Grief-Stricken in a Crowd: The Language of Bereavement and Distress in Social Media

Jed R. Brubaker[1], Funda Kivran-Swaine[2], Lee Taber[1], and Gillian R. Hayes[1]

[1]School of Information and Computer Sciences
University of California, Irvine
Irvine, CA, USA
{jed.brubaker, ltaber, hayesg}@uci.edu

[2]School of Communication and Information
Rutgers University
New Brunswick, NJ, USA
funda@rutgers.edu

Abstract

People turn to social media to express their emotions surrounding major life events. Death of a loved one is one scenario in which people share their feelings in the semi-public space of social networking sites. In this paper, we present the results of a two-part investigation of grief and distress in the context of messages posted to the profiles of deceased MySpace users. We present coding system for identifying emotion distressed content, followed by a detailed analysis of language use that lays a foundation for natural language processing (NLP) tasks, such as automatic detection of bereavement-related distress. Our findings suggest that in addition to words bearing positive or negative sentiment, linguistic style can be an indicator of messages that demonstrate distress in the space of post-mortem social media content. These results contribute to research in computational linguistics by identifying linguistic features that can be used for automatic classification as well as to research on death and bereavement by enumerating attributes of distressed self-expression in a post-mortem context.

Introduction

Social media-based interactions are an increasingly important to many people's social lives. Life's significant events, from birthdays to marriage proposals, are now experienced in part via social network sites (SNS). This remains true of when those lives come to an end. Grieving a loved one's death—once limited to the physical world—now extends to the social networking profiles of the deceased. Following a user's death, friends often express their shock and grief on the deceased's profile page (Brubaker & Vertesi 2010; Getty *et al*. 2011). Over time, survivors continue to return to social media and keep those who have passed close through continued updates on events and feelings that would have been shared with the deceased were they alive (Brubaker & Hayes 2011).

Analyzing content from sites such as MySpace provides insight into the role that social media plays in contemporary experiences of death, as well as the impact of death on these digital spaces. So far, much of the previous work in the area of post-mortem interactions in social media has been qualitative in approach, employing content and thematic analyses to understand grieving practices. However, quantitative analyses of the content and linguistic style of post-mortem comments are crucial in understanding practices during bereavement. Moreover, user-generated content posted via social media provides a unique opportunity to observe grieving practices in a naturalistic setting. Messages in SNS are in their intended social environments as opposed to interview or lab settings, where expressions of grief have historically been studied.

This work focuses on extreme expressions of grief and mourning in SNS following the death of a friend or loved one. In the first part of this paper we present a codebook for identifying distressed comments based on a thorough review of the literature on grieving practices and bereavement. This coding scheme identifies comments exhibiting emotional distress as a potential indicator of unhealthy mourning and/or the need for additional support. Based on comments coded through this scheme, in the second part of this paper we present the results of a detailed linguistic analysis of comments posted to deceased MySpace profiles in order build explanatory models based on linguistic features that contrast comments exhibiting emotional distress with those that display more conventional funerary language. To this end, we examined 2213 post-mortem comments posted to the profiles of 652 MySpace users following their deaths.

The detailed linguistic analysis of distressed comments presented here expands our current knowledge around the use of language in online grieving. Existing research in this space has predominantly focused on the benefits of such spaces that enable individuals to express their grief. A review of content, however, reveals that some of the

bereaved could benefit from additional social or clinical support. Likewise, the continued presence of the dead amidst lists of friends appears to result in prolonged distress (Brubaker & Hayes 2010). Our code system identifies MySpace comments demonstrating such distress while our computational investigation of language indicates which linguistic features are most informative for identifying emotional distress in this context. As such, this work contributes to enabling future research in natural language processing (such as automatic detection of emotional distress in social media). Detection of comments expressing emotional distress can be beneficial for directing appropriate support during times of social and psychological hardship.

Theoretical Background

Death in the context of social media

Memorializing the dead online has been a practice since the early days of the Internet. Initially memorials consisted of individually maintained webpages and later "virtual cemeteries" hosting "cyber memorials" that allowed visitors to post messages (Roberts & Vidal 2000). With the rise of social media, users also started reappropriating the profile pages of deceased users (Brubaker & Hayes 2010) and creating SNS groups to memorialize the deceased (DeGroot 2008; Marwick & Ellison unpublished).

As Brubaker & Hayes (2011) note, however, SNS profile pages are distinct from other memorial spaces as they are created by the deceased while they are alive, and not by friends or family post-mortem. Following a user's death, profiles are flooded with comments posted by friends of the deceased, expressing shock and communicating their remorse. A number of scholars have noted that commenters rarely interact with each other on the deceased's profile page and instead speak directly to the deceased (DeGroot 2008; Dobler 2009). Brubaker & Hayes (2011) argue that this tendency, in conjunction with the continued presence of deceased friends and their profile pages in SNS, provides the ability for the bereaved to engage in "post-mortem social networking" as a way of maintaining connections with the deceased.

In a classification of 200 MySpace comments, Carroll & Landry (2010) found similar patterns showing that users turn to post-mortem profiles to write expressions of grief and admiration, as well as share memories of the deceased. Meanwhile, Getty *et al*. (2011) performed a linguistic study of the wall posts of memorialized Facebook profiles and found that that post-mortem Wall posts show higher rates of negative emotion than pre-mortem equivalents.

Across the existing research, the ability for the bereaved to visit post-mortem profiles is generally regarded positively as it provides a space for "disenfranchised grievers" (Doka 1989) to mourn, such as friends (Skylar & Hartley 1990) or others who are not family or members of the immediate community. However, these spaces can also be sources of pain. When discussing comments that serve to maintain connections with the deceased, Brubaker & Hayes (2010) noted that some authors repeatedly post distressed comments over long periods of time. The continued presence of the deceased in an individual's social network may have ramifications given that the survivor must make the active choice to remove the deceased from their network (Brubaker, Hayes, & Dourish in press).

Grief & Mourning

While there are many definitions of both grief and mourning (*e.g.*, Kastenbaum 1979; Klass *et al*. 1996; Kubler-Ross 2005), Rando (1993) provides a useful delineation between these two terms with definitions that focus on the potential psychological states and complications experienced by the bereaved. *Mourning* refers to the overall process through which an individual comes to acknowledge and accommodate the loss, while *acute grief reactions* are defined as psychological, behavioral, social, and physical reactions to the perception of that loss. Thus, while mourning is an ongoing process, posting a comment to a friend's MySpace page is an acute event serving as a manifestation of grief.

Current approaches to death and dying generally agree that the bereaved do not follow a uniform or finite grieving process. As such, there is no ideal duration for mourning or way in which to grieve. Instead, the objective of mourning is for the bereaved to incorporate the reality of the death into their relationship with, and understanding of, the deceased (Rando 1993). This objective is held while acknowledging that the bereaved often continue to have bonds with the deceased (Klass *et al*. 1996).

To this end, researchers and clinicians have created a variety of frameworks that enumerate objectives for the bereaved. Rando, for example, outlines six requirements for healthy mourning: recognition of the loss; reaction to the separation; recollections of the deceased; relinquishing of attachments to the deceased; readjustment to the world without the deceased; and a reinvestment of energy into other people, activities, etc. In Rando's framework, these requirements assist the bereaved in reorienting themselves to a reality that incorporates their loss.

Studying the writing of the bereaved

Previous research has found that writing to and about the deceased allows the bereaved to express their grief and gain perspective on their loss (see Roberts & Vidal 2000 for review). Emotional disclosure via writing following the

love of a loved one can have positive effects, such as successful coping (Cable 1996). Translating traumatic events such as death into written language can help individuals encode and structure their experiences in an organized way (Pennebaker *et al*. 1997). These benefits are particularly apparent when the death is sudden or tragic (Roberts & Vidal 2000) as if often the case with social media where, until recently, users have been young. However, as researchers have had little access to personal writing, these claims tend to be based on historical journals, content produced for public consumption, and self-reflection and survey studies (Roberts & Vidal 2000). In contrast, SNS provide a naturalistic setting to observe the writing of the bereaved. Messages in SNS are in their intended social environments as opposed to interview or lab settings, where expressions of grief were most frequently studied in past.

Data

The data used for the work we present in this paper are a subset of a larger dataset of comments collected from profiles of 1369 deceased MySpace users in April 2010. Deceased profiles were identified using MyDeathSpace (MDS)[1], a website dedicated to connecting obituaries and/or news of deaths to existing MySpace profiles. MDS contains more than 15,000 user-submitted entries, as well as comments on individual entries, including additional information and links to other online content (*e.g.*, newspaper articles). The deceased status of each profile owner was verified by one of the researchers based on comments posted by friends or profile content.

For this study, we limited our sample to comments posted to profiles of users who lived in the United States and who had been dead for at least three years. Additionally, only comments from publicly visible profiles were collected. Publicly visible profiles made up the majority of profiles given the lack of privacy features on MySpace during the timeframe in which profile owners died.

Finally, two types of profiles that are substantially different from typical profiles were omitted from our sample: those belonging to celebrities (*e.g.*, Elvis) and profiles that had been created as or repurposed into "issue-profiles" by survivors of the deceased (*e.g.*, those focused on issues such as substance abuse or war rather than on a specific individual who had died from causes related to those issues).

The subset used for this study consists of 2213 randomly selected comments from a total dataset of 206,068 comments. Comments in this subset were posted to 652 profiles by 1660 unique comment authors. The deceased, on average, died at age 27.68 (SD=13.06) and 66.7% of the deceased users were male.

[1] www.mydeathspace.com

Coding for Emotional Distress

We generated a codebook for manual coding of comments exhibiting emotional distress on post-mortem MySpace profiles following a review of literature on expressions of grief and sorrow during mourning and bereavement. The development of the codebook began with an acknowledgement that much of the language in post-mortem comments is performative and uses formulaic language. Thus, we distinguished between common funerary sentiments (*e.g.,* "I miss you already...") and those expressing what we ultimately labeled as emotional distress (ED).

Two authors initially identified twenty extreme comments that unambiguously were written by an author expressing distress, such as the following:

> *i cant stop thinking about u lately .. its been crazy especially friday considering it was ur 100 day.. ppl keep telling me to stop counting the days but i wont.. its the only way i feel close to you ... ah i've been so confused lately over everything even the smallest things.. like the whole time i was i the hospital alls i could do was think of you nd lately thats all i can do.. its crazy er i hate this so much i miss you like crazy u were so easy to talk to its like no matter wut u always understood.. idk w.e. i just needed to "talk" to you i guess.. love you babii*

Using a grounded approach, we open coded comments to identify specific features that distinguished them from non-ED comments. These features were then used to develop tentative rules that specify what does and does not qualify a comment as ED. For example, "Sharing Memories of the Deceased" was a common feature of ED comments, but not unique to them, resulting in rule #3: "*Memories of the deceased in and of themselves do not constitute ED*."

Codebook Rules

The codebook started by defining ED as negative emotions, specifically pain and sadness associated with the deceased. Language associated with ED involves words like hurt, pain, sorrow, and metaphors related to these words (*e.g.,* "my heart has been broken", "my life has been shattered", etc.).

The codebook included six rules to aid in distinguishing between comments, each including example comments that did and did not qualify as ED based on that rule (see Table 1 for summary). The codebook was conservative in attributing ED to comments, with Rule #1 establishing

Rule	Example
Rule 1 *Comments that are ambiguous are to be coded as not ED*	*"Hope you are having a good time up there… miss you forever… "* **(Not ED)**
Rule 2 *Ignore comment content without any emotional component*	*"Hey just had my first day of high school It was ok Still miss you "* **(First two sentences ignored.)**
Rule 3 *Memories of the deceased in and of themselves do not constitute ED*	*"I remember how great you were at ice hockey Whenever I play I know you are watching over me We sure are going to miss you "* **(First sentence disqualified.)**
Rule 4 *Common funeral styled and prosaic content does not qualify as ED*	*"I love you and thinking about you all the time We will never forget you "* **(All statements disqualified.)**
Rule 5 *Comments that include distressed language but that that is positively reframe the loss do not qualify as ED*	*"Thanks to your death I've gained a new appreciation for life I tried to live every day to the fullest like you always did Sometimes it still hurts to think of you and I know I will always miss you but I am still glad that I was fortunate enough to be in your life "* **(Not ED)**
Rule 6 *Comments that include references to ED in the past but do not include subsequent positive reframes are to be coded as ED*	*"This is the third anniversary of that terrible night I miss you more than ever we all do… for a while we could not even speak your name You are s0 loved and you will always be missed mom"* **(ED)**

Table 1 Summary of codebook rules for classifying post mortem comments as expressing emotional distress (ED)

non-ED as the default code: *"Comments that are ambiguous are to be coded as not ED."* For example, simple comments such as *"Miss you!"* were not coded as ED, while those unambiguously indicating distress were:

> *Happy 22nd Birthday Baby!! There are no words to describe how much I miss you, this day this weekend nothing is the same. The world for me has lost it's magic, it's beauty. I love you Matt-Matt there is not one single minute that passes that I am not thinking of you. Momma loves you baby!!!!!*

Post-mortem comments often include non-emotive content, such as memories or personal updates. Rule #2 instructed coders to ignore content that does not have an emotional component, while rules #3 and #4 stated that memories of the deceased and common funerary styled content (*e.g., "We will never forget you.")* do not, in and of themselves, constitute ED. Rather, coders were instructed to focus on the interpretation and context of the content. Under these rules, the following comment would qualify as ED:

> *I miss you so much every day. I can barely live my life without you, I know I will miss you til the day I die*

Whereas a similar comment would not:

> *We miss you a lot, but I know you are watching over us from heaven. Love you!*

Finally, our coding was also sensitive to the ways in which the loss of a loved one can be reframed over time. Rule #5 specifies that if the comment includes content about how the deceased or the experience of their death enriched the author's life, it is unlikely that the comment is ED. For example, the following would not be coded as ED:

> *Thanks to your death I've gained a new appreciation for life. I tried to live every day to the fullest, like you always did. Sometimes it still hurts to think of you, and I know I will always miss you, but I am still glad that I was fortunate enough to be in your life*

While this comment would qualify as ED:

> *Every day has been terrible since you've been gone. I feel like anything could just push me over the edge. I hold the memories of your smiles dear and I try to make my own like I know you would want me to, but I still can't think of you without the pain.*

Finally, rule #6 specifies that comments discussing times in the past when the author was distressed, but that do not positively reframe the loss, can be coded as ED (subject to other rules). For example, the following message, while discussing the past, provides no subsequent reframe, and as such would be coded as ED:

> *When we were at your casket, Mom and I had to hold each other up. We cried so hard during most of the wake, and its still painful to think about. I love you and miss you very much.*

Throughout, coders were instructed to weigh emotional content and give preference to the more potent and/or substantial aspects of the message.

Reliability Assessment

To assess inter-rater reliability, two of the authors independently coded 100 comments randomly selected from 10 deceased individuals' profiles. Inter-rater reliability on this coding was good with an agreement of 90% and kappa of .71. Next, 20 comments were coded by eight coders with 84.4% agreement. Upon the achievement of good inter-rater reliability, 2113 more comments were coded as ED or non- ED by one author. Our conservative definition resulted in low rates of occurrence with 12.5% of the comments including expressions of ED.

The Language of Emotional Distress

Having established a reliable way of distinguishing comments expressing emotional distress, our next objective was to determine whether linguistic patterns

underlie expressions of emotional distress. In this part of our study, we aimed to find and articulate the linguistic style of emotional distress.

Linguistic style, how language is used by individuals or groups of people (Eckert 1996), provides information about the characteristics of the people using language as well as the conditions of their social surroundings. One way to quantitatively study linguistic style is to observe the use of linguistic features, such as parts of speech. This type of analysis focuses less on context and more on how language use varies between groups (see Pennebaker, Mehl, & Niederhoffer 2002 for review).

Previous research in both computer-mediated communication and traditional settings indicates that mental and physical well-being exhibit strong relationships with language use. Particularly, associations between certain parts of speech and emotional state have been observed in several previous studies. As reviewed by Pennebaker and his colleagues (*ibid*.), studies have reported that use of first person singular pronouns were elevated in the language use of individuals exhibiting signs of depression and suicide ideation. Researchers interpreted these heightened instances of pronoun use as high degree of self-preoccupation.

Patterns of word use bearing positive or negative sentiment, meanwhile, were also found to relate to health outcomes in the context of bereavement (Pennebaker *et al.* 1997). Specifically, higher rates of use of positive emotion words were linked to better physical and mental health.

Analysis of Language Use

To capture linguistic differences between ED and non-ED comments, we computed variables encapsulating the frequency of use of linguistic style features and sentiment expression.

In order to analyze the sentiment content as well and linguistic style of the comments in our dataset, we used "Language Inquiry Word Count" (LIWC)[2], a common language analysis package that provides dictionaries for parts of speech and punctuation, as well as psychological and social processes. LIWC has previously been used for analyzing social media content (*e.g.*, Getty *et al*. 2011; Kivran-Swaine & Naaman 2011). For each body of text analyzed, LIWC provides a score between 1 and 100 per category, specifying the proportion of words in the text contained in the dictionary for that category.

The *linguistic style* features we examined were pronouns, conjunctions, negations, adverbs, and tense use. Heightened use of first person pronouns and decreased use of second and third person pronouns have previously been found to relate to depression (Pennebaker, Mehl, & Niederhoffer 2002). Additionally, previous research suggests that use of first person singular pronouns elevate during emotional upheavals (*ibid*.). High use of adverbs and negations is an indicator of female style language, which is more socially aware (Mulac *et al*. 1998), while use of prepositions and conjunctions suggests a common frame of reference between the conversing parties. As such, uses of these parts of speech can be expected to elevate in more informal and intimate settings (Pennebaker, Mehl, & Niederhoffer 2002). Finally, tense use has been found to be an indicator of informality as well as need states in previous work (*ibid*.).

To assess *expression of sentiment*, we used scores from the following categories: positive emotion, sadness, anger, and social processes. The category, *positive emotion*, has been found to provide a basis for detecting expression of joy in social media content (Kivran-Swaine & Naaman, 2011), and includes words such as "happy", "love", "nice". *Sadness*, meanwhile, contains words such as "cry", "sad", and "grief", which all relate to the emotional state, sadness. Similarly, the category, *anger* contains words related to the emotional process of anger and aggression, such as "hate" and "kill". Finally, the *social processes* category contains social verbs such as "talk" as well as references to social relationships such as "family" and "friend."

For each comment in our dataset, we computed the LIWC scores for the following categories: word count, pronouns (first person singular, first person plural, second person, third person singular, third person plural), conjunctions, negations, adverbs, and tense use (present, future, and past tense), positive emotion, sadness, anger, and social processes.

Method

The unit of analysis in this work is a single comment. We first conducted multiple independent samples T tests to examine whether linguistic style and expression of sentiment differ between post-mortem comments that were coded ED or non-ED. The independent variables for the T-tests were the computed variables for each linguistic and sentiment category, and the grouping factor was the binary variable *ed*, (whether a comment was coded as ED). To check for the assumption of homogeneity of variances, we conducted Levene's tests. To account for our practice of conducting repeated tests looking at the mean differences between the same groups, we used the Bonferroni correction, adjusting significance levels to α/n when conducting n tests simultaneously.

Although independent-samples T tests give us information about the significance of mean differences, they do not account for higher order interactions between multiple language variables. Therefore, as the second step

[2] http://www.liwc net/liwcdescription.php

of our analysis, we constructed two binary logistic regression models. The first model, Model 1, uses measures of linguistic style variables while Model 2 uses measures of expressions of sentiment as independent variables. For both models, the binary variable *ed* is the response variable.

Results

The results from the independent samples T-Test indicate that linguistic style can be a strong marker for expression of ED. T-values, significance values (after the Bonferroni correction), and group-level descriptive statistics for the T-tests can be seen in Table 2. The odds ratios, significance levels, β values as well as standardized errors for the variables in binary logistic regression models can be seen in Table 3.

Greater use of first person singular pronouns, past tense verbs, adverbs, prepositions, conjunctions, and negations, as well as higher word count per comment were observed in ED comments; non-ED comments demonstrated higher levels of use of second person pronouns. Significant differences in sentiment content were also observed between ED and non-ED comments: non-ED comments showing higher use of social process words and words expressing positive emotion and sadness, and ED comments with higher use of anger words.

Model 1, explaining the tendency of a comment to be ED by looking at linguistic style variables, showed that higher word count, as well as higher use of first person singular pronouns, adverbs, conjunctions, and negations increase the likelihood of a comment to be classified as ED, whereas higher use of second person pronouns, third person singular and plural pronouns, and future tense decrease that likelihood. Model 2, which takes content features into account, showed that higher levels of use of positive emotion and sadness words decrease the likelihood of a comment to be ED, whereas higher use of anger words increases that likelihood.

Discussion

The results of the analysis of language use show that the sentiment and linguistic style of comments exhibiting ED differ significantly from other post-mortem comments.

ED comments on average are longer than non-ED comments; nearly double the length of non-ED comments. This difference may indicate that those experiencing ED spend more time constructing messages and in working through their feelings through writing. Word length for post-mortem writing is also influenced by the system into which it is posted. For example, Getty *et al*. (2011) reported an average of 39.9 words per message in a sample from Facebook, while Roberts & Vidal's (2000) analysis

Variable	t	df	p-value	d
Word Count	10 94	306 96	***	0 81
F P S Pronouns	9 05	569 37	***	0 45
S P Pronouns	5 08	645 75	***	0 24
Past Tense	4 73	461 57	***	0 25
Adverbs	8 61	493 53	***	0 46
Prepositions	6 6	533 25	***	0 34
Conjunctions	10 03	442 54	***	0 56
Negations	6 75	419 46	***	0 39
Social Proc	6 6	705 98	***	0 31
Pos Emotion	10 76	953 08	***	0 46
Anger	3 74	400 47	**	0 23
Sadness	3 6	1053 42	**	0 15
	ED Comments		**Non-ED Comments**	
	Mean	S D	Mean	S D
Word Count	122 4	102 3	53 57	62 03
F P S Pronouns	9 95	3 52	7 65	6 26
S P Pronouns	6 62	3 79	8 06	7 48
Past Tense	4 21	2 62	3 37	3 78
Adverbs	7 29	3 28	5 33	5 09
Prepositions	8 77	3 24	7 24	5 44
Conjunctions	6 59	3 01	4 55	4 13
Negations	1 88	1 68	1 13	2 15
Social Proc	14 17	5 45	16 94	11 54
Pos Emotion	4 19	2 76	6 7	7 2
Anger	0 46	0 87	0 24	1 04
Sadness	2 55	2 08	3 2	5 79
p< 1 * p< 05 ** p< 01 *** p< 001 **** p< 0001				

Table 2 ndependent Samples T Test Results and Descriptive Statistics

Variable	b	S.E.	p-value	O.R.
Model 1				
ntercept	4	0 32	****	
Word Count	0 01	0	****	1 01
F P S Pronouns	0 08	0 01	****	1 09
S P Pronouns	0 04	0 01	***	0 96
T P S Pronouns	0 14	0 07		0 87
T P P Pronouns	0 35	0 17	**	0 7
Future Tense	0 11	0 05	**	0 9
Adverbs	0 06	0 02	****	1 06
Conjunctions	0 08	0 02	****	1 08
Negations	0 15	0 03	****	1 16
Model 2				
ntercept	1 28	0 13	****	
Pos Emotion	0 09	0 02	****	0 91
Sadness	0 03	0 01	**	0 96
Anger	0 13	0 05	**	1 14
p< 1 * p< 05 ** p< 01 *** p< 001 **** p< 0001				

Table 3 Details for Binary Logistic Regression Models

of cybermemorials found a word count similar to the ones we report here.

Understandably, ED comments express less positive emotion than non-ED post-mortem comments. That ED comments include significantly less positive emotion words than non-ED comments may suggest that those expressing ED are more likely to be in a depressed state. This is consistent with previous research that looked at linguistic signs of depression (Pennebaker, Mehl, & Niederhoffer 2002). ED comments also use more words expressing anger. This suggests that for people experiencing ED, post-mortem profiles are providing a space for venting frustration, in contrast to the more traditional funerary language employed by non-ED comments.

One might imagine that to detect individuals on SNS who are struggling to cope with the loss of a friend, sentiment analysis would be a sufficient approach. Our results suggest otherwise. Post-mortem spaces are sites of social performance in which individuals (distressed or not) utilize highly emotive language, making sentiment analysis alone insufficient. Our findings suggest, however, that comments expressing emotional distress can be identified through the way in which sentiment is expressed.

Linguistic features demonstrate promise in differentiating ED comments from the rest of the post-mortem comments. People expressing ED showed higher use of first person singular pronouns and lower use of second person pronouns. This indicates that, overall, comments containing declarations of ED portray a linguistic style indicative of self-focus and isolation, deviating from the socially aware linguistic style of other post-mortem comments. Meanwhile, the decreased use of second person pronouns in comments expressing ED can be a sign that, for these authors, the comment space serves more as an environment for conveying individual yearning or pain, rather than continuing bonds by addressing the deceased.

The increased use of past tense verbs in distressed comments may suggest that looking back and reflecting on the past may be a form of coping for those who are experiencing intense episodes of sorrow, post-mortem. In a previous study, Pennebaker, Mayne & Francis (1997) found that decreased use of past tense words is one of the predictors of perceived emotional distress for those who are bereaving the loss of a life partner. The fact that in this study we see those expressing ED using higher rates of past tense words may suggest that those in distress are in fact using this space to help with their coping. Finally, future tense use contributed negatively to Model 1. It is possible that during ED, people are less inclined to talk about the future and are more focused on their past or present experiences.

ED comments score higher on use of adverbs and negations. Use of adverbs, specifically intensifier adverbs, which dominate the adverb dictionary of LIWC, have been found to be indicators of "female-style" language, in previous research (Mulac *et al*. 1998). Use of conjunctions and negations in language are also potential signs of intimacy between the speaker and the addressee. All of these parts of speech show higher occurrences in ED comments suggesting that ED comments possess more signs of intimacy than non-ED comments, which use more funerary-styled language.

Conclusion & Future Work

In this paper, we reported on the development and implementation of a coding system for classifying post-mortem SNS comments as expressing emotional distress (ED) or not. Using this coding system, we then conducted a thorough linguistic analysis of comments posted on MySpace profiles belonging to deceased users to identify differences between ED and non-ED comments with respect to linguistic style and sentiment content. We noted that significant linguistic style differences exist between ED and non-ED comments.

To account for possible higher order interactions between variables, we constructed two binary logistic regression models, identifying linguistic style and content features that are strong predictors for a comment being ED. The regression models show that within the linguistic style features we considered, levels of negations, pronouns, and adverbs show potential to be strong markers for detecting emotional distress within the context of grieving in social media outlets.

The major contributions of this work are in defining a systematic way to identify and classifying content that exhibits grief-related emotional distress, as well as laying the groundwork for characterizing the nature of comments in post-mortem spaces based on language use. Moreover, the results from this work provide indications of a relationship between language and psychological well-being, as observed in natural public environments.

This works marks an initial step towards using content in SNSs to identify emotionally distressed at a large scale. As such, this could result in tools to classify social media content similar to NLP-based work on classifying suicide letters (Pestian *et al*. 2010).

Identifying users who are struggling with the loss of a loved one can also result in development of tools or programs to intervene. Such interventions are already beginning to appear. For example, Facebook recently added functionality that allows users to report content generated by their friends as potentially suicidal, resulting in an e-mail sent to the content author with support

resources. It is not hard to imagine that users experiencing ED following their friends' passing might also benefit from similar support systems. Moreover by identifying users experiencing ED, system designers can create web experiences sensitive to the emotional states of mourners.

This study also contributes to our understanding of the writing of the bereaved in social media. While most individuals author non-distressed content, often utilizing funerary-styled language, individuals expressing emotional distress write in ways indicative of self-focus and isolation, despite the public nature of the SNS. This may suggest that even while SNS can benefit disenfranchised grievers by allowing them to participate in a space of communal mourning (Carroll & Landry 2010), the public nature of the SNS alone may be insufficient for distressed individuals whose language suggests they remain isolated.

As an initial step at analyzing the language in post-mortem spaces in social media, our study has limitations. First, the data we used for both the development of the coding system and the language analysis consisted of content posted to MySpace between 2006 and 2007. Analysis of similar but more recent content from additional sites such as Facebook is necessary. Second, we limited the analysis of language use to exploratory models. Future research should build on this work to develop and test predictive models, possibly on other sources of post-mortem content. Additionally, the coding system for classifying emotional distress is not comprehensive; it is limited to the content that people post. For example, the pain associated with losing a loved one may prevent some users from interacting with a post-mortem profile entirely. Finally, LIWC, while a frequently used tool for language analysis has its weaknesses. Because it is a dictionary-based tool, it fails to capture contextual differences, and certain subtleties of language such as sarcasm.

While it can be a sensitive topic, death and mourning are unavoidable parts of life. As such, studying grief in the context of SNS is valuable. As the adoption of social media continues to grow, grief and mourning will increasingly become important aspects of our social experiences online.

References

Brubaker, J. R. and Hayes, G. R. 2011. "We will never forget you [online]": An empirical investigation of post- mortem MySpace comments. In Proceedings of the 14th Conference on Computer Supported Cooperative Work. 123-132. Hangzhou, China.

Brubaker, J. R., Hayes, G. R., and Dourish, P. In Press. Beyond the Grave: Interpretation and participation in peri-mortem behavior on Facebook. *The Information Society*.

Brubaker, J. R. and Vertesi, J. 2010. Death and the Social Network. Presented at the HCI End of Life workshop at the 28th Conference on Human Factors in Computing Systems.

Cable, D. G. 1996. Grief counseling for survivors of traumatic loss. In Doka, K. J., eds., *Living with Grief After Sudden Loss*. Bristol, PA: Taylor & Francis.

Carroll, B., & Landry, K. 2010. Logging On and Letting Out: Using Online Social Networks to Grieve and to Mourn. *Bulletin of Science, Technology & Society* 30(5): 341

Degroot, J. M. 2008. Facebook Memorial Walls and CMC's Effect on the Grieving Process. Paper presented at the annual meeting of the National Communication Association, San Diego, CA.

Dobler, R. 2009. Ghosts in the Machine: Mourning the MySpace Dead. In Blank, T. J., ed., *Folklore and the Internet: Vernacular Expression in a Digital World*. Utah State Univ. Press. 175-194.

Doka, K. J. 1989. *Disenfranchised Grief: Recognizing Hidden Sorrow*. England: Lexington Books.

Eckert, P. 1996. Vowels and nail polish: the emergence of linguistic style in the pre-adolescent heterosexual marketplace. In *Gender and Belief Systems*. Berkeley Woman & Language Group.

Getty, E., Cobb, J., Gabeler, M., Nelson, C., Weng, E. and Hancock, J. 2011. I said your name in an empty room: grieving and continuing bonds on Facebook. In Proceedings of CHI 2011. 997-1000. Vancouver, B.C.

Kastenbaum, R., and Aisenberg, R. 1972. *The psychology of death*. Springer.

Kivran-Swaine, F., Naaman, M. 2011. Network properties and social sharing of emotions in social awareness streams. In Proceedings of CSCW 2011. 379-382. Hangzhou, China.

Klass, D., Silverman, P., and Nickman, S. 1996. *Continuing Bonds: New Understandings of Grief*. Philadelphia: Taylor & Francis.

Kübler-Ross, E., and Kessler, D. 2005. *On grief and grieving: finding the meaning of grief through the five stages of grief*. New York: Simon and Schuster.

Marwick, A., and Ellison, N. Unpublished manuscript. "There Isn't Wifi in Heaven!": Negotiating Visibility on Facebook Memorial Pages.

Mulac, A., Wiemann, J.M., Widenmann, S.J., and Gibson, T.W. 1998. Male/female language differences and effects in same-sex and mixed-sex dyads: The gender-linked language effect. *Communication Monographs* 55: 315-335.

Pennebaker, J.W., Mayne, T.J., and Francis M.E. 1997. Linguistic predictors of adaptive bereavement. *Journal of Personality and Social Psychology* 72: 863- 871.

Pennebaker, J.W., Mehl, M.R., and Niederhoffer, K.G. 2002. Pyschological aspects of natural language use: Our words, our selves. *Annual Review of Psychology* 54: 547-477.

Pestian, J., Nasrallah, H., Matykiewicz, P., Bennett, A., and Leenaars, A. 2010. Suicide Note Classification Using Natural Language Processing: A Content Analysis. *Biomedical Informatics Insights* 3: 19-28.

Rando, T. A. 1993. *Treatment of complicated mourning*. Champaign, IL: Research Press.

Roberts, P., and Vidal, L. 2000. Perpetual care in cyberspace: A portrait of Web memorials. *OMEGA: The Journal of Death and Dying* 40(4): 57-76.

Skylar, F., and Hartley, S. F. 1990. Close Friends as Survivors: Bereavement patterns in a "hidden" population. *OMEGA: The Journal of Death and Dying* 21(2): 103-112.