others, at a rate of 16:1, whereas the realized writers indulge in blame rather

than praise at a rate of just over 2:1. The prosodic effect of this disparity is that the non-realized

corpus appears far more incriminatory, or, put in colloquial terms more in line with Martin and

White's (2005) metaphor of mortal sin, more 'fire and brimstone.' Conversely—and again,

somewhat counterintuitively—positive propriety (e.g., *forgive*, *protecting*, *serves*) is eight times

more likely to appear in the realized texts than the non-realized, softening the condemnatory tone

of writings which turned out to contain real-world intent.

4.2.1. Who is Unethical?

The substance of a great deal of the negative tokens present in the two corpora are predictable

from the generic character of these pledges. For example, third parties are said to *exploit*, *hate*,

and *steal* from others. They are portrayed as *evil*, *disgusting*, *heartless*, and so on. Virtually none

of these texts, either inter- or intra-corporally, share a concern with a single, homogeneous type

of 'bad' person or 'bad' action. Instead, as this small sampling shows, the ethical stances enacted

by the authors in the dataset are diverse—drawing from social issues, political grievance,

criminal activity, etc. However, certain patterns are apparent within this widespread negativity.

According to Foster (2003: 36), ethics is tied inextricably to justice, and justice "is about

receiving one's due or getting what one deserves." Or, as he says elsewhere, "[j]ustice served is

ethics realized" (Foster, 2003: 35). Thus, exploring the various meanings of impropriety begins

with discovering who, in these texts, is deserving of what. A second question naturally flows

from this, this time from a threat assessment perspective: are the perceived victims of future

violence also the objects of judgement in these texts, and, if so, how are they judged?

'Victims,' of course, refers to those parties who are specified (to whatever degree) as the targets

of violence. Identifying Victim as a category of referring expressions is to identify it as a topic

of, or semantic macro-structure in, these data. As van Dijk (1977: 16) defines it, "[t]he topic of

discourse is a semantic structure which we take to be identical with the macro-structure of the

discourse," and is recognizable through a small range of properties, for example by the

possibility of using anaphoric pronouns and definite articles when no "co-referential

expression…has occurred in the previous part of the discourse." Establishing a topic of discourse

is to set boundaries on "the kind of possible events and actions which may take place in an

episode" (van Dijk, 1977: 16), thereby forming a kind of touchstone for which propositional

content is judged coherent or not. Brahm NR's text provides the clearest example, in that every

single sentence somehow relates to the bombing of the NFL stadiums, i.e., each proposition is sensible and coherent with its neighbors because it "originate[s] in the same range of semantic space" (van Dijk, 1977: 6), that of a specific terrorist attack and its aftermath. The topical Victim(s) of each text have been isolated according to these parameters and are shown in Table 4.

Table 4. Primary Target by Text

Corpus	Author	Victim(s)
Non-realized	Brahm	NFL stadiums (metonym for the fans present)
	Dickens	all white cops nationwide
	LA Schools NR	students at every school in the L.A. Unified district
	McKelvey	every black person i see (at Kean University)
	Rodger	all of the men who have had pleasurable sex lives
	Skyline	people in the commons (area of Skyline High School)
	Valle	her (Victim-1)
Realized	Hribal	students of one of the "best schools in Pennsylvania"
	Kinkel	people
	Long	bad cops, good cops
	Rodger	Alpha Phi Sorority
	Shaw	every Asian Woman by herself

Of course, not every token of propriety is aimed only at variations of these expressions. Authors in both corpora judge a range of other human categories, both positively and negatively. For instance, when Shaw R says that *people hurt my feelings*, this token is better logged under a header like 'other,' because 'people' in general are not the topical objects of Shaw R's ire. But when he says *I don't think Asian Women like me*, thus casting aspersion on the preferences of Asian women, this token *is* catalogued as a judgement of his intended 'Victim(s).'

Figure 5. Objects of Propriety (Frequency per 1000 words)
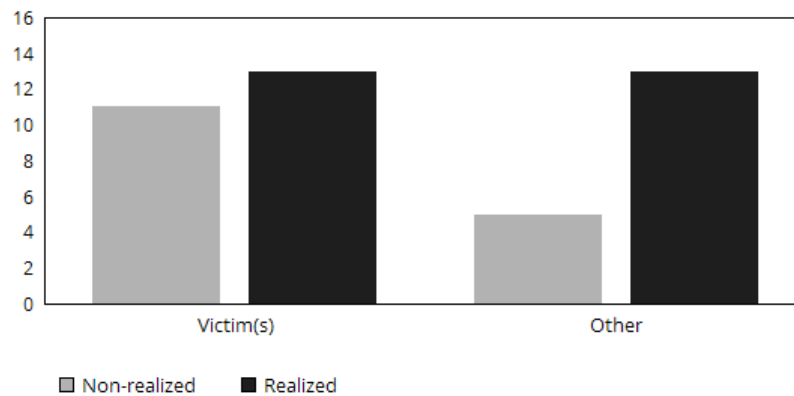
Clearly, there is no great difference in how focused the authors in each corpus are on the ethical qualities of their imagined victim(s). The non-realized authors are almost as likely to judge these people as somehow unethical as their realized counterparts. Remembering that propriety is one of the strongest resources a writer may call on to advocate for punitive measures, it is interesting to find so little daylight here between the two threat categories. Figure 5 also shows that the non-realized authors are more likely to focus on their imagined victims to the exclusion of other, outside entities, where the realized authors split their focus evenly between the two.

The data itself suggests a possible taxonomy of these Objects of Propriety. As illustrated in Figure 6, the category of Other may be further subdivided into three separate, roughly delineated areas, the first two being Agent-less yet human-driven events (such as the process noun phrases *attacks* and *civil wars*), and whatever non-victims are named by the authors (e.g., *jihadist cell*, *justice system leader's*, *everyone who hates blacks people*). The third subcategory, referred to simply as 'generic,' encompasses expressions whose extensions are so broad as to be essentially global. Such indefiniteness may be signaled via grammatical devices like the 1st Person plural (*we are all selfish*), or the 'generic you' (*making your fellow man suffer*), as well as through non-

delimited noun phrases (*humans* should be nice, *people* stain the world with sins), and 'allness' quantifiers (*the whole World Hates me*).
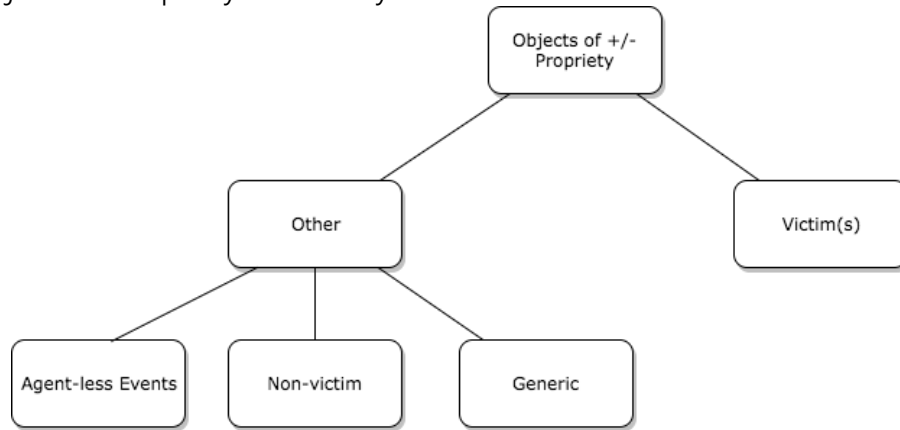
Figure 6. Objects of Propriety Taxonomy

Figure 7 captures what results when the three leaves of Other are viewed in terms of token frequency. Beginning with non-victims, the realized writers are twice as likely as their counterparts to aim ethical judgements at this subcategory. With this in mind, the realized texts read as less single-minded in their focus on some 'bad' entity the author imagines harming—demonstrating a more nuanced worldview than is apparent in the non-realized texts, which are far more victim-oriented.

Figure 7. 'Other' Objects of Propriety (Frequency per 1000 words)
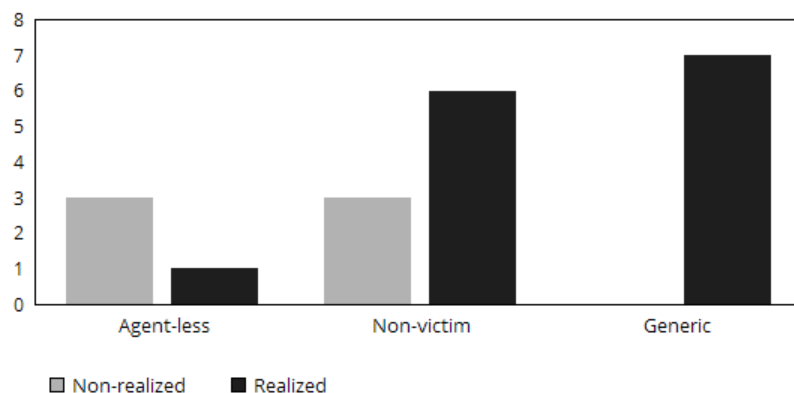
An additional pattern occurs with Agent-less events. The realized texts include just one token of propriety in such a thematically passive construction, i.e., there are technically fewer than one

Agent-less instances per 1000 words. In the non-realized texts, such Agent-less events constitute

fully half of the Other items present (three of the six tokens per 1000 words). The prosodic effect

of this is that the realized texts leave far less room for doubt concerning who the authors blame

for what.

Finally, and perhaps most significantly, fully half the tokens of Other in the realized writings (7

per 1000 words out of a total of 14) are generic, essentially addressing society or the world at

large. Interestingly, the non-realized texts employ none of the same wide-angle devices to deliver

judgements of propriety, limiting themselves rather to people or communities that are either

more circumscribed or simply unnamed. The realized texts are, thus, far more likely to let their

ethical discussions range across the entire spectrum of possible targets, up to—at its widest

end—a status quo which can include society in general or even humanity itself.

5. Conclusion

The question of intent and its absence is thorny, not least because of the real-world consequences

which hang on how the question is answered. In the U.S. and many other countries, for instance,

criminal intent, or *mens rea*, is often weighed as an equal counterpart to whatever criminal act, or

actus reas, is being prosecuted, so that two otherwise identical offenses are punished differently

depending on whether the defendant meant the resulting harm, e.g., involuntary manslaughter vs

a potential charge of murder (18 U.S.C. § 1111-1112). Similarly, the threat assessment literature

treats intent as a necessary precursor to violence, situating it directly between the idea of

violence and the act itself (Meloy & Hoffmann, 2013). Intent is thus a quality the legal system not only assumes to exist, but one that can be both detected and measured.

Theorizing intent as an element which is added to violent ideation "based on the intensity of the emotional impact of an experience" (Mueller & Dyer, 1985: 5) invites a comparison between expressions of imagined future violence which were indeed realized against those where no (or different) action was taken. Viewed through the lens of the Appraisal subsystem of judgement, discriminating patterns of linguistic difference do occur between realized and non-realized pledges to harm. The differences between the two groups are not, however, necessarily intuitive. For instance, an examination of incapacity, a variable which is ideally suited to coding descriptions of violence, reveals that the non-realized pledges in the dataset actually contain more violent ideation than their realized counterparts, and in several regards. First, non-realized texts are over four times more likely to employ a token of violent incapacity. Second, these writings also feature a higher ratio of violent to non-violent lexemes, at 2:1 where the realized texts opt for *non*violent incapacity at 1.5:1. Third, a non-realized author is between six and ten times more likely to place this imagined violence in a grammatical future time. Thus, a heightened sense of impending harm across a pledge should not reflexively be conflated with the author's intent to see that harm done.

Further to this, the path to violent ideation is said to begin with a grievance (Calhoun & Weston, 2015), i.e., the belief that an external actor is responsible for some injustice which the author imagines rectifying through violence. The variable of propriety seems to verify this basic assumption, uncovering similar levels of opprobrium aimed at the imagined victims of both

realization categories. However, a wider exploration shows that the realized texts also discuss the ethical qualities of non-victim actors and institutions at twice the rate, including more global stand-ins for the status quo such as society and the world at large. Taken together, the non-realized pledges' relentless negativity and laser-like focus on their imagined victims reveals a stance which is far more single-minded in its condemnation.

Thus, although "the literature suggests that all threateners find themselves on a pathway between an idea to cause harm and the actual implementation" (Geurts et al., 2016: 55), pinpointing just how far along an author may be, and whether that location qualifies as *mens rea*, is an important question. The long-term goal of this research is to develop as concise an instrument as possible for identifying the presence of linguistic features indicative of intent. Any such diagnostic tool must, of course, take a holistic approach to the data. A first step towards this will be to analyze these pledges according to the remaining areas of attitude (affect and appreciation) and the two other systems of Appraisal (engagement and graduation) to better understand how intent may or may not be encoded across a pledge to harm as a unit of *discourse*. While the currently available data limits the degree to which broad generalizations can be drawn from this project, there is nevertheless a legal need to evaluate intent in pledges to harm. The research presented here thus provides a proof of concept that elements of Appraisal analysis can indeed be used to address this aim and so provide a linguistic analysis to bridge the gap between the expression of a pledge to harm and the psychological intent to realize that harm.

Appendix

Below is a sample coding of the five variables of judgement (normality, capacity, tenacity, veracity, and propriety) performed on Rodger NR's pledge to harm. This is offered as a more comprehensive example of the data and how these texts may be understood via this subsystem of Appraisal. Attitudinally laden phrases and lexemes appear in italics beside their classification and polarity.

> Text 1: Rodger Non-realized Pledge to Harm
>
> On the day before the Day of Retribution, I will start the First Phase of my *vengeance* [-prop]: Silently *killing* [-cap] as many people as I can around Isla Vista by *luring* [-cap] them into my apartment through some form of trickery. After that, I will start *luring* [-cap] people into my apartment, *knock them out* [-cap] with a hammer, and *slit their throats* [-cap]. I will *torture* [-cap] some of the *good looking* [+norm] people before I *kill* [-cap] them, assuming that the *good looking* [+norm] ones had the best sex lives. All of that pleasure they had in life, I will *punish* [-cap] by bringing them *pain* [-cap] and *suffering* [-cap]. I have lived a life of pain and suffering, and it was time to bring that pain to people who actually *deserve* [-prop] it. I will *cut* [-cap] them, *flay* [-cap] them, *strip all the skin off their flesh* [-cap], and *pour boiling water all over them while they are still alive* [-cap], as well as any other form of *torture* [-cap] I could possibly think of. When they are *dead* [-cap], I will *behead* [-cap] them and *keep their heads in a bag* [-cap], for their heads will play a major role in the final phase. This First Phase will represent my *vengeance* [-prop] against all of the men who have had pleasurable sex lives while I've had to suffer. Things will be fair once I make them *suffer* [-cap] as I did. I will finally even the score.

References

Bell, A. (2014). *The guidebook to sociolinguistics.* John Wiley & Sons.

Borum, R., Fein, R., Vossekuil, B., & Berglund, J. (1999). Threat assessment: Defining an

    approach to assessing risk for targeted violence. *Behavioral Sciences & the Law*, *17*(3),

    323.

Borum, R., & Reddy, M. (2001). Assessing violence risk in Tarasoff situations: A fact-based

    model of inquiry. *Behavioral sciences & the law, 19*(3), 375-385.

Branson-Potts, H., Ceasar, S., & Blume, H. (2015). L.A. schools to reopen Wednesday; threat

    against schools was 'not credible,' officials say. *Los Angeles Times*. Retrieved from:

    http://www.latimes.com/local/lanow/la-me-ln-all-lausd-schools-closed-threat-20151215-

    story.html

Calhoun, F. S., & Weston, S. W. (2015). Perspectives on threat management. *Journal of Threat

    Assessment and Management, 2*(3-4), 258.

Foster, G. D. (2003). Ethics: Time to revisit the basics. *The Humanist, 63*(2), 30.

Freeman, L., & Kupfermann, K. (1988). *The power of fantasy: Where our daydreams come from

    and how they can help or harm us*. New York: Continuum.

Gales, T. (2010) 'Ideologies of Violence: A Corpus and Discourse Analytic Approach to Stance

    in Threatening Communications', unpublished dissertation, University of California,

    Davis.

Gales, T. (2011). Identifying interpersonal stance in threatening discourse: An appraisal analysis.

    Discourse Studies, 13(1), 27-46.

Gellerman, D. M., & Suddath, R. (2005). Violent fantasy, dangerousness, and the duty to warn

    and protect. *Journal of the American Academy of Psychiatry and the Law Online*, *33*(4),

    484-495.

Geurts, R., Granhag, P. A., Ask, K., & Vrij, A. (2016). Taking threats to the lab: Introducing an experimental paradigm for studying verbal threats. *Journal of Threat Assessment and Management, 3*(1), 53.

Harmon, A. G. (2008). Back from wonderland: A linguistic approach to duties arising from threats of physical violence. *Cap. UL Rev., 37*, 27.

Holmes, E. A., Craske, M. G., & Graybiel, A. M. (2014). A call for mental-health science. *Nature, 511*(7509), 287.

Lock, G. (1996). *Functional English grammar: An introduction for second language teachers.* Cambridge University Press.

Manslaughter, 18 U.S.C. § 1112 (1948)

Mann, J. J., Apter, A., Bertolote, J., Beautrais, A., Currier, D., Haas, A., ... & Mehlum, L. (2005). Suicide prevention strategies: a systematic review. *Jama, 294*(16), 2064-2074.

Martin, J. R., & White, P. R. (2005). *The language of evaluation.* Palgrave Macmillan.

Meloy, J. R., & Hoffmann, J. (Eds.). (2013). *International handbook of threat assessment.* Oxford University Press.

Meloy, J. R., & O'Toole, M. E. (2011). The Concept of Leakage in Threat Assessment. *Behavioral Sciences & The Law, 29*(4), 513-527.

Mueller, E. T., & Dyer, M. G. (1985). Daydreaming in Humans and Computers. *Proceedings of the Ninth International Joint Conference on Artificial Intelligence*, 278-280.

Murder, 18 U.S.C. § 1111 (1948)

National Conference of State Legislatures (2015, September 28) *Mental health professionals' duty to warn*. Accessed July 20, 2017: http://www.ncsl.org/research/health/mental-health-professionals-duty-to-warn.aspx

O'Donnell, M. (2007). UAM Corpus Tool 3.3g.

Regis, M. (2013). The Major Models of Daydreaming. In *Daydreams and the Function of Fantasy* (pp. 31-60). Palgrave Macmillan UK.

Rodger, E. (2014). *My Twisted World.* (Unpublished) Retrieved from http://www.scribd.com/doc/225960813/Elliot-Rodger-Santa-Barbara-mass- shooting-suspect-My-Twisted-World-manifesto

Schacter, D. L., Addis, D. R., & Buckner, R. L. (2008). Episodic simulation of future events. *Annals of the New York Academy of Sciences, 1124*(1), 39-60.

Schacter, D. L., Benoit, R. G., De Brigard, F., & Szpunar, K. K. (2015). Episodic future thinking and episodic counterfactual thinking: Intersections between memory and decisions. *Neurobiology of learning and memory, 117*, 14-21.

Searle, J. R. (1976). A classification of illocutionary acts1. *Language in society, 5*(1), 1-23.

Singer, J. L. (2014). *Daydreaming and Fantasy (Psychology Revivals)*. Routledge.

Shaw, T. (2015). REJECTED TOO MANY TIMES. *Mr. Talented.* Retrieved from https://mrtalented.wordpress.com/2015/06/18/rejected-too-many-times/

Shuy, R. (1993). *Language crimes: The use and abuse of language evidence in the courtroom.* Cambridge, MA: Blackwell.

Van Dijk, T. A. (1977). Semantic macro-structures and knowledge frames in discourse comprehension. *Cognitive processes in comprehension*, 3-32.

Marlon Hurt is a PhD candidate with Aston University's Forensic Linguistics research program, and an Adjunct Instructor of Linguistics at Hofstra University. In July 2017, he presented portions of this work at the 13th Biennial Conference of the International Association

of Forensic Linguistics. Further findings were presented to the 45th International Systemic Functional Congress in July, 2018. In addition to academic pursuits, he has worked in a support capacity on active legal cases with the Institute for Forensic Linguistics, Threat Assessment and Strategic Analysis on an as-needed basis.

Prof. Tim Grant is the Director of the Centre for Forensic Linguistics and Aston University's 50th Anniversary Chair in Forensic Linguistics. He publishes mostly in the area of forensic authorship analysis and most recently has been researching how forensic linguistic techniques can assist in darkweb investigations. As a practitioner, his casework has helped resolve numerous cases involving murder, sexual crime and terrorism. He has appeared in the press, on television and on radio programmes in the UK and internationally, including appearances on BBC Crimewatch and BBC Radio 4 Word of Mouth.