

AN ABSTRACT OF THE DISSERTATION OF

Kristeen E. Elrod for the degree of Doctor of Philosophy in Counseling presented on September 20, 2019.

Title: A Corpus Linguistic Analysis of Pro-Nonsuicidal Self-Injury and Pro-Anorexia Public Tumblr Posts Which Use Spanish (Un Analisis de Lingüística de Corpus de Publicaciones Pública que usan Español en Tumblr que Respalda a las Pró Autolesiones No Suicidas y Pró-Anorexia)

Abstract approved:

Cass Dykeman

On social media platforms and on blog sites, disordered behavior as a lifestyle choice is a growing trend that supports and promotes behaviors related to self-harm and anorexia. Web content advocating self-harm and anorexia can be found easily through search engines or can be concealed in the social media sphere through deliberately misused language and creative hashtags. Existing scholarly literature includes information regarding the language used on social media platforms and on blogs sites when material is written in English; however, there is little research exploring these phenomena as they are communicated in Spanish.

The present study, which used a synchronic corpus linguistic design, has two arms, both of which examined public Tumblr posts. The first arm focused on pro-nonsuicidal self-injury (NSSI) posts, and the second arm scrutinized pro-anorexia (pro-ana) posts. Posts were scraped when they included Spanish language and/or hashtags and when they also utilized hashtags identifying the post as pro-NSSI or pro-ana. The research questions for the arms were as follows:

First, in public Tumblr posts written in Spanish, what was the use rate of specific words for individuals posting content of a pro-NSSI or pro-ana nature? Second, in public pro-NSSI or pro-ana Tumblr posts written in English, what was the use rate of pronoun categories? Third, in public pro-NSSI or pro-ana Tumblr posts written in English, did the use rate of pronoun categories differ from norms for such posts? And, finally, in public pro-NSSI or pro-ana Tumblr posts written in Spanish, did the use rate of the psychological process categories differ from norms for such posts?

Following an analysis of the data with Linguistic Inquiry and Word Count (LIWC) software, the first arm of the study revealed the use of high levels of negative emotion words in the dominant language. The results also revealed possible depression, perceived lower social status and users' efforts to distance themselves from their emotions when writing in the non-dominant language (English). Overall, these results demonstrated that people posting on Tumblr about NSSI are suffering and in need of help. The results of the second arm of the study revealed that reliance on pronouns increased users' focus on people and objects, which has been linked to decreased likelihood of recovery. The psychological properties of the posts had a negative emotional tone and significant use of ingestion and health words, signaling potential symptoms of anorexia.

Considering the extent of both research arms, the results show that these overlapping populations struggle to manage intense negative emotions. This research points to the benefits both of assessing bilingual clients for NSSI and anorexia symptoms and of evaluating their online posting behaviors. So, too, do these studies underscore the need for improved evidence-based treatment for trauma in bilingual individuals and the importance of devising Spanish and

codeswitching blog norms to inform future research. Finally, these results can have profound implications for clinical treatment by informing the counseling theories used to treat these populations, by transforming how counselors are trained to assess and respond to NSSI and anorexia behaviors, and by investigating how research can reveal the nuances of language such that scholars better appreciate the difficult issues confronting teens and young adults.

©Copyright by Kristeen E. Elrod
September 20, 2019
All Rights Reserved

A Corpus Linguistic Analysis of Pro-Nonsuicidal Self-Injury and Pro-Anorexia Public Tumblr Posts Which Use Spanish (Un Analisis de Lingüística de Corpus de Publicaciones Pública que usan Español en Tumblr que Respaldan a las Pró Autolesiones No Suicidas y Pró-Anorexia)

by
Kristeen E. Elrod

A DISSERTATION

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

Presented September 20, 2019
Commencement June 2020

Doctor of Philosophy dissertation of Kristeen E. Elrod presented on September 20, 2019

APPROVED:

Major Professor, representing Counseling

Dean of the College of Education

Dean of the Graduate School

I understand that my dissertation will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my dissertation to any reader upon request.

Kristeen E. Elrod, Author

ACKNOWLEDGEMENTS

Undertaking and completing this program has been a life-changing experience for me, and it would not have been possible without the guidance and support of many people. First, I would like to extend my deepest gratitude to my advisor, Dr. Cass Dykeman. You have been a bottomless well of positive energy, and you were always there with a backup plan—the two things I needed most. I could not have finished this dissertation without you. I am a better clinician, supervisor, educator, researcher, advocate, and person because of your support.

I am also extremely grateful to my committee: Dr. Gene Eakin, Dr. Deborah Rubel, Dr. Catherine Beckett, and Dr. Richard Nafshun. As outstanding teacher-scholars, you offered me extensive guidance, and I respect and admire your dedication to the profession. Thank you for your time and goodwill throughout this process.

Cohort 63 deserves recognition for everything our group brought to the table. Thanks to each of you for studying with me, challenging me, and inspiring me to do better. I must also acknowledge my friends, family, the gym, my therapist, and other supporters at OSU. I appreciate all the assistance, hand-holding, barbells, and interventions that got me to this point. Thank you for believing in me.

To Ginger, the sweet cat who wanted to type on my computer as much as I did, I appreciate the comfort you offered when I needed it most. And, finally, thank you to my husband, Matt Dunn. You have witnessed many highs and lows throughout this experience, and you always did everything you could to support me.

CONTRIBUTION OF AUTHORS

Cass Dykeman assisted with the methodology, research design, and statistical computation of the data, while also refining the dissertation's prose and narrative structure.

TABLE OF CONTENTS

	<u>Page</u>
Chapter 1: A General Introduction	1
Overview	2
Importance to the Counseling Profession	3
Literature Review.....	4
Description of the Research Manuscripts	11
Thematic Links Between Manuscripts.....	18
Organization of the Dissertation	18
Chapter 2: A Research Manuscript	27
Method	33
Results.....	38
Discussion	39
Chapter 3: A Research Manuscript.....	56
Method	64
Results.....	69
Discussion	70
Chapter 4: A General Conclusion.....	86
Overview	87
Summary of the Research Manuscripts	88
Thematic Links Between Manuscripts and Contributions to the Literature	94
Future Research Agenda.....	97
Bibliography	101
Appendix A: IRB Document 1	114
Appendix B: IRB Document 2.....	115

LIST OF TABLES

<u>Table</u>	<u>Page</u>
1. Nonsuicidal Self-Injury Item Word Count	50
2. English Pronoun Word Count.....	51
3. English Pronoun Test Results	52
4. Spanish Psychological Process Category Test Results	53
5. English Psychological Process Category Test Results	54
6. Mixed Psychological Process Category Test Results	55
1. Anorexia Item Word Count	80
2. English Pronoun Word Count.....	81
3. English Pronoun Test Results	82
4. English Psychological Process Category Test Results	83
5. Spanish Psychological Process Category Test Results	84
6. Mixed Psychological Process Category Test Results	85

Chapter 1: A General Introduction

Overview

Disordered behavior as a lifestyle choice is a growing trend on social media and on blog sites. This social movement supports and promotes behaviors related to self-harm and anorexia, among other destructive behaviors. Content advocating nonsuicidal self-injury (NSSI), referred to as pro-NSSI material, describes, supports, and promotes such behavior as a form of self-help that can regulate one's affect and alleviate one's negative emotions (Thyssen & van Camp, 2014). In a relatedly harmful way, pro-anorexia (pro-ana) content promotes anorexia and anorexic activities as fashionable and beautiful, often including tips and methods for becoming anorexic and maintaining anorexia as a lifestyle choice, while resisting the notion that behavior is associated with mental illness (Syed-Abdul et al, 2013).

Anorexia and NSSI typically begin in adolescence and young adulthood (Klonsky, Victor, & Saffer, 2014). Adolescents and young adults spend an average of nine hours per day between engaging on social media and watching online videos (Tsukayama, 2015), making pro-ana and pro-NSSI content a potential public health concern (Syed-Abdul et al., 2013). Content related to self-harm and anorexia and the continued maintenance of these behaviors can be found easily through search engines or can be concealed in the social media sphere through specific language and hashtag terms.

Existing scholarly literature includes information on the language used on social media and blogs sites when material is presented in English; however, there is little research exploring these phenomena as they occur and arise in Spanish-language forms. Currently, there are more Hispanic users than Caucasian users on Instagram and other social media platforms (Hull, Kim, & Stilwell, 2018), and the population of Spanish-speaking Hispanic individuals in the United States is expected to continue rising in the future. Because of the significance of this topic and

the lack of research on the relationship between its particular elements, the present study evaluated data from pro-NSSI and pro-ana public posts written in Spanish and placed on Tumblr (a popular microblogging site) to identify and analyze the language users relied on to describe their NSSI and anorexia behaviors.

Importance to the Counseling Profession

Adolescents engage with the world and interact through social media. Social networking has become a new way to seek information and exchange user-generated content with others. News, events around the world, art, and so on are accessed and created through social media microblogs like Twitter and through image-sharing platforms such as Tumblr and Instagram. Adolescents use social media to communicate with their friends and family, meet new people, find job opportunities, promote themselves or their interests, and even to make money. Because social media is such a large part of adolescents' lives, mental health clinicians must be aware of the landscape of social media as well as current trends to guide regular discussions in counseling sessions.

Previous studies have looked at language used on social media and in mental health posts and blogs, but comparatively little research has focused on communities interacting in Spanish. Many studies have applied a corpus linguistic analysis approach to social media posts related to depression, anxiety, postpartum depression, and eating disorders, however very few have examined posts drafted in Spanish. Of the work that exists in both English and Spanish, scholars have struggled to access research and to build on previous work, given the variability in terms associated with this growing trend online. The manuscripts in this dissertation explored the precise language of pro-NSSI and pro-ana content, focusing on the psychological markers involved in the language within posts. The results of this work afford a different perspective on

individuals who engage in NSSI and anorexia behaviors, and these findings can help mental health professionals improve treatment protocols.

Literature Review

This chapter provides background for the present study by, first, defining and describing NSSI and anorexia, before offering an overview of pro-NSSI and pro-ana blog posts, related issues over language and bilingualism, NSSI and anorexia in Hispanic communities, the influence of social media and blogs on adolescents, and the use of Linguistic Inquiry and Word Count (LIWC) software as a new method of understanding language online. A summary of the entire dissertation will follow the overview of these concepts, terms, and concerns.

There are varying definitions and terms to describe NSSI; Favazza (1998) defined it as “the deliberate, non-suicidal destruction of one’s own body tissue” (p. 260), a description that is widely cited and accepted around the world. Favazza’s differentiation of nonsuicidal intent has become the fundamental distinction between self-harm and suicidal behaviors. Within this research, actions classified as NSSI behaviors include cutting, burning, scratching, and biting, although behaviors can also reach as far as scarring, ingesting poisons, breaking bones, picking, and more (Thyssen & van Camp, 2014; Whitlock, Powers, & Eckenrode, 2006). Terms used to describe NSSI in the literature vary greatly and across languages, making it difficult to access prior works in the field and, consequently, limiting the dissemination of literature written in Spanish and Portuguese (Thyssen & van Camp, 2014). Because NSSI is identified as an area in need of further research in the fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (American Psychiatric Association, 2013), it is imperative that researchers devise and implement more united and easily accessible means of accessing information.

Studies show that NSSI usually begins in early adolescence, with onset occurring at around 12 years of age. The prevalence rates of NSSI among adolescents range from a reported 7.5% to 46.5% in English-language articles (Cipriano, Cella, & Cotrufo, 2017) and from 13% to 45% in literature from other parts of the world (Vega et al., 2018). The most common function of NSSI is reported to be the regulation of affect, where negative emotions (anger, anxiety, depression, etc.) tend to occur before the self-injury and where an increase in positive emotions and a decrease in negative emotions occur after.

Not surprisingly, many young people managing such emotions look for outlets for their feelings. Pro-NSSI sites and blogs are online communities where those who engage in NSSI behavior go to share their experiences and engage with others who also self-harm (Lewis, Heath, Michal, & Duggan, 2012). These sites and blogs can trigger urges to engage in self-harm, and they can also normalize self-harming behavior. Social reinforcement can encourage members to post more self-harming content or to escalate their self-harming, because, sadly, more severe wounds tend to elicit more comments and “likes” (Brown et al, 2018). Members of these sites view self-harm as an acceptable way to manage and express emotions and to provide one another with support on how to keep these acts hidden from others (i.e., where to inflict harm on the body and how best to hide marks).

Like the urge to self-harm, an eating disorder is a complex condition that can lead to serious lifelong health concerns. Indeed, anorexia claims the highest mortality rate of any psychiatric disorder (Arcelus, Mitchell, Wales, & Nielson, 2011). Anorexia is characterized by a significantly low body weight, fear of gaining weight or “becoming fat,” a tendency to engage in behaviors that interfere with weight gain (extreme dieting, starvation, etc.), and a disturbance in the way individuals perceive their body weight or the way their shape is experienced (American

Psychiatric Association, 2013; Rodrigues, 2017). Individuals suffering from anorexia often disguise and hide their symptoms, making this high-risk population difficult to reach.

The age of onset for anorexia is early adolescence, from 13 to 18 years old, with prevalence rates at about .5% to 2% of the population. Demographics have been shifting and are showing higher rates of eating disorders being diagnosed in younger children, boys, and minority groups (Campbell & Peebles, 2014). Medical complications are common for full and sub-clinical cases, because anorexia affects every organ system in the body. Due to the damage that anorexia behaviors can have on the body, they are often considered a kind of self-harming behavior.

Pro-ana is an online social movement that honors anorexia behaviors as a lifestyle and that promotes such behaviors as an effective way to diet, despite acknowledged risks to health (Castro & Osorio, 2013). These online communities are popular because they afford individuals suffering from eating disorders a safe and neutral space where they can find relief and express their feelings. Typical content found on these sites and blogs includes information about diets, information about drugs, tips and tricks to avoid hunger (e.g., drink lots of water, chew gum, smoke cigarettes, or sleep), exercise routines, photos with proof of progress, words or images offering “thinspiration,” and so on.

Research has focused on anorexia-related behaviors and concerns in English-speaking communities, but scholars know comparatively little about pro-ana content written in Spanish. With the increasing size of the Hispanic population, Spanish has become the most common non-English language spoken in the home, reaching a total of 37.6 million people over age 5 (PEW Research Center, 2013). Moreover, bilingualism is common in the United States (Perry, 2013); even more, as Santiago-Rivera and Altarriba (2002) have suggested, bilingual individuals represent emotions differently between their two languages, with emotion words stored at a

deeper level of representation. Emotion words in the first language have been experienced in more contexts and applied in a variety of ways, but it is important to stress that bilingual individuals also demonstrate more emotional expressiveness when speaking in their primary language (Ramos-Sanchez, 2007). Language switching, or *code-switching*, can indicate distancing from emotional experiences (Fuertes, 2004).

For the purposes of the present study, social media presents the unique opportunity to “manage the terms of interpersonal linguistic engagement with others” (Baron, 2008, p.84), affording users the potential to connect with other speakers and to shape concepts of self and self-perceptions. Importantly, language use impacts levels of acculturation and social bonds; it is an important means of transmitting information regarding beliefs and cultural traditions. So, too, does it convey an individual's heritage and serve as a source of identity and pride. Emotion-laden words or experiences are usually better expressed in an individual's dominant or native language, while one's non-dominant language can be used to express thoughts or beliefs not associated with emotions. Note, as well, that Gutfreund (1990) found that counseling participants expressed more affect in Spanish than they did in the English treatment condition, regardless of whether Spanish was their dominant language.

Bringing the elements of online exposure and language together, research has shown that 88% of bilingual young adults reported some use of Spanish on social media (Velazquez, 2017). Spanish is a global language, but it has a status as a minority language in the United States. There has been limited research on psychological and social issues contributing to NSSI in the case of Hispanic individuals (Croyle, 2007; Gulbas et al., 2015). Of the research that does exist, several common environmental factors that contribute to NSSI among Hispanic adolescents have been identified, including fragmented family structures, family conflict, parental criticism,

transnational stress, acculturation, bullying, experiences of violence, and academic challenges (Gulbas, Hausmann-Stabile, De Luca, Tyler, & Zayas, 2015). For those engaging in self-harm, family conflicts often center around different perspectives on autonomy, symbolizing a cultural or generational gap that makes connections with parents difficult.

As Gulbas et al (2015) demonstrated, among those with no history of self-harm, adolescents were more able to empathize with their parents' point of view and were more cognizant of social and economic factors impacting their parents. It is also the case that Hispanics with a self-harm history often feel emotions as more immediate and overwhelming and typically have more destabilized relationships with parents and others (Gulbas et al., 2015). Acculturation issues, particularly within the family, can predict self-harm among Hispanic females, as can immigration stress and separation, suggesting that family is a powerful factor in this population (Cervantes et al., 2014). Such life events induce powerful affective states, and self-harming adolescents believe they can control their negative feelings with self-harm behaviors.

Hispanic adolescents who develop eating disorders face similar life stressors as those who self-harm, including family dysfunction and family rigidity. Research has also shown that additional factors, including body dissatisfaction, sexual abuse, and acculturation can contribute to the development of eating disorders (Rodrigues, 2017). Social comparisons and perceived pressure to adhere to Western ideals of appearance are especially troubling factors impacting body dissatisfaction or the desire to lose weight, and they can be predictors of disordered eating. Body image and body dissatisfaction are complex issues Hispanic adolescents face as they function within a dominant culture in North America that equates beauty with success and where beauty means being thin (Schooler, 2008). Equating beauty with thinness can conflict with

cultural and family values that embrace a curvier female shape as the ideal, or those that put less emphasis on appearance and have more flexible physical ideals (Warren, Schoen, & Schafer, 2010).

While anorexia and NSSI are discrete issues, they do overlap. It is estimated that 30–40% of adolescents with anorexia also engage in self-harming behaviors (Koutek, Kocourkova, & Dudova, 2016). Eating disorders are already associated with severe medical problems, and engaging in self-harm behaviors increases the health risks of the condition (Olatunji, Cox, Ebesutani, & Wall, 2015). Adolescents with eating disorders might engage in self-harm behaviors to manage their emotions and to communicate their distress, which could indicate a lack of social problem-solving skills (Nock & Prinstein, 2005). Anorexia and NSSI behaviors implicate difficulty regulating and expressing emotions; moreover, given the scale of their associated problems in society today, understanding the characteristics and contextual features of these conditions is both a public health necessity and a social justice obligation.

Because so many adolescents live their lives through online media, young people are less inclined to watch the traditional television news; instead, they learn about world and local events through video blogs and microblogs, particularly sites such as YouTube, Twitter, and Instagram. Unfortunately, however, microblogs and video-based social media outlets are heavier on opinions than objective facts and are also rife with misinformation (Li & Du, 2014; Syed-Abdul et al., 2013). Chancellor (2016) argued that visual images, like those shared on social media, influence and perpetuate distorted thinking. The use of images and the prevalence of opinions in pro-NSSI and pro-ana posts can potentially influence others to continue and even increase their behaviors through social reinforcement. The reality is that pro-NSSI and pro-ana blogs and social

media pages, while purporting to provide support, end up promoting disordered behaviors and discourage people from seeking help or taking steps to recover.

The influence of social media and blogs on individuals, particularly adolescents, is not intrinsically harmful (Weinstein, 2018). Adolescents report conflicting and co-existing feelings about their engagement with social media. Weinstein (2018) found that adolescents reported that social media facilitates closeness but also encourages negative emotions and feelings of being left out. In the same sense, social media can promote self-expression while also causing feelings of stress about being judged. So, too, can social media inspire a sense of exploration while heightening feelings of distress about the state of the world—simultaneously eliciting admiration and encouraging envy among users. Among more vulnerable populations, such an emotional rollercoaster can be overwhelming, meaning that the community that initially provided relief and support can quickly become part of why users continue engaging in the self-destructive behaviors that led them to look for support in the first place.

As participation in social media increases, research is being done to explore the language used online, such as noting the psychological markers in the language used by people with mental health issues (Choudhury, Counts, & Horvitz, 2013). It is now possible to identify people who are struggling with mental health issues based on how they post on social media platforms such as Tumblr and Instagram. Because of this prevalent online engagement among adolescents, and given the potential for significant positive and negative effects of exposure and interaction, research must continue to explore how to identify high-risk users. The need for additional study is amplified by the fact that, despite efforts to curtail the activities of pro-NSSI and pro-ana groups, they continue to thrive. Specifically, online platforms have taken steps to ban or block

the language associated with these communities, but groups have evaded these restrictions with language that has evolved to survive in this space.

To this point, research has yet to adequately account for the nature and significance of this evolution in language, particularly as words are expressed in Spanish by pro-NSSI and pro-ana online communities. One relatively new methodological tool for conducting this kind of research is corpus linguistics, a method where a corpus is constructed using large amounts of data. To address the above-described gap in the literature and to advance society's understanding of, and eventual response to, the public health concerns posed by self-harm and anorexia behaviors, the present studies culled pro-NSSI and pro-ana public posts from the website Tumblr. Using a Neo-Firthian theoretical framework for analysis of the corpus, which views language variation as systematic and able to be described through empirical, quantitative methods, this research assumes that the production of language is a result of psychological, academic, social and personal skills that one acquires through life (Biber, Reppen, & Friginal, 2010). Tumblr was chosen because its users tend to be younger than the users of similar microblogs, portending a better potential for locating users at the age of onset for NSSI and anorexia behaviors. The LIWC software that this research relied on can help analyze the language specific to communities associated with self-harm (Pennebaker, Booth, Boyd, & Francis, 2015), and these studies deployed it to explore three domains: specific NSSI and pro-ana word content, linguistic processes, and psychological processes.

Description of the Research Manuscripts

For the first research manuscript included in this dissertation, data from public pro-NSSI Tumblr posts written in Spanish were used to analyze the NSSI-specific words used, the linguistic properties of the texts, and the psychological properties found in the language of the

posts. Tumblr is considered a microblog website where individuals can post photos and a short blog on a preferred topic. To curtail pro-NSSI, pro-eating disorder, and pro-suicide materials on its platform, Tumblr has banned all posts related to these topics; however, adolescents have discovered ways around the ban and continue to post on these topics using hidden hashtags and slang terms.

The *Journal of Adolescent Health* is the target journal for this manuscript, because it publishes research in the field of adolescent medicine and health, ranging from the basic biological and behavioral sciences to public health and policy. The *Journal* is committed to improving the lives of adolescents and young adults, and a recent article explored the complex relationship between health outcomes and social media use by adolescents. The article focused on specific site use and suggested that patterns of use of specific sites might be more important than overall frequency of use of social media. The analysis found that young adults showed distinct patterns for social media use, emphasizing that those patterns were associated with differences in health-related risk factors and behaviors. This work suggests that the anonymity of posting on Tumblr could contribute to the site's high degree of disclosure and sharing around sensitive topics (Ilakkuvan, Johnson, Villanti, Evans, & Turner, 2019).

With the above information as background, the present study addressed the following questions regarding the language of NSSI on a public social media platform:

1. In public pro-NSSI Tumblr posts written in Spanish, what was the use rate of NSSI-related word categories?
2. In public pro-NSSI Tumblr posts written in English, what was the use rate of pronoun categories?

3. In public pro-NSSI Tumblr posts written in English, did the use rate of pronoun categories differ from norms for such posts?

4. In public pro-NSSI Tumblr posts written in Spanish, did the use rate of the psychological process categories differ from norms for such posts?

To answer these questions, this study employed a synchronic corpus linguistic design for exploring public Tumblr posts related to NSSI (Weisser, 2017), analyzing the content with LIWC software. For the first research question, regarding the use rate of NSSI-related words, the following categories were analyzed: methods of NSSI, cutting-specific terms, NSSI terms, instruments used, reasons for NSSI, and hidden-hashtag terms. The second and third research questions, considering the use rate of specific pronouns, focused on the following categories: first-person singular pronoun, first-person plural pronoun, third-person singular pronoun, and third-person plural pronoun. The variables for the fourth research question, examining psychological process categories, included (a) negative emotion, (b) anxiety, (c) anger, (d) sadness, (e) cognitive process, (f) insight, (g) body, (h) health, (i) ingestion, and (j) death.

For the second research manuscript, data from public pro-ana Tumblr posts written in Spanish were used to analyze the use of specific pro-ana words, the linguistic properties of the texts, and the psychological properties associated with the language of these posts. The target journal for this manuscript is the *International Journal of Eating Disorders*, an outlet advancing scientific knowledge to understand, treat, and prevent eating disorders. The *Journal* publishes research related to the theory, methodology, etiology, clinical practice, and policies related to eating disorders. A recent article, for example, demonstrated strategies for using big-data approaches and the potential for use of big data in social media by examining a pro-eating-disorder thread on Reddit (Moessner, Feldhege, Wolf, & Bauer, 2018). The approaches outlined

in this article were predominantly exploratory and the results were descriptive, suggesting the *Journal* would be a welcoming venue for the present research.

Like the first manuscript, the second manuscript addressed public Tumblr posts written in Spanish; but the focus of the second work was the language of pro-ana behavior, with specific attention to the following questions:

1. In public pro-ana Tumblr posts written in Spanish, what was the use rate of anorexia-related word categories?
2. In public pro-ana Tumblr posts written in English, what was the use rate of pronoun categories?
3. In public pro-ana Tumblr posts written in English, did the use rate of pronoun categories differ from norms for such posts?
4. In public pro-ana Tumblr posts written in Spanish, did the use rate of the psychological process categories differ from norms for such posts?

As was the case with the first manuscript, the second manuscript employed a synchronic corpus linguistic design (Weisser, 2017) to assess content, scrutinizing public pro-ana Tumblr posts with LIWC software. For the first research question, the specific pro-ana content word-use categories evaluated were pro-ana specific terms, methods of anorexia, reasons for anorexia, and hidden-hashtag terms. The variables for the second and third research questions were first-person singular pronoun, third-person singular pronoun, first-person plural pronoun, and third-person plural pronoun. Finally, the fourth research question attended to the following word categories: (a) negative emotion, (b) anxiety, (c) anger, (d) sadness, (e) cognitive process, (f) insight, (g) body, (h) health, (i) ingestion, and (j) death.

Glossary

This study involved a specialized vocabulary perhaps unfamiliar to most readers, and thus key terms are defined below.

Code-Switching

Code-switching refers to “texts that contain terms in two or more different languages” (Vilares, Alonso, & Gomez-Rodriguez, 2016). An example of code-switching is “Spanglish,” a mix of Spanish and American English.

Corpus/Corpora

Corpus/Corpora refer(s) to a collection of texts, typically in a machine-readable form, organized to represent a kind of language and offering some amount or degree of annotation (essex.ac.uk, 1998).

Corpus Linguistics

In the words of McEnery and Hardie (2012), *corpus linguistics* could mean “either (a) any approach to language that uses corpus data and methods, or (b) an approach to linguistics that uses corpus methods but does not subscribe to principles of the so-called *corpus-driven* approach” (p. 241).

Hashtag

While originally designed for categorizing posts, the term *hashtag* can now signify material intended to be coy or witty. The word *tag* can indicate “a word or phrase used for a description or identification,” while *hash* is “short for hash mark, a term for what we more commonly call a pound sign” (Merriam-webster.com, n.d.).

Lexical Variants

Lexical variants (LV) refer to different forms of a word that represent the same expression, forms that could emerge from variations in spelling, grammar, or abbreviation (Zeng, 2003).

LIWC

As Tausczik and Pennebaker (2010) explain, LIWC is a “transparent text analysis program that counts words in psychologically meaningful categories,” adding that empirical results using LIWC “demonstrate its ability to detect meaning in a wide variety of experimental settings, including to show attentional focus, emotionality, social relationships, thinking styles, and individual differences” (p. 24).

Log-Likelihood Test

A log-likelihood test is a significance test that is “similar to the chi-square test, but generally considered more reliable, especially when working with small values” (McEnery & Hardie, 2012, p. 246).

Nonsuicidal Self-Injury

Nonsuicidal self-injury refers to the “deliberate, non-suicidal destruction of one’s own body tissue” (Favazza, 1998, p. 260). Those engaging in NSSI typically cut, bite, scratch, or burn themselves.

Out-of-Vocabulary Words

As Han and Baldwin (2011) put it, *out-of-vocabulary words* (OOV) refer to “neologisms and proper nouns like *hopeable* or *WikiLeaks* which have not made their way into the dictionary” (p. 370).

Pro-NSSI

To be pro-NSSI is to exude positivity toward or demonstrate encouragement of nonsuicidal self-injurious behaviors.

Pro-Ana

To be *pro-ana* is to exude positivity toward or demonstrate encouragement of behaviors related to anorexia.

Synchronic

Synchronic, as McEnery and Hardie (2011) describe it, relates to “the study of a language or languages as they exist at a particular moment in time, without reference to how they might change over time,” meaning that a derivative construct, *synchronic corpus*, “contains texts drawn from a single period—typically the present or the very recent past” (p. 251).

Textese

Textese is a “form of written language as used in text messages and other digital communications, characterized by many abbreviations and typically not following standard grammar, spelling, punctuation, and style” (dictionary.com, n.d.)

Thinspiration

Thinspiration is a term referring to “something or someone that serves as motivation for a person seeking to maintain a very low body weight” (dictionary.com, n.d.), with examples including photos of very thin celebrities, song lyrics, images of protruding bones, and so on.

Token

Token is “the smallest unit that each corpus divides to. Therefore, corpora contain more tokens than words. Spaces between words are not tokens.”

(https://www.sketchengine.eu/my_keywords/token/, n.d.)

Tumblr

Tumblr is “a cross between a social networking site (like Facebook and Twitter) and a blog.” It often functions as a “microblog” accommodating posts with short snippets of text and quick snaps, as opposed to the longer diary-style entries of traditional blogs. Compared to other social networks, Tumblr has a younger base of users (Webwise.ie, n.d.).

Thematic Links Between Manuscripts

These studies both seek to refine researchers’ understandings of the linguistic process by focusing on the use of specific terms deployed in public Tumblr posts written in Spanish. These data, results, and associated discussion will be presented in two research manuscripts, similar in design but different in terms of the behaviors and language being assessed. The first manuscript, Chapter 2, used data from public pro-NSSI Tumblr posts written in Spanish to better assess specific language markers for pro-NSSI posts. The second manuscript, Chapter 3, used data from public pro-ana Tumblr posts written in Spanish to more carefully evaluate pro-ana behaviors, with a focus on specific language, linguistic process, and psychological-process language markers. The research in these manuscripts fills a gap in the counseling literature by providing data and insight into pro-NSSI and pro-ana behaviors for those posting in Spanish. It also clarifies the thought processes of those who engage in NSSI and/or anorexia behaviors, eliciting more information to assist in the treatment of individuals who use such behaviors to cope with distressing feelings.

Organization of the Dissertation

This dissertation is divided into four chapters. Chapter 1 provides a general introduction that includes an overview of the major research on the topics of NSSI and anorexia among Hispanic adolescents, while explaining major themes that have emerged in earlier research. Chapter 2 is an original manuscript that uses data from public Tumblr posts written in Spanish to

identify NSSI-specific terms, linguistic processes, and psychological processes in the language of the posts. Chapter 3 is a separate original manuscript that uses data from public pro-ana Tumblr posts written in Spanish to identify pro-ana specific terms, linguistic processes and psychological processes in the language of the posts. Finally, Chapter 4 concludes this dissertation, reviewing the findings from the research manuscripts, summarizing their significance, attending to thematic links between the studies, and setting forth a research agenda building on this work.

References

- Aalai, A. (2018, June 8). The social contagion of suicide [Blog post]. Retrieved from <https://www.psychologytoday.com/us/blog/the-first-impression/201806/the-social-contagion-suicide>
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.
- Arcelus, J., Mitchell, A. J., Wales, J., & Neilson, S. (2011). Mortality rates in patients with anorexia nervosa and other eating disorders: A meta-analysis of 36 studies. *Archives of General Psychiatry*, 68(7), 724–731. <https://doi.org/10.1001/archgenpsychiatry.2011.74>
- Baron, N. (2008). *Always on: Language in an online and mobile world*. New York, NY: Oxford University Press.
- Biber, D. E., Reppen, R., & Fringinal, E. (2012). Research in corpus linguistics. In *The Oxford Handbook of Applied Linguistics*, (2nd ed.) Oxford University Press. <https://doi.org/10.1093/osfordhb/9780195384253.013.0038>
- Brown, R. C., Fischer, T., Goldwich, A. D., Keller, F., Young, R., & Plener, P. L. (2018). #cutting: Non-suicidal self-injury (NSSI) on Instagram. *Psychological Medicine*, 48(2), 337–346. <https://doi.org/10.1017/S0033291717001751>
- Campbell, K., & Peebles, R. (2014). Eating disorders in children and adolescents: State of the art review. *Pediatrics*, 134(3), 582–592. <https://doi.org/10.1542/peds.2014-0194>
- Castro, T. S., & Osorio, A. J. (2013). “I love my bones!”¹ - self-harm and dangerous eating youth behaviors in Portuguese written blogs. *Young Consumers*, 14(4), 321–330. <https://doi.org/10.1108/YC-03-2013-00351>

- Cervantes, R. C., Goldback, J. T., Varela, A., & Santisteban, D. A. (2014). Self-harm among Hispanic adolescents: Investigating the role of culture-related stressors. *Journal of Adolescent Health, 55*, 633–639. <https://doi.org/10.1016/j.jadohealth.2014.05.017>
- Chancellor, S., Lin, Z., Goodman, E. L., Zerwas, S., & Choudhury, M. D. (2016). Quantifying and predicting mental illness severity in online pro-eating disorder communities. *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, 1171–1184. <https://dx.doi.org/10.1145/2818048.2819973>
- Choudhury, M. D., Counts, S., & Horvitz, E. (2013). Major life changes and behavioral markers in social media. *Proceedings of the 2013 Conference on Computer-Supported Cooperative Work*, 1431–1442. <https://dx.doi.org/10.1145/2441776.2441937>
- Cipriano, A., Cella, S., & Cotrufo, P. (2017). Nonsuicidal self-injury: A systematic review. *Frontiers in Psychology, 8*(1946), 1–14. <https://doi.org/10.3389/fpsyg.2017.01946>
- Croyle, K. L., (2007). Self-harm experiences among Hispanic and Non-Hispanic White young adults. *Hispanic Journal of Behavioral Sciences, 29*(2), 242–253.
- Explainer: What is Tumblr? (2018, August 08). Retrieved from <https://www.webwise.ie/parents/explainer-what-is-tumblr-2/>
- Favazza, A. R. (1998). The coming of age of self-mutilation. *The Journal of Nervous & Mental Disease, 186*, 259–268. <https://dx.doi.org/10.1097/00005053-199805000-00001>
- Fuertes, J. N. (2004). Supervision in bilingual counseling: Service delivery, training, and research considerations. *Journal of Multicultural Counseling and Development, 32*, 84–94. <https://doi.org/10.1002/j.2161-1912.2004.tb00363.x>
- Gonzalez-Barrera, A., & Lopez, M. H. (2013). Spanish is the most spoken non-English language in U.S. homes, even among non-Hispanics. Retrieved from <http://pewrsr.ch/11S8uek>

- Gulbas, L. E., Hausmann-Stabile, C., De Luca, S. M., Tyler, T. R., & Zayas, L. H. (2015). An exploratory study of non-suicidal self-injury and suicidal behaviors in adolescent Latinas. *American Journal of Orthopsychiatry*, *84*(4), 302–314.
<http://dx.doi.org/10.1037/ort0000073>
- Guttfreund, D. G. (1990). Effects of language usage on the emotional experience of Spanish-English and English-Spanish bilinguals. *Journal of Consulting and Clinical Psychology*, *58*(5), 604–607. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/2254507>
- Hashtag. (n.d.). Retrieved from <https://www.merriam-webster.com/dictionary/hashtag>
- Ilakkuvan, V., Johnson, A., Villanti, A. C., Evans, W. D., & Turner, M. (2019). Patterns of social media use and their relationship to health risks among young adults. *Journal of Adolescent Health*, *64*(2), 158–164. <https://doi.org/10.1016/j.jadohealth.2018.06.025>
- Jarvi, S., Jackson, B., Swenson, L., & Crawford, H. (2013). The impact of social contagion on non-suicidal self-injury: A review of the literature. *Archives of Suicide Research*, *17*, 1–19. <https://doi.org/10.1080/13811118.2013.748404>
- Klonsky, D. E., Victor, S. E., & Saffer, B. Y. (2014). Nonsuicidal self-injury: What we know, and what we need to know. *Canadian Journal of Psychiatry*, *59*(11), 565–568.
<https://dx.doi.org/10.1177%2F070674371405901101>
- Koutek, J., Docourkova, J., & Dudovva, I. (2016). Suicidal behavior and self-harm in girls with eating disorders. *Neuropsychiatric Disease and Treatment*, *12*, 787–793.
<http://dx.doi.org/10.2147/NDT.S103015>
- Lewis, S. P., Heath, N. L., Michal, N. J., & Duggan, J. M. (2012). Non-suicidal self-injury, youth, and the internet: What mental health professionals need to know. *Child and*

- Adolescent Psychiatry and Mental Health*, 6(13), 1–9. <https://doi.org/10.1186/1753-2000-6-13>
- Li, F., & Du, T. C. (2014). Listen to me — Evaluating the influence of micro-blogs. *Decision Support Systems*, 62, 119–130. <https://dx.doi.org/10.1016/j.dss.2014.03.008>
- McEnery, T., & Hardie, A. (2012). *Corpus linguistics: Method, theory and practice*. Cambridge, UK: Cambridge University Press.
- Moessner, M. M., Feldhege, J., Wolf, M., & Bauer, S. (2018). Analyzing big data in social media: Text and network analyses of an eating disorder forum. *International Journal of Eating Disorders*, 51, 656–667. <https://doi.org/10.1002/eat.22878>
- Nock, M. K., & Prinstein, M. J. (2005). Contextual features and behavioral functions of self-mutilation among adolescents. *Journal of Abnormal Psychology*, 114(1), 140–146. <https://psycnet.apa.org/doi/10.1037/0021-843X.114.1.140>
- Olatunji, B. O., Cox, R., Ebesutani, C., & Wall, D. (2015). Self-harm history predicts resistance to inpatient treatment of body shape aversion in women with eating disorders: The role of negative affect. *Journal of Psychiatric Research*, 65, 37–46. <https://doi.org/10.1016/j.jpsychires.2015.03.016>
- O'Reilly, M., Dogra, N., Whiteman, N., Hughes, J., Eruyar, S., & Reilly, P. (2018). Is social media bad for mental health and wellbeing? Exploring the perspectives of adolescents. *Clinical Child Psychology and Psychiatry*, 23, 601–613. <https://dx.doi.org/10.1177/1359104518775154>
- Pennebaker, J. W., Booth, R. J., Boyd, R. L., & Francis, M. E. (2015). *Linguistic Inquiry and Word Count: LIWC2015* (Operator's Manual). Austin, TX: Pennebaker Conglomerates, Inc.

- Perry, S. (2013). The bilingual brain. Retrieved from <http://www.brainfacts.org/Archives/2008/The-Bilingual-Brain>
- Ramos-Sanchez, L. (2007). Language switching and Mexican Americans' emotional expression. *Journal of Multicultural Counseling and Development, 35*, 154–168. <https://psycnet.apa.org/doi/10.1002/j.2161-1912.2007.tb00057.x>
- Rodrigues, M. (2017). Do Hispanic girls develop eating disorders? A critical review of the literature. *Hispanic Health Care International, 15*(4), 189–196. <https://doi.org/10.1177%2F1540415317744500>
- Santiago-Rivera, A. L., & Altarriba, J. (2002). The role of language in therapy with the Spanish-English bilingual client. *Professional Psychology: Research and Practice, 33*(1), 30–38. <http://dx.doi.org/10.1037/0735-7028.33.1.30>
- Self-injurious behavior, self-injury treatment. (2010, February 22). Retrieved from <https://www.healthyplace.com/abuse/articles/self-injurious-behavior-self-injury-treatment>
- Syed-Abdul, S., Fernandez-Luque, L., Jian, W., Li, Y., Crain, S., Hsu, M., . . . Liou, D. (2013). Misleading health-related information promoted through video based social media: Anorexia on YouTube. *Journal of Medical Internet Research, 15*(2), e30. <https://doi.org/10.2196/jmir.2237>
- Tausczik, Y. R., & Pennebaker, J. W. (2009). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology, 29*, 24–54. <https://dx.doi.org/10.1177/0261927x09351676>
- Textese. (n.d.). Retrieved from <https://www.dictionary.com/browse/textese>
- Thinspiration. (n.d). Retrieved from www.dictionary.com

- Thyssen, L. S., & van Camp, I. (2014). Non-suicidal self-injury in Latin America. *Salud Mental*, 4(37), 153–157. <https://psycnet.apa.org/doi/10.17711/SM.0185-3325.2014.019>
- Token. (n.d.). Retrieved from https://www.sketchengine.eu/my_keywords/token/
- Vega, D., Sintes, A., Fernandez, M., Punti, J., Soler, J., Santamarina, P., . . . Pascual, J. C. (2018). Revision y actualizacion de la autolesion no suicida: Quien, como y por que? *Actas Espanolas de Psiquiatria*, 46(4), 146–155.
- Velazquez, I. (2017). Reported literacy, media consumption and social media use as measures of relevance of Spanish as a heritage language. *International Journal of Bilingualism*, 21(1), 21–33. <https://doi.org/10.1177%2F1367006915596377>
- Vilares, D., Alonso, M. A., & Gomez-Rodriguez, C. (2016). *EN-ES-CS: An English-Spanish code-switching Twitter corpus for multilingual sentiment analysis*. Paper presented at the Tenth International Conference on Language Resources and Evaluation, Paris, France. Abstract retrieved from https://www.researchgate.net/publication/303666796_EN-ES-CS_An_English-Spanish_Code-Switching_Twitter_Corpus_for_Multilingual_Sentiment_Analysis
- Warren, C. S., Schoen, A., & Schafer, K. J. (2010). Media internalization and social comparison as predictors of eating pathology among Latino adolescents: The moderating effect of gender and generational status. *Sex Roles*, 63, 712–724. <https://doi.org/10.1007/s11199-010-9876-1>
- Weisser, M. (2016). *Practical corpus linguistics: An introduction to corpus-based language analysis*. Hoboken, NJ: John Wiley & Sons, Inc.
- Zeng, M. L. (2005). Semantic relationships used in controlled vocabularies. Retrieved from <http://marciazeng.slis.kent.edu/Z3919/42equivalency.htm>

Chapter 2: A Research Manuscript

A Corpus Linguistic Analysis of Pro-Nonsuicidal Self-Injury Public Tumblr Posts Which Use Spanish (Un Analisis de Lingüística de Corpus de Publicaciones Pública que Usan Español en Tumblr que Respaldan a las Pró Autolesiones No Suicidas)

Kristeen E. Elrod and Cass Dykeman

Oregon State University

Author Note

Kristeen E. Elrod, Counseling Academic Unit, Oregon State University; Cass Dykeman, Counseling Academic Unit, Oregon State University.

The research contained in this manuscript was conducted under the approval of the Oregon State University Institutional Review Board (Study ID No. 8721) and was part of the first author's dissertation research project.

Correspondence concerning this article should be addressed to Cass Dykeman, Counseling Academic Unit, Oregon State University, College of Education, Joyce Collin Furman Hall, 200 SW 15th Street, Corvallis, OR 97331-3502. Contact: dykemanc@onid.oregonstate.edu

Abstract

Rates of nonsuicidal self-injury (NSSI), especially high among adolescents, are associated with comorbid disorders including depression, anxiety, post-traumatic stress disorder, and eating disorders. Hispanic adolescents engage in NSSI at rates similar to or higher than their Caucasian counterparts, but little research exists on Spanish-language markers of mental health. This study, harnessing data from the social media platform Tumblr, collected pro-NSSI public blog posts written using Spanish and analyzed their content with Linguistic Inquiry and Word Count (LIWC) software. Pro-NSSI linguistic properties, psychological linguistic properties, and signature words were analyzed, and results revealed that, of pro-NSSI terms, *reasons for NSSI* was the most frequently used category in posts. This result aligns with the literature because words associated with negative emotions are commonly presented in the dominant language within trauma narratives. Evaluation of linguistic properties indicated that users could be code-switching to avoid encountering overwhelming emotions, a practice common among bilingual individuals. The psychological properties of these blog posts had an emphatically negative emotional tone when written in Spanish and a more concrete and less emotionally charged tone when users switched to English. This research suggests benefits to assessing bilingual clients for pro-NSSI and online-posting behaviors, while underscoring both the need for improved evidence-based practices for treating trauma in bilingual clients and, broadly speaking, the importance of Spanish and code-switching norms for blogs and social media platforms.

Keywords: NSSI, LIWC, linguistic analysis, mental health, self-injury

Palabras Claves: lingüística de corpus, pró autolesiones no suicidas (ANS), salud mental, LIWC

A Corpus Linguistic Analysis of Pro-Nonsuicidal Self-Injury Public Tumblr Posts Written in Spanish (Un Estudio de Lingüística de Corpus de Publicaciones Públicas de Tumblr a las Pró Autolesiones No Suicidas en Español)

Online communities use distinctive linguistic practices to define terms of membership, build solidarity among members, share practices internally, and distinguish themselves from outsiders. Language plays an important role in signaling a shared identity and is subject to ongoing adaptation. In online communities where membership is discouraged by content-banning or blocking, linguistic variations of hashtags grow more popular, complex, and detached from original spellings, usage, and context. For parents, teachers, and clinicians, this can be confusing, diminishing the understanding enjoyed by these individuals, as well as their access to the community and their ability to provide care and support.

The literature on the social media posting habits of pro-nonsuicidal self-injury (NSSI) communities centers on five basic themes: (a) the prevalence of NSSI, (b) the prevalence of NSSI in Hispanic communities, (c) the prevalence of social media use, (d) the presence of pro-NSSI communities on social media, and (e) the lack of education for parents, teachers, and clinicians on pro-NSSI terms in Spanish. The research questions driving this study will be presented and considered in more detail following a thorough review of each of the above themes.

Favazza (1998) has described NSSI as the “deliberate destruction of one’s own body tissue in the absence of conscious suicidal intent” (p. 260); and, as others have demonstrated, it is not a new phenomenon (Moreno, Ton, Selkie, & Evans, 2016; Nock & Favazza, 2009). Self-harm and NSSI involve behaviors such as cutting, burning, picking, scratching, interfering with wound healing, and hitting an object until injuring oneself; they could also take on potentially

more dangerous forms, such as breaking bones or poisoning oneself (Thyssen & van Camp, 2014). One study found 15% of adolescents engaging in such behaviors, with the rate for college-age individuals approaching 35% (Kerr, Muehlenkamp, & Turner, 2010). Importantly, NSSI is associated with comorbid disorders such as depression, post-traumatic stress disorder, generalized anxiety disorder, and eating disorders (Bentley, Nock, & Barlow, 2014), warranting NSSI's classification as a transdiagnostic issue for further study in the fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5, American Psychiatric Association, 2013). Particularly for migrant Hispanics, the focus of the second of the five primary themes considered in this review of the literature, these risk factors include fragmented family structure, family conflict, parental criticism, transnational stress, bullying, and experiences with violence, as well as acculturation issues for the individual and family (Benjet et al., 2017).

Regarding the third theme in this discussion of the research surrounding NSSI, social media has changed ways that people seek support from and interact with others. Online support groups can provide guidance and a sense of community for both positive and negative behaviors. Some support groups promote self-harm as a lifestyle and help their members maintain this lifestyle at great cost to their health and relationships (Chancellor, Kalantidis, Pater, De Choudhury, & Shamma 2017). The pervasiveness of social media affects the behavior and choices of adolescents and young adults who could be engaging in self-harm and who are experiencing ongoing mental health issues. Research suggests adolescents and young adults spend about nine hours per day on social media, while checking social media sites approximately 100 times per day (Wallace, 2015). On social media platforms, NSSI is especially interesting given how peer and media influences can encourage ongoing behaviors (Moreno et al., 2016).

Turning to the fourth theme in this review, among the most common social media platforms for teens and young adults are Instagram (Pew Research Center, 2018) and Tumblr (Pew Research Center, 2015). Both sites have been criticized for continuing to allow online communities promoting NSSI behaviors, eating disorders, and suicide. Given its secretive nature, NSSI can be tremendously isolating; thus, a community can offer understanding and/or a sense of belonging for those in need of succor (Joiner, Van Orden, Witte, & Rudd, 2009). While Instagram and Tumblr have attempted to eliminate these communities, warning users about their content, blocking hashtags, and so on, the communities continue to thrive (Matsakis, 2018). Banning or eliminating these communities entirely could be dangerous for their members, because online cohorts could be a user's only lifeline. That said, pro-NSSI communities can also trigger NSSI urges and normalize or reinforce NSSI behaviors (Brown et al., 2018).

Turning to the fifth and final theme, Linguistic Inquiry and Word Count (LIWC, pronounced "Luke") software can analyze language used by and specific to pro-NSSI communities (Pennebaker, Booth, Boyd, & Francis, 2015). In a manner akin to the way species must adapt to survive in nature, online pro-NSSI communities have evolved, changing the language they use to communicate and interact with one another. And yet, research has yet to appreciate this transformation, especially regarding Spanish content.

Thus, this study focused on words and phrases associated with NSSI, written in Spanish and appearing in public social media posts on Tumblr. Tumblr was chosen because it is used more by teens and young adults (between 13 and 24), ages coinciding, in most cases, with the onset of NSSI behaviors. Understanding characteristics and contextual factors influencing Spanish-speaking pro-NSSI communities is a critical mental health and social justice matter,

especially given the growth of the Hispanic population in the United States (Gulbas, Hausmann-Stabile, De Luca, Tyler, & Zayas, 2015).

Inspired by these themes and building on the abovementioned literature, four questions guided this study. First, in public pro-NSSI Tumblr posts written in Spanish, what was the use rate of pro-NSSI-related word categories? Second, in public pro-NSSI Tumblr posts written in English, what was the use rate of pronoun categories? Third, in public pro-NSSI Tumblr posts written in English, did the use rate of pronoun categories differ from norms for such posts? And, fourth, in public pro-NSSI Tumblr posts written in Spanish, did the use rate of the psychological process categories differ from norms for such posts?

Method

Design

This study employed a synchronic corpus linguistic design (Weisser, 2017). The unit of analysis was single words (Bjekić, Lazarević, Živanović, & Knežević, 2014). The variables for the first research question included the following word categories: (a) methods of NSSI, (b) cutting-specific terms, (c) NSSI terms, (d) instruments used, (e) reasons for NSSI, and (f) hidden-hashtag terms. The variables for the second and third research questions involved the word categories first-person singular pronoun, third-person singular pronoun, first-person plural pronoun, and third-person plural pronoun. The variables for the fourth research question included (a) negative emotion, (b) anxiety, (c) anger, (d) sadness, (e) cognitive process, (f) insight, (g) body, (h) health, (i) ingestion, (j) risk, and (k) death. All variables were reported as a percentage of total words. There were three sub-corpora in this study: (a) Spanish-only, (b) English-only, and (c) Mixed. In terms of the third and fourth research questions, the comparison norm for all corpora was Twitter's set of blog norms in English. There are currently no blog norms for

Spanish or mixed-language situations that might serve as respective baselines for the Spanish-only and Mixed sub-corpora.

The inferential analysis used in this study was a log-likelihood (LL) ratio test, a test that evaluates goodness-of-fit to a hypothesized distribution. Since the LL ratio can be approximated by Pearson's χ^2 , a priori power analysis for a χ^2 test was employed using G*Power 3.1 (Faul, Erdfelder, Buchner, & Lang, 2009). The proper effect size is Cohen's w (Rosnow & Rosenthal, 2003). The effect size input was the average of the Cohen's w reported in a study on self-harm that was relevant to the present study (Harned, Najavits, & Weiss, 2006). Given the number of comparisons planned, the α level was set at .0033. The specific input parameters were: (a) test family = χ^2 tests, (b) statistical test = goodness-of-fit tests: contingency tables, (c) type of analysis = a priori: Compute required sample size given α , power, and effect size, (d), $w = 0.52$, (e) power (1- β error probability) = 0.9, (f) $\alpha = .0033$, and (g) degrees of freedom = 1. The G*Power 3.1 output included a sample size of 78 and an actual power of 0.90.

Corpus

Scope. Tumblr's Application Programming Interface (API) was used to collect public posts occurring between October 23, 2017 and October 23, 2018. The API was instructed to scrape posts that included Spanish in the post and/or hashtags, that deployed hashtags identifying the post as pro-NSSI (e.g., #autolesiones signified *self-harm*), and that originated in a North, Central, or South American country. The API was instructed to eliminate the following while collecting text: usernames, URLs, location, names, photographs, comments left by other users, and direct quotes. Between October 23, 2017 and October 23, 2018, 1,868 posts and associated hashtags were collected (<https://www.tumblr.com/docs/en/api/v2>).

Construction. Captions for and hashtags associated with posts were scrutinized, and collected posts were entered in an Excel document. First, posts written in any language other than English or Spanish were excluded, and then posts written entirely in English were excluded. The corpus, which contained posts in Spanish, posts where whole sections were in English or Spanish, and posts with mixed language (switching between English and Spanish in the same section), were then separated into three sub-corpora for analysis at the section level. These sub-corpora were: (a) English-only, including sections of the post, either its caption or hashtags, which were entirely in English, while the rest of the post had Spanish language, (b) Spanish-only, including sections of the post where either the caption or hashtags were entirely in Spanish, and (c) Mixed, including sections of the post where code-switching occurred in the caption or hashtag.

Preprocessing. Posts in all sub-corpora were pre-processed using procedures outlined in the LIWC operator's manual and a supplement to this manual (Department of Psychological Services, 2015; Pennebaker, Booth et al., 2015). This cleaning involved the following processes: (a) all abbreviations were spelled out (e.g., "max" became *maximum*), (b) NSSI slang (words used to hide NSSI-related Instagram posts) was not normalized (Moreno et al., 2016), and (c) textese was translated into standard spelling. Once the text was ready for analysis, the file was processed by LIWC software.

Size. The size of the Spanish-only sub-corpus was 42,661 tokens. The English-only sub-corpus contained 9,814 tokens. The Mixed sub-corpus included 4,299 tokens.

Measures

The measures used were selected scales from the LIWC scales and the first author's Spanish adaptation of the Greaves NSSI Linguistic Scales (Greaves & Dykeman, 2019). Scores

for all variables represented a percentage of all words used. The LIWC scales were pre-set scales contained in the LIWC software. Pennebaker, Boyd, Jordan, and Blackburn (2015) reported adequate reliability and validity for these scales. The reference norms used for comparison to the present results were Twitter norms contained in the LIWC psychometric manual (Pennebaker, Boyd et al., 2015). Internal reliability and external validity of LIWC software were well documented (Pennebaker, Boyd et al., 2015). Six NSSI-specific measures were employed.

GNLS_S-Methods of NSSI. Methods of NSSI are ways people engage in nonsuicidal self-injury, a broad category including all manner of self-harming behavior. Some examples of NSSI methods include biting (*mordiendo*), burning (*quemando*), erasing (*borrando*), hitting (*golpeando*), and picking (*raspando*).

GNLS_S-Cutting-Specific Terms. Cutting is a specific category under methods of NSSI. Since cutting is one of the most common methods of NSSI, it has its own sub-category. Two examples for this category are the seed words *cut* (*cortar*) and *cutting* (*cortando*)

GNLS_S- NSSI Terms. Some examples of NSSI terms include *nonsuicidal* (*no suicida*), *NSSI* (*ANS*), *suicide* (*suicida*) and the seed word *self-mutil* (*automutilacion*).

GNLS_S-Instruments Used. This category includes tools people use to engage in NSSI behaviors, instruments such as blades (*navaja*), bleach (*blanqueador*), erasers (*borrador*), and fingernails (*unas*).

GNLS_S-Reasons for NSSI. There are many reasons why people choose to engage in NSSI behaviors, including feeling angry (*enojo*), anxious (*ansiedad*), or empty (*vacio*) and seeking attention (*atencion*) or desiring control (*control*).

GNLS_S-Hidden Hashtag Terms. To evade recently imposed restrictions, those who engage in NSSI behaviors and who post about their activities on social media sites have become

more creative with the hashtags they use to communicate with others engaging in NSSI. Some of their new “hidden” hashtags include #mifamiliasecreta (mysecretfamily), and #munecas (dolls).

Pronouns

LIWC was employed to assess the frequency of the following four pronoun types: first-person singular, third-person singular, first-person plural, and third-person plural.

Psychological Processes

The LIWC software was also used to assess 10 psychological processes. These processes, with the words that represented them in parentheses, included negative emotion (nasty), anxiety (worried), anger (hate), sadness (crying), cognitive processes (know), insight (consider), body (hands), health (flu), ingestion (dish), and death (coffin).

Comparison Norms

English Twitter norms contained in the LIWC psychometric manual were used to evaluate the three sample corpora (Pennebaker, Boyd et al., 2015). There are currently no Spanish or mixed-language blog norms available for comparison. While English and Spanish are syntactically dissimilar languages, it is reasonable to assume a roughly one-to-one correspondence between English nouns and Spanish nouns in terms of word count (Bailey, McIntyre, Arreola, & Venta, 2019). Thus, the second sub-corpus (Spanish-only) was compared to English Twitter norms for psychological process categories.

Apparatus

The LIWC software has been updated several times, but this study relied on the 2015 English version and the 2007 Spanish version. LIWC2015 offered new dictionaries, including more social words and cognitive-process words (Pennebaker, Booth, et al., 2015). The LIWC program can be utilized to review and analyze many kinds of written text. For the English sub-

corpus, the English LIWC2015 dictionary was used; the Spanish LIWC2007 dictionary was used for the Spanish sub-corpus; and, finally, the Mixed sub-corpus was analyzed using both the English LIWC2015 and Spanish LIWC2007 dictionaries, as recommended by Pennebaker (personal communication, October 30, 2018). LIWC analysis returned results for 90 output variables (i.e., LIWC categories), and LIWC category scores represented the percentage of total words used for the sample. Words could belong to more than one category.

Data Analysis

For the first and second research questions, raw count and percentage of total words used were reported for each variable. For the third and fourth questions, the LL test was used to compare the sub-corpora to norms (Rayson & Garside, 2000). The Bayesian Information Criterion (BIC) was the effect-size measure reported (Brezina, 2018). Analyses were completed using *R*. Given the large number of hypothesis tests conducted, a Bonferroni correction was used to reduce the chance of a Type I error.

Results

The first research question looked at word-usage rates, specifically among pro-NSSI word categories, in three groups of sub-corpora. Across all sub-corpora, the most common NSSI-specific category was *reasons for NSSI*. Complete results can be found in Table 1. For the second and third research questions, the linguistic process category that differed most from the English Twitter blog norms was *first-person singular* (LL = 40.21). The effect size for this category was BIC = 30.74, where anything greater than 10 was considered large. The linguistic process category *third-person singular* also differed from English Twitter blog norms (LL = 27.78), with the effect size of BIC = 18.3. So, too, did the linguistic process category *first-person plural* differ from English Twitter blog norms (LL = 20.85), yielding an effect size of BIC = 11.38. The

results for these research questions can be found in Tables 2 and 3. Finally, regarding the fourth research question, in the English sub-corpus, the psychological process category that differed most from English Twitter blog norms was *health* (LL = 151.79, BIC = 142.32). In the Spanish sub-corpus, the psychological process category that differed most from English Twitter blog norms was *negative emotions* (LL = 261.61, BIC = 250.87). In the Mixed sub-corpus, the psychological process category that differed most from English Twitter blog norms was *negative emotions* (LL = 256.38, BIC = 247.46). The complete results for this research question can be found in Tables 4, 5, and 6.

Discussion

This study examined language used in public pro-NSSI blogs written in Spanish and posted on Tumblr, focusing specifically on differences in the use of language in Spanish pro-NSSI blogs as compared to blogs overall. Regarding the first research question, where *reasons for NSSI* was used most across sub-corpora, two explanations can account for these results. The first is that words in the *reasons for NSSI* category (e.g., anger, anxiety, and sadness) are more associated with words suggestive of negative emotions. These negative-emotion words, particularly during the recall or recollection of traumatic incidents, arise more often when communicants use their dominant language (i.e., Spanish) (Bailey et al., 2019), which is likely the case during posts about self-injury. The alternative explanation is that engaging in this community is a way to belong and feel understood among peers, facilitating discussions about the impulse to harm oneself. Between these two explanations, the former is the most plausible, because past research has demonstrated a relationship between self-injurious behavior and negative emotion-generating experiences such as childhood abuse.

The results from the second and third research questions revealed that pronoun use differed from English Twitter norms most frequently and significantly as *first-person singular*, followed by *third-person plural* and *first-person plural*. Two explanations could account for these results. First, the use of the first-person singular pronoun could indicate depression and perceived lower social status (Pennebaker, 2011). The English sub-corpus included language taken from posts where Spanish had been used, which assumes users are bilingual and also that they switched languages within the post. When bilingual individuals use their non-dominant language to discuss emotional issues, it can also be a way for them to create distance from or avoid their emotional experience and associated vulnerability. Self-harm is a way to avoid directly confronting intense emotions, so it would make sense that bilingual individuals who engage in self-harm might code-switch when discussing their experience in order to avoid those same intense emotions.

Second, it could be that the presence of code-switching denotes greater acculturation, thereby influencing pronoun use. Between these two explanations, the former is more likely because the sample was taken from Tumblr posts originating in North, Central, or South America. While acculturation, or adaptation of attitudes and behaviors to a host culture, might be influencing posts from some countries in this region, it does not accurately explain the code-switching occurring in posts from Spanish-dominant countries with native dominant culture.

The results from the fourth research question revealed the psychological process word-use patterns across all three pro-NSSI sub-corpora. The most striking result was for *negative emotion*. With the sub-corpora that contained Spanish, this category ranked first in effect size; however, in the case of the English sub-corpora, *health* ranked first, with *negative emotion* a distant second.

There could be two explanations for this result. First, research has shown that expression in the non-dominant language might be more concrete and less emotionally charged (Bailey et al., 2019), explaining why *health* could be substantially greater than *negative emotion* in the second-language-only sub-corpus. Second, it could be that the obtained results were encountered because Romance languages permit greater expression of emotionality than Germanic languages. Between these options, the former is most likely the correct explanation for two reasons. First, on a psycho-lexical level, these language groups are similar (Ashton et al., 2004). Second, the former conjecture is consistent with prior research suggesting that people who engage in self-harm are in immense pain and would be expected to use a negative emotional tone to express their sadness. Pennebaker (2013) suggested that those who use negative emotional language do not benefit from writing about it and continue to dwell on it. Writing or posting on Tumblr, then, does not help individuals process these emotions; instead, it can contribute to a cycle where they carry on with the same feelings and behaviors.

While these data are compelling, there are a few admitted limitations to this study. First, words can be polysemous, and words in LIWC could appear in more than one category, somewhat skewing the data. To control for this potential and to avoid confusion, the data were reviewed before being analyzed by LIWC software. Second, only public Tumblr posts from North, Central, and South American countries were considered in this study. It is possible that people who post on Tumblr are doing so through a private account; but they could also be Spanish speakers who do not happen to reside in countries within North, Central, or South America. This combination of factors could fail to account for the entirety of the population. Third, the meanings of words can vary across languages. This study used a Spanish-language corpus, but it also relied on Spanish, English, and Mixed sub-corpora because of the

phenomenon of code-switching. And there were no Spanish or mixed-language blog norms to use for comparison. As such, while the effect size for this study was large, inferences drawn from these data should be colored by the above concerns.

This study reveals two implications for practitioners and one implication for researchers. First, this research suggests what to inquire about when encountering client NSSI behaviors or ideation—that is, which lines of inquiry about specific NSSI behaviors and online activities should be pursued. These could include examples of posts made online, examples of sites clients visit, or discussing where clients learn about NSSI or who they are talking to about it. Asking clients about NSSI and online behaviors is a necessary skill for counselors because these may not be topics adolescent clients would bring up on their own, and these topics might not be covered by standard assessment questions. This is important if clients are bilingual and if their dominant language is Spanish.

Additionally, this study points to the need to consider bilingualism with clients with NSSI issues. These clients often feel isolated and might be reaching out to online communities where they are not shamed and where they might feel more understood. The language of their distress could be seen in the frequency of their negative emotional tone or in their sadness. This implication is further highlighted by the shortage of bilingual or Spanish-speaking counselors. This study, like prior research, suggests that narration in trauma-focused therapies might be less successful, leaving counselors feeling stuck when treating these clients (Karp & Vogeleson, 2016).

A second implication for counselors would involve the Stages of Change (or Transtheoretical) Model (Prochaska & DiClemente, 1983), specifically allowing practitioners to better incorporate information about linguistic markers in pro-NSSI and recovery oriented posts. By monitoring clients' language use to indicate stages of change, and to more accurately target

interventions for progression, counselors can, therefore, provide tailored and more individualized treatment to increase recovery rates and lengths of wellness.

With increased access to online platforms and the growing number of Spanish-speaking adolescents in the United States, researchers and practitioners must continue to explore the language used by these communicants and must adapt in new ways to responsibly serve these individuals. Software such as LIWC can reveal psychological processes and patterns of behavior around a variety of mental health issues, and these results can show parents, teachers, and counselors how to intervene, perhaps also shedding insight on the process of recovery.

References

- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.
- Ashton, M. C., Lee, K., Perugini, M., Szarota, P., de Vries, R. E., Di Blas, L., . . . De Raad, B. (2004). A six-factor structure of personality-descriptive adjectives: Solutions from psycholexical studies in seven languages. *Journal of Personality and Social Psychology*, 86, 356–366. <http://dx.doi.org/10.1037/0022-3514.86.2.356>
- Bailey, C., McIntyre, E., Arreola, A., & Venta, A. (2019). What are we missing? How language impacts trauma narratives. *Journal of Child & Adolescent Trauma*, 1–9. <https://doi.org/10.1007/s40653-019-00263-3>
- Benjet, C., Gonzalez-Herrera, I., Castro-Silva, E., Mendez, E., Borges, G., Casanova, L., & Medina-Mora, M. E. (2017). Non-suicidal self-injury in Mexican young adults: Prevalence, associations with suicidal behavior and psychiatric disorders, and DSM-5 proposed diagnostic criteria. *Journal of Affective Disorders*, 215, 1–8. <https://doi.org/10.1016/j.jad.2017.03.025>
- Bentley, K. H., Nock, M. K., & Barlow, D. H. (2014). The four-function model of nonsuicidal self-injury: Key directions for future research. *Clinical Psychological Science*, 2(5), 638–656. <https://psycnet.apa.org/doi/10.1177/2167702613514563>
- Bjekić, J., Lazarević, L. B., Živanović, M., & Knežević, G. (2014). Psychometric evaluation of the Serbian dictionary for automatic text analysis-LIWCser. *Psihologija*, 47, 5–32. <https://doi.org/10.2298/PSI1401005B>
- Brezina, V. (2018). *Statistics in corpus linguistics: A practical guide*. Cambridge, UK: Cambridge University Press.

- Brown, R. C., Fischer, T., Goldwisch, A. D., Keller, F., Young, R., & Plener, P. L. (2018). #cutting: Non-suicidal self-injury (NSSI) on Instagram. *Psychological Medicine, 48*(2), 337–346. <https://doi.org/10.1017/S0033291717001751>
- Chancellor, S., Pater, J., Clear, T., Gilbert, E., & De Choudhury, M. (2016). #thyghgapp: Instagram content moderation and lexical variation in pro-eating disorder communities. *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, 1201–1213. <https://doi.org/10.1145/2818048.2819963>
- De Muth, J. E. (2014). *Basic statistics and pharmaceutical statistical applications*. Danvers, MA: Chapman and Hall/CRC.
- Department of Psychological Services. (2015). *Reformatting transcripts for LIWC 2015*. Lubbock, Texas; Author. Retrieved from: <http://www.depts.ttu.edu/psy/lusi/files/liwc2015transcriptioncleaning.pdf>
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*, 1149–1160. <http://dx.doi.org/10.3758/BRM.41.4.1149>
- Feder, P. I. (1975). The log likelihood ratio in segmented regression. *The Annals of Statistics, 3*(1), 84–97. <https://www.jstor.org/stable/2958080>
- Greaves, M. M., & Dykeman, C. (2019, February 3). A corpus linguistic analysis of public Tumblr blog posts on non-suicidal self-injury. <https://doi.org/10.31234/osf.io/k4qt3>
- Gulbas, L. E., Hausmann-Stabile, C., De Luca, S. M., Tyler, T. R., & Zayas, L. H. (2015). An exploratory study of non-suicidal self-injury and suicidal behaviors in adolescent Latinas. *American Journal of Orthopsychiatry, 84*(4), 302–314. <http://dx.doi.org/10.1037/ort0000073>

- Harned, M. S., Najavits, L. M., & Weiss, R. D. (2006). Self-harm and suicidal behavior in women with comorbid PTSD and substance dependence. *The American Journal on Addictions, 15*, 392-395. <https://doi.org/10.1080/10550490600860387>
- Hull, K., Kim, J. K., & Stillwell, M. (2018): Fotos de béisbol: An examination of the Spanish-language Instagram accounts of Major League Baseball teams. *Howard Journal of Communications, 30*, 249–264. <https://doi.org/10.1080/10646175.2018.1471756>
- Joiner, T. E., Jr., Van Orden, K. A., Witte, T. K., & Rudd, M. D. (2009). *The interpersonal theory of suicide: Guidance for working with suicidal clients*. <http://dx.doi.org/10.1037/11869-000>
- Karp, M., & Vogele, C. Does anyone still understand me? Psychotherapy and multilingualism. *Verhaltenstherapie, 26*(3), 156–157. <https://doi.org/10.1159/000448818>
- Kerr, P., Muehlenkamp, J., & Turner, J. (2010). Nonsuicidal self-injury: A review of current research for family medicine and primary care physicians. *Journal of the American Board of Family Medicine 23*(2), 240–259. <https://doi.org/10.3122/jabfm.2010.02.090110>
- Matsakis, L. (2018). How pro-eating disorder posts evade filters on social media. Retrieved from <https://www.wired.com/story/how-pro-eating-disorder-posts-evade-social-media-filters/>
- McNamara, P., McLaren, D., & Durso, K. (2007). Representation of the self in REM and NREM dreams. *Dreaming, 17*(2), 113–126. <http://dx.doi.org/10.1037/1053-0797.17.2.113>
- Moreno, M. A., Ton, A., Selkie, E., & Evans, Y. (2016) Secret society 123: Understanding the language of self-harm on Instagram. *Journal of Adolescent Health, 58*(1), 78–84. <http://dx.doi.org/10.1016/j.jadohealth.2015.09.015>

- Nock, M. K., & Favazza, A. R. (2009). Non-suicidal self-injury: Definition and classification. In M. K. Nock (Ed.), *Understanding non-suicidal self-injury: Origins, assessment, and treatment* (pp. 9–18). Washington DC: American Psychological Association.
- O’Dea, B., Larsen, M. E., Batterham, P. J., Calear, A. L., & Christensen, H. (2017). A linguistic analysis of suicide-related Twitter posts. *Crisis: The Journal of Crisis Intervention and Suicide Prevention*, *38*, 319–329. <https://doi.org/10.1027/0227-5910/a000443>
- Pennebaker, J. W. (2011). *The secret life of pronouns: What our words say about us*. New York, NY: Bloomsbury Press.
- Pennebaker, J. W., Booth, R. J., Boyd, R. L., & Francis, M. E. (2015). *Linguistic Inquiry and Word Count: LIWC2015* (Operator’s Manual). Austin, TX: Pennebaker Conglomerates, Inc.
- Pennebaker, J. W., Boyd, R. L., Jordan, K., & Blackburn, K. (2015). *The development and psychometric properties of LIWC2015*. Austin, TX: Pennebaker Conglomerates, Inc.
- Pew Research Center (2015). Demographics of Tumblr. Retrieved from http://www.pewinternet.org/2015/08/19/mobile-messaging-and-social-media-2015/2015-08-19_social-media-update_05
- Pew Research Center (2018). Social media use in 2018. Retrieved from <http://www.pewinternet.org/2018/03/01/social-media-use-in-2018>
- Prochaska, J. O., & DiClemente, C. C. (1983). Stages and processes of self-change of smoking: Toward an integrative model of change. *Journal of Consulting and Clinical Psychology*, *51*(3), 390–395.

- Rayson, P., & Garside, R. (2000). Comparing corpora using frequency profiling. *Proceedings of the Workshop on Comparing Corpora*, 9, 1–6.
<http://dx.doi.org/10.3115/1117729.1117730>
- Rosnow, R. L., & Rosenthal, R. (2003). Effect sizes for experimenting psychologists. *Canadian Journal of Experimental Psychology*, 57, 221–237. <http://dx.doi.org/10.1037/h0087427>
- Thyssen, L. S., & van Camp, I. (2014). Non-suicidal self-injury in Latin America. *Salud Mental*, 4(37), 153–157. <https://psycnet.apa.org/doi/10.17711/SM.0185-3325.2014.019>
- Wallace, K. (2015, November 3). Teens spend a ‘mind boggling’ 9 hours a day using media.
Retrieved from <https://www.cnn.com/2015/11/03/health/teens-tweens-media-screen-use-report/index.html>
- Wang, S., Eccleston, C., & Keogh, E. (2015). The role of spatial frequency information in the recognition of facial expressions of pain. *Pain*, 156, 1670–1682.
<https://doi.org/10.1097/j.pain.0000000000000226>
- Wardecker, B. M, Edelstein, R. S., Quas, J. A., Cordo, I. M, & Goodman, G. S. (2017). Emotion language in trauma narratives is associated with better psychological adjustment among survivors of childhood sexual abuse. *Journal of Language and Social Psychology*, 36(6), 628–653. <http://dx.doi.org/10.1177/0261927X17706940>
- Washburn, J. J., Richardt, S. L., Styer, D. M., Gebhardt, M., Juzwin, K. R., Yourek, A., & Aldridge, D. (2012). Psychotherapeutic approaches to non-suicidal self-injury in adolescents. *Child and Adolescent Psychiatry and Mental Health*, 6(14).
<https://doi.org/10.1186/1753-2000-6-14>
- Weisser, M. (2016). *Practical corpus linguistics: An introduction to corpus-based language analysis*. Hoboken, NJ: John Wiley & Sons, Inc.

Table 1

Nonsuicidal Self-Injury Item Word Count

	Pro- NSSI English %	Pro- NSSI English Ct.	Pro- NSSI Spanish %	Pro- NSSI Spanish Ct.	Pro- NSSI Mixed %	Pro- NSSI Mixed Ct.
Reasons for NSSI	.0288	283	.0482	2056	.221	950
NSSI Terms	.0058	57	.049	2090	.1868	803
NSSI Instruments Used	.0001	1	.014	597	.0126	54
Cutting-Specific Terms	0	0	.0005	21	0	0
Hidden-Hashtag Terms	0	0	0	0	0	0

Table 2

English Pronoun Word Count

	Pro- NSSI English %	Pro- NSSI English Ct.	Blog Norms English %	Blog Norms English Ct.	Total %	Total Ct.
All Words	75.37	9814	24.63	3206	100	13020
First-Person Singular	9.89	971	6.26	201	9.00	1171
Third-Person Singular	0.50	49	1.50	48	0.75	97
Third-Person Plural	0.53	52	0.68	22	0.57	74
First-Person Plural	0.25	25	0.91	29	0.41	54

Table 3

English Pronoun Test Results

Pronoun	Pro- NSSI English Obs. Ct.	Pro- NSSI English Exp. Ct.	Blog Norms English Obs. Ct.	Blog Norms English Exp. Ct.	Family- Wise Error Rate	Adj. Error Rate	<i>df</i>	<i>cv</i>	<i>LL</i>	<i>BIC</i>
First- Person Singular	971	882.63	201	288.37	0.05	0.003	1	8.6	40.21	30.74
Third- Person Singular	49	73.11	48	23.89	0.05	0.003	1	8.6	27.78	18.3
First- Person Plural	25	40.7	29	13.3	0.05	0.003	1	8.6	20.85	11.38
Third- Person Plural	52	55.78	22	18.22	0.05	0.003	1	8.6	1	-8.48

Table 4

Spanish Psychological Process Category Test Results

Category	Pro- NSSI Span. Obs. Ct.	Pro- NSSI Span. Exp. Ct.	Blog Norms Eng. Obs. Ct.	Blog Norms Exp. Ct.	Family- Wise Error Rate	Adj. Error Rate	<i>df</i>	<i>cv</i>	<i>LL</i>	<i>BIC</i>
Neg. E.	4049	3827.33	66	287.67	0.05	0.0036	1	8.49	261.61	250.87
Sadness	1898	1779.27	14	133.73	0.05	0.0036	1	8.49	182.03	171.29
Cog. Pro.	9245	8943.78	371	672.22	0.05	0.0036	1	8.49	171.44	160.71
Health	1851	1740.2	20	130.8	0.05	0.0036	1	8.49	153.38	142.65
Death	1135	1060.31	5	79.69	0.05	0.0036	1	8.49	126.84	116.11
Anger	1506	1421.18	22	106.82	0.05	0.0036	1	8.49	105.08	94.34
Anxiety	431	409.24	9	30.76	0.05	0.0036	1	8.49	22.53	11.8
Insight	1566	1524.42	73	114.58	0.05	0.0036	1	8.49	18.46	7.73
Body	678	652.93	24	49.07	0.05	0.0036	1	8.49	16.77	6.03
Ingestion	375	365.53	17	27.47	0.05	0.0036	1	8.49	2.87	-7.86

Table 5

English Psychological Process Category Test Results

Category	Pro- NSSI Eng. Obs. Ct.	Pro- NSSI Eng. Exp. Ct.	Blog Norms Eng. Obs. Ct.	Blog Norms Exp. Ct.	Family- Wise Error Rate	Adj. Error Rate	<i>df</i>	<i>cv</i>	<i>LL</i>	<i>BIC</i>
Health	464	364.81	20	119.19	0.05	0.0033	1	8.62	151.79	142.32
Neg. E.	604	505	66	165	0.05	0.0033	1	8.62	95.29	85.82
Sadness	218	174.87	14	57.13	0.05	0.0033	1	8.62	56.75	47.27
Death	96	76.13	5	24.87	0.05	0.0033	1	8.62	28.49	19.01
Anger	158	135.67	22	44.33	0.05	0.0033	1	8.62	17.32	7.84
Anxiety	94	77.63	9	25.37	0.05	0.0033	1	8.62	17.31	7.84
Cog. Pro.	1339	1288.89	371	421.11	0.05	0.0033	1	8.62	8.14	-1.34
Body	109	100.25	24	32.75	0.05	0.0033	1	8.62	3.32	-6.15
Insight	265	254.76	73	83.24	0.05	0.0033	1	8.62	1.72	-7.75
Ingestion	65	61.81	17	20.19	0.05	0.0033	1	8.62	.7	-8.78

Table 6

Mixed Psychological Process Category Test Results

Category	Pro-NSSI Mixed Obs. Ct.	Pro-NSSI Mixed Exp. Ct.	Blog Norms Eng. Obs. Ct.	Blog Norms Exp. Ct.	Family- Wise Error Rate	Adj. Error Rate	<i>df</i>	<i>cv</i>	<i>LL</i>	<i>BIC</i>
Neg. E.	1519	907.86	66	677.14	0.05	0.0036	1	8.49	1256.38	1247.46
Sadness	1131	655.84	14	489.16	0.05	0.0036	1	8.49	1133.16	1124.24
Death	910	524.1	5	390.9	0.05	0.0036	1	8.49	960.63	951.7
Health	627	370.59	20	276.41	0.05	0.0036	1	8.49	554.37	545.44
Anger	554	324.2	22	241.8	0.05	0.0036	1	8.49	457.68	448.76
Cog. Pro.	261	362.57	371	270.43	0.05	0.0036	1	8.49	63.03	54.11
Anxiety	98	61.29	9	45.71	0.05	0.0036	1	8.49	62.75	53.82
Ingestion	50	38.95	17	29.05	0.05	0.0036	1	8.49	6.76	-2.17
Body	58	46.97	24	35.03	0.05	0.0036	1	8.49	6.32	-2.6
Insight	126	113.98	73	85.02	0.05	0.0036	1	8.49	3.01	-5.91

Chapter 3: A Research Manuscript

A Corpus Linguistic Analysis of Pro-Anorexia Public Tumblr Posts Which Use Spanish (Un
Analisis de Lingüística de Corpus de Publicaciones Pública que Usan Español en Tumblr que
Respaldan a la Pró-Anorexia)

Kristeen E. Elrod and Cass Dykeman
Oregon State University

Author Note

Kristeen E. Elrod, Counseling Academic Unit, Oregon State University; Cass Dykeman, Counseling Academic Unit, Oregon State University.

The research contained in this manuscript was conducted under the approval of the Oregon State University Institutional Review Board (Study ID No. 8720) and was part of the first author's dissertation research project.

Correspondence concerning this article should be addressed to Cass Dykeman, Counseling Academic Unit, Oregon State University, College of Education, Joyce Collin Furman Hall, 200 SW 15th Street, Corvallis, OR 97331. Contact: dykemanc@onid.oregonstate.edu

Abstract

Eating disorders have the highest mortality rate of any mental illness, affecting all races and ethnicities and with symptoms typically beginning in adolescence and early adulthood. Still, little is known about the language markers of mental health in content using Spanish language. This study collected pro-anorexia (pro-ana) public Tumblr blog posts written using Spanish and analyzed their content using Linguistic Inquiry and Word Count software. Pro-ana-specific words, pronoun use, and psychological linguistic properties were analyzed, and results indicated that *reasons for anorexia* was the most frequently occurring category within pro-ana posts. Words in this category were associated with intense and negative emotions commonly found in trauma narratives in the dominant language (Spanish). Users' pronoun use suggested an increased focus on people and objects, which has been linked to decreased likelihood of recovery. The psychological properties of the posts had a negative emotional tone and were marked by significant use of ingestion and health words, signaling the symptoms of anorexia. This research points to benefits in assessing bilingual clients for anorexia behaviors, symptoms, and online-posting behaviors, while also revealing the need for improved evidence-based treatment for trauma in bilingual individuals, as well as the need for Spanish and code-switching blog norms to inform and advance future research.

Keywords: pro-ana, LIWC, corpus linguistics, Tumblr

Palabras Claves: pró-anorexia, LIWC, lingüística de corpus, Tumblr

A Corpus Linguistic Analysis of Pro-Anorexia Public Tumblr Posts Written in Spanish (Un Estudio de Lingüística de Corpus de Publicaciones Públicas de Tumblr a las Pró-Anorexia en Español)

User engagement with social media has grown considerably and is expected to keep growing in the coming years. Nearly four billion people worldwide have Internet access, and the vast majority maintain some form of social media involvement, with the median user engaging on three different platforms (Pew Research Center, 2018). Social media outlets can provide a space for free and honest communication and connection with others without stigma or consequences. Mental health-related conversations occur frequently within this space (Gaur et al., 2018). Online communities use distinctive language to define themselves and to distinguish insiders from outsiders, but pro-anorexia (pro-ana) groups, in particular, can find their efforts discouraged by policies of social media platforms that block or ban certain words or hashtags central to a community's sense of identity. When certain options are blocked or banned, users make their language choices more complex and vary their spellings or renderings to conceal their activities and to maintain solidarity within the group. Because of such obfuscation, pro-ana group members can evade scrutiny, becoming more difficult to reach for parents, teachers, and clinicians and thereby finding themselves limited in terms of access to care.

From a review of the literature on pro-ana communities' public posts on social media, five themes stand out. These include the significant rates of anorexia in American society, the prevalence of anorexia in Hispanic communities, the frequency of social media use, the presence of pro-ana communities on social media, and the lack of education, awareness, or training for parents, teachers, and clinicians with respect to pro-ana terms written in Spanish. Following an explication of these themes, the research questions driving the present study will be presented.

Eating disorders are complex biopsychosocial illnesses; they can lead to serious health consequences, and they register the highest mortality rate of any psychiatric disorder (Arseniev-Koehler, Lee, McCormick, & Moreno, 2016). In the United States, 2.7% of those 13 to 18 years old experience eating disorders, and an additional 3.3% struggle with subthreshold eating disorder symptoms or behaviors (Merikangas et al., 2010; Swanson, Crow, Legrange, Swendsen, & Merikangas, 2011). Those experiencing anorexia often attempt to conceal this disorder and the behaviors associated with it, making this population difficult to reach; given the anonymity of social media, however, those with anorexia have an outlet for social support and resources.

While studies have addressed eating disorders as experienced by Caucasians, the Hispanic community has often been underrepresented in research on anorexia, perpetuating the misconception that the illness is one experienced mostly by Caucasian women, especially those living in the Western world. Prevalence rates are consistent across Latin American countries; however, Hispanics in North America, and especially second-generation Hispanics, experience rates on par with their Caucasian counterparts (Kolar, Mejia-Rodriguez, Chams, & Hoek, 2016).

Immigration status and level of acculturation are unique factors affecting many Hispanics in the United States (PEW Hispanic Center, 2018). Eating disorders arise from the adoption of American beauty standards and the balancing of multiple cultures (Perez, Ohrt, & Hoek, 2016). Food, culture, and identity are intertwined within the Hispanic community: Meals are social affairs, food is a way to express love, and women are taught to cook as a way of contributing to the family and household. Moreover, while stigma attaches to mental illness throughout the United States, the phenomenon is heightened in the Hispanic community, making members less likely to seek services (Kolar et al., 2016). With the growing Hispanic population in the United States, and considering the risk factors facing future generations, research on pro-ana social

media communication by Spanish speakers is crucial to counselors' efforts to engage and treat members of this community.

Adolescents are regular social media users notoriously difficult to connect with when it comes to health issues. Adolescents are also relatively less likely to take advantage of health care services, often preferring to gather information and seek support via the Internet (Wong, Merchant, & Moreno, 2014). Finding a group of similarly situated individuals online can provide young people with a sense of relief and encouragement; but, unfortunately, some pro-ana groups promote self-harm and the continuation of dangerous symptoms of anorexia as a lifestyle choice, encouraging their members to continue behaviors associated with anorexia and leading to devastating effects on health and relationships (Chancellor, Kalantidis, Pater, De Choudhury, & Shamma, 2017). Those suffering from anorexia can be identified online by their use of certain key words characterizing and promoting their condition (Arseniev-Koehler et al., 2016; Want, Brede, Ianni, & Mentzakis, 2017). Language plays an important role in signaling a shared identity, and the words that groups use are constantly evolving and being developed by members.

While anorexia is not a new phenomenon, pro-ana online communities did not become popular until the early 2000s. With their increasing members and reach, pro-ana communities have exacerbated problems associated with poor body image, have encouraged unhealthy methods of weight loss, and have established or reinforced conceptions of identity around anorexia (Chancellor et al., 2016). Social media outlets have influenced the behavior and choices of adolescents and young adults who could have ongoing mental health issues, including anorexia. Pro-ana behavior on social media is of particular interest given the combination of peer influence, social learning, and media that coexist in these communities (Moreno, Ton, Selkie, & Evans, 2016), and the negative potential of social media is amplified by the fact that teens and

young adults are those most likely to go online, spending up to nine hours per day with digital technology, an amount of time that could exceed the number of hours they sleep on a nightly basis (Wallace, 2015).

Tumblr is a preferred social media platform for teens and young adults (Pew Research Center, 2015). Even more, for the purposes of this study, Hispanics are among the most prevalent users of the platform, interacting with Tumblr more than they do with any other social media outlet (Hull, Kim, & Stillwell, 2018). Tumblr has been criticized for the continued presence of pro-ana communities—communities that continue to thrive despite the site’s efforts to break up selected groups, to show warnings about content, and to block certain hashtags and words (Mastakis, 2018). An imposition on an online community could be dangerous to members; users might rely on such groups for support, especially considering the isolating nature of anorexia, a disease where individuals often suffer in secrecy or hide their activities. An understanding community and online place of belonging could, therefore, have a healing effect (Joiner, Van Orden, Witte, & Rudd, 2009). That said, pro-anorexia communities can also trigger anorexia urges, can normalize anorexia-related activities, and can reinforce negative behaviors to the point where the harm far outweighs any perceived benefits for pro-ana members (Brown et al, 2018).

One of the central problems with pro-ana communities on social media platforms is that they promote behaviors associated with anorexia and depict anorexia as a lifestyle choice rather than an illness (Chancellor et al., 2016). This misrepresentation of the disorder is facilitated by the visual nature of Tumblr, whereby pro-ana communities can disseminate images ostensibly representing “ideal” body shapes while promoting a warped notion of “health” that the groups refer to as *thinspo* or *thinspiration*, terms that refer to images, slogans, videos, and other

materials that inspire the pursuit of extreme thinness. Within this community, such a degree of thinness exemplifies hard work, willpower, control, dedication, and perfection. The large majority of thinspo images include white women, thus racializing the issue, layering on assumptions of status and privilege, and creating a false dichotomy where being thin means being in control and being a curvaceous “other” is construed as giving in to bodily desires. Racial hierarchies are associated with beauty, and thinness works as a means of improving social status and enabling the navigation of elite spaces for Hispanic and African American communities (Schott, 2017).

There is a gap in the literature around linguistic and psychological properties of language used in the pro-ana community online, specifically regarding materials written in Spanish. Linguistic Inquiry and Word Count (LIWC—pronounced “Luke”) software can evaluate many kinds of written text, and in this study it was used to analyze the language specific to the pro-ana community (Pennebaker, Booth, Boyd, & Francis, 2015). Tumblr was chosen as the social media platform for the present research because of its popularity among teens and young adults, groups whose approximate age (13–24) coincides with the typical age of onset for anorexia. With the continuing growth of the Hispanic population in the United States, understanding the characteristics of language deployed in public Spanish-speaking pro-ana online posts will encourage greater social justice and will aid in the addressing of mental health issues facing this population (Gulbas, Hausmann-Stabile, De Luca, Tyler, & Zayas, 2015).

With the abovementioned literature as a foundation, this study was organized around four research questions. First, in public pro-ana Tumblr posts written in Spanish, what was the use rate of pro-ana word categories? Second, in public pro-ana Tumblr posts written in English, what was the use rate of pronoun categories? Third, in public pro-ana Tumblr posts written in English,

did the use rate of pronoun categories differ from norms for such posts? Fourth, in public pro-ana Tumblr posts written in Spanish, did the use rate of the psychological process categories differ from norms for such posts?

Method

Design

This study employed a synchronic corpus linguistic design (Weisser, 2017) with single words as the unit of analysis (Bjekić, Lazarević, Živanović, & Knežević, 2014). The variables for the first research question included the following word categories: methods of anorexia, anorexia-specific terms, reasons for anorexia, and hidden-hashtag terms. The variables for the second and third research questions included the categories of first-person singular pronoun, third-person singular pronoun, first-person plural pronoun, and third-person plural pronoun. Finally, the variables for the fourth research question were psychosocial and physical categories including (a) negative emotion, (b) anxiety, (c) anger, (d) sadness, (e) cognitive process, (f) insight, (g) body, (h) health, (i) ingestion, and (j) death.

All variables were reported as a percentage of total words, and there were three sub-corpora in this study: (a) Spanish-only, (b) English-only, and (c) Mixed. In terms of the third and fourth research questions, the comparison norm for all corpora was the assemblage of Twitter blog norms in English. There are currently no Spanish or mixed-blog norms to serve as respective baselines for the Spanish-only and Mixed corpora.

The inferential analysis used in this study was a log-likelihood (LL) ratio test. This technique tests for goodness-of-fit to a hypothesized distribution. A priori power analysis for a χ^2 test was employed using G*Power 3.1, because the LL ratio can be approximated by Pearson's χ^2 (Faul, Erdfelder, Buchner, & Lang, 2009). Cohen's w is the proper effect size (Rosnow &

Rosenthal, 2003). Given the planned number of comparisons, the α level was set at .0033, and the input parameters were: (a) test family = χ^2 tests, (b) statistical test = goodness-of-fit tests: contingency tables, (c) type of analysis = a priori: Compute required sample size given α , power, and effect size, (d) $w = 0.52$, (e) power (1- β error probability) = 0.9, (f) $\alpha = .0033$, and (g) degrees of freedom = 1. The G*Power 3.1 output included a sample size of 78 and an actual power of 0.90.

Corpus

Register and scope. The register was computer-mediated communication and the sub-register was microblogging about anorexia with Spanish hashtags and captions. The researcher used Tumblr's Application Programming Interface (API) to collect public posts between October 23, 2017 and October 23, 2018. The API was instructed to scrape posts meeting three conditions: (a) the caption and/or hashtag included Spanish language, (b) the post used hashtags identifying the content as pro-ana (e.g., #hastaloshuesos signified *to the bone*), and (c) the post emanated from a North, Central, or South American country. While collecting text, the API was instructed to eliminate usernames, URLs, location, names, photographs, comments left by other users, and direct quotes. During the range of time noted above, 1,165 posts and associated hashtags were collected (<https://www.tumblr.com/docs/en/api/v2>).

Construction. The scraped posts were compiled into an Excel document to be cleaned. Posts (caption and hashtags) using any language other than English or Spanish were excluded, as were those composed entirely in English. The corpus, which contained posts in Spanish, posts where whole sections were in English or Spanish, and posts with mixed-language code-switching (switching between English and Spanish in the same section), were then separated into three sub-corpora for analysis at the section level. These sub-corpora were English-only, which included

sections of the post (either caption or hashtag) that were written entirely in English, while the rest of the post contained Spanish language; Spanish-only, which included sections of the post where either the caption or hashtags were written entirely in Spanish; and Mixed, which included sections of the post where code-switching occurred in the caption or hashtag.

Preprocessing. Posts were cleaned using the procedures outlined in the LIWC operator's manual and supplement (Department of Psychological Services, 2015; Pennebaker, Booth et al., 2015), meaning all abbreviations were spelled out (e.g., “max” became *maximum*), “anorexia slang” (words used to hide anorexia-related Tumblr posts) was not normalized (Moreno et al., 2016), and “textese” was translated into standard spelling. The word count was sufficient for both power and reliability requirements (Faul et al., 2009; Wardecker, Edelstein, Quas, Cordon, & Goodman, 2017). Once the cleaning procedures were concluded, the file was processed by LIWC software.

Size. The size of the Spanish-only sub-corpus was 3,234 tokens, the size of the English-only sub-corpus was 46,921 tokens, and the Mixed sub-corpus amounted to 531 tokens.

Measures

The measures used were selected scales from the LIWC scales and the Elrod Pro-Ana Linguistic Scales (EPALS), the first author's Spanish adaptation of the Greaves NSSI Linguistic Scales (Greaves & Dykeman, 2019). Scores for all variables represented a percentage of all words used. The EPALS scales were developed for this study; the LIWC scales were pre-set scales contained in the LIWC software. Pennebaker, Boyd, Jordan, and Blackburn (2015) reported adequate reliability and validity for these LIWC scales. The reference norms used for comparison to the present results were the Twitter norms contained in the LIWC psychometric manual (Pennebaker, Boyd et al., 2015). Internal reliability and external validity of LIWC were

well-documented (Pennebaker, Boyd et al., 2015), and four anorexia specific measures were employed.

EPALS-Methods of Anorexia. Methods of anorexia are ways people engage in behaviors associated with anorexia. This is a broad category encompassing many types of anorexia-related behaviors, such as calorie restrictions, over-exercise, measuring oneself, and weighing oneself (in Spanish as *nocomer*, *vomitar*, and *ejercicio*).

EPALS-Anorexia Specific Terms. Anorexia is a specific diagnosis under the category of eating disorders, and examples within this grouping included words such as *thighgap*, *thinspo*, and *skinny* (in Spanish as *flaco*, *delgado*, and *huesos*).

EPALS-Reasons for Anorexia behaviors. There are many reasons why people choose to engage in anorexic behaviors, such as shame, anger, anxiety, attention, control, and guilt (in Spanish as *triste*, *ansiedad*, or *enojo*).

EPALS-Hidden Hashtag Terms. To evade new social media restrictions, those who engage in anorexic behaviors and who desire to post about it on social media, have had to become creative in the hashtags they use to communicate with other pro-ana individuals and groups. Some of these users' "hidden" hashtags include *#princesasanaymia* and *#putavidaa*.

Pronouns

LIWC was employed to assess the frequency of the following four pronoun types: first-person singular, third-person singular, first-person plural, and third-person plural.

Psychological Processes

LIWC was also used to assess 10 psychological processes. These processes, with representative words included in parentheses, included negative emotion (*hurt*), anxiety (*fearful*),

anger (kill), sadness (grief), cognitive processes (cause), insight (think), body (spit), health (clinic), ingestion (eat), and death (bury).

Comparison Norms

The baseline norm for the three sample corpora was the set of English Twitter norms contained in the LIWC psychometric manual (Pennebaker, Boyd et al., 2015). There are currently no Spanish or Mixed blog norms available for comparison. English and Spanish are syntactically dissimilar languages, but it is reasonable to assume a 1:1 correspondence between English nouns and Spanish nouns in terms of word count (Bailey, McIntyre, Arreola, & Venta, 2019); thus, the second sub-corpus, Spanish-only, was compared to English Twitter norms for psychological process categories.

The entire corpus for this study was analyzed with LIWC software (Pennebaker, Booth, et al., 2015). LIWC analysis returns results for 90 output variables (i.e., LIWC categories), and LIWC category scores represent the percentage of total words used for the sample. Words could be part of more than one category. The English LIWC2015 dictionary was used for the English sub-corpus; the Spanish LIWC2007 dictionary was used for the Spanish sub-corpus; and, both the English LIWC2015 and Spanish LIWC2007 dictionaries were employed for the Mixed sub-corpus, pursuant to Pennebaker's recommendation (personal communication, October 30, 2018). New dictionaries were available in LIWC2015, with more cognitive-process and social words included (Pennebaker, Booth, et al., 2015).

Data Analysis

For the first and second research questions, and for each variable, raw count and percentage of total words used have been reported below. For the third and fourth research questions, the sub-corpora were compared to norms by way of the LL test (Rayson & Garside,

2000). The Bayesian Information Criterion (BIC) was the effect-size measure reported (Brezina, 2018), and analyses were completed using *R*. Given the large number of hypothesis tests conducted, a Bonferroni correction was used to reduce the chance of a Type I error.

Results

The first research question focused on identifying word-use rates, specifically pro-ana word categories, in three groups of sub-corpora. Across all sub-corpora, the most frequently used pro-ana-specific category was *reasons for anorexia*. Complete results can be found in Table 1. With respect to the second and third research questions, the linguistic process category that differed most from English Twitter blog norms was *third-person singular* (LL = 49.13). The effect size for this category was BIC = 38.31, where anything greater than 10 was considered large.

The linguistic process category of *first-person plural* also differed from English Twitter blog norms (LL = 46.27), registering an effect size of BIC = 35.45. Likewise, the linguistic process category of *third person plural* differed from the English Twitter blog norms (LL = 11.42), showing an effect size of BIC = .6. The results for these research questions can be found in Tables 2 and 3. Finally, regarding the fourth research question, in both the English sub-corpus and the Spanish sub-corpus, the psychological process category that differed most from English Twitter blog norms was *ingestion* (LL = 485.14, BIC = 474.32, for the English sub-corpus; LL = 135.71, BIC = 126.94 for the Spanish sub-corpus). In the Mixed sub-corpus, by contrast, the psychological process category that differed most from English Twitter blog norms was *health* (LL = 316.68, BIC = 308.46), followed by *ingestion* (LL = 302.09, BIC = 293.87). The complete results for this research question can be found in Tables 4, 5, and 6.

Discussion

This study examined language used in public pro-ana Tumblr blog posts written in Spanish, relying on four research questions to discern and analyze differences in the use of language in Spanish pro-ana blogs compared to blogs overall. The results from the first research question revealed that the highest use rate across sub-corpora was *reasons for anorexia*. There are two potential explanations for this result. The first is that words included in the *reasons for anorexia* category (e.g., hate, control, depression) are associated with negative emotions, a sentiment that is likely to be more prominent in the dominant language (Spanish) while individuals are in the midst of retrieving traumatic memories and in the course of offering explanations or justifications for their behavior. Specifically, anorexia behaviors are a way for individuals to manage their distress and to avoid other social forms of expressed anger (Espeset, Gulliksen, Nordbo, Skarderud, & Holte, 2012). Alternatively, it could be that users were undergoing a change and indicating openness to recovery, given that engagement with emotional expression is linked to signs of recovery. Between these two explanations, the former is most compelling, because research has demonstrated a link between traumatic experiences, such as childhood sexual abuse and eating disorder behaviors.

The results from the second and third research questions revealed the pronoun use rate differed from English Twitter norms most frequently and significantly in the category of *third-person singular*, followed by *first-person plural* and *third-person plural*. There are two potential explanations for the obtained results. Consistent with results from other studies, those who exhibit less attention on themselves (for example, using the word *we* more than the word *I*) are more popular in a social community (Wang, Brede, Ianni, & Mentzakis, 2018). Gaining “likes” and followers could be a contributing factor in posting behavior of members within this

community, especially given that acceptance and encouragement are part of the allure of membership in the group. An alternative explanation is that, because this is a pro-ana community, it is reasonable to expect that language will reflect a greater focus on people and objects, an effect that has been linked to a decrease in likelihood of recovery. Between these explanations, the latter is most likely, because other research has shown that, over time, only half of users posting about anorexia in English on Tumblr will have a “recovery event” (Chancellor, Mitra, & De Choudhury, 2016).

For the fourth research question, where the most common and statistically significant psychological process category in both the English and Spanish sub-corpora was *ingestion* and where the most common category in the Mixed sub-corpus was *health* followed by *ingestion*, there are two probable explanations for the obtained results. The first is that the results are consistent with the focus on body, health, and ingestion revealed by prior research on English-dominant pro-ana groups online. An alternative explanation is that users’ posts were more likely to contemplate recovery. Research has shown that the increased use of words such as *health* and *ingestion* on Tumblr posts was linked to a greater likelihood of recovery, with the increased use of health-related words being the most likely predictor (Chancellor et al., 2016). Between these explanations, the former is more convincing, because pro-ana users who indicate concerns about body image and ingestion, especially when they express themselves with a more negative tone, are more likely to be signaling symptoms of anorexia (Wang et al., 2018). These results are compatible with prior research that pro-ana communities exacerbate and reinforce harmful ideals about weight and weight-loss practices.

While these findings are robust, there are three limitations to this study that warrant consideration. First, words in LIWC can exist within more than one category and can have more

than one meaning. This could skew the results to some degree. Before being analyzed by LIWC software, the data were reviewed to clarify any confusing meanings and to control for this concern as much as possible. Second, this study considered neither private posts nor posts originating from beyond North, Central, or South American countries. Because of the practice of banning hashtags and blocking content on social media sites, it is possible that users have private accounts where they screen followers to ensure their content will not be reported by other users. This sort of decision would limit the available data and could have affected the results. Third, this study began by focusing on a Spanish-language corpus. Due to the code-switching phenomenon, however, it ended up including Spanish, English, and Mixed sub-corpora. There were no Spanish-language or mixed-language blog norms to use for comparison, thereby constraining the external validity of the study.

The above limitations notwithstanding, this study reveals significant implications for both practitioners and researchers. With respect to practitioners, this study suggests areas of additional inquiry when encountering a Spanish-speaking or bilingual client engaging in behaviors associated with anorexia. Counselors must be equipped to speak with their clients about anorexia, especially in the case of adolescents, who are less likely to raise the topic on their own. Gathering more information about online behaviors, trauma history, and engagement with emotions would reveal even more opportunities to assess the needs of these clients.

A second implication for counselors would involve the Stages of Change (or Transtheoretical) Model (Prochaska & DiClemente, 1983), specifically allowing practitioners to better incorporate information about linguistic markers in pro-ana and pro-recovery posts. Counselors are committed to facilitating health and wellness, and research continues to demonstrate linguistic markers among groups. By monitoring clients' language use to indicate

stages of change, and to more accurately target interventions for progression, counselors can, therefore, provide tailored and more individualized treatment to increase recovery rates and lengths of wellness. Finally, and regarding research implications, this study emphasized the ongoing need for a deeper understanding of the relationship between language and anorexia behaviors, particularly in Spanish-speaking and bilingual individuals. To generate more and better data in this domain, and to cultivate a more refined and inclusive research agenda, scholars must develop Spanish blog norms and code-switching corpora.

References

- Arseniev-Koehler, A., Lee, H., McCormick, T., & Moreno, M. A. (2016). #Proana: Pro-eating disorder socialization on Twitter. *Journal of Adolescent Health, 58*, 659–664.
<https://doi.org/10.1016/j.jadohealth.2016.02.012>
- Bjekić, J., Lazarević, L. B., Živanović, M., & Knežević, G. (2014). Psychometric evaluation of the Serbian dictionary for automatic text analysis-LIWCser. *Psihologija, 47*, 5–32.
<https://doi.org/10.2298/PSI1401005B>
- Brown, R. C., Fischer, T., Goldwich, A. D., Keller, F., Young, R., & Plener, P. L. (2018). #cutting: Non-suicidal self-injury (NSSI) on Instagram. *Psychological Medicine, 48*(2), 337–346. <https://doi.org/10.1017/S0033291717001751>
- Chancellor, S., Kalantidis, Y., Pater, J., De Choudhury, M., & Shamma, D. A. (2017). *Multimodal classification of moderated pro-eating disorder content*. Paper presented at the 2017 CHI Conference. <http://dx.doi.org/10.1145/3025453.3025985>
- Chancellor, S., Lin, Z., Goodman, E. L., Zerwas, S., & Choudhury, M. D. (2016). Quantifying and predicting mental illness severity in online pro-eating disorder communities. *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, 1171–1184. <https://dx.doi.org/10.1145/2818048.2819973>
- Chancellor, S., Mitra, T., & De Choudhury, M. (2016). Recovery amid pro-anorexia: Analysis of recovery in social media. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 2111–2123. <https://doi.org/10.1145/2858036.2858246>
- De Muth, J. E. (2014). *Basic statistics and pharmaceutical statistical applications*. Danvers, MA: Chapman and Hall/CRC.

Department of Psychological Services. (2015). *Reformatting transcripts for LIWC 2015*.

Lubbock, TX: Author. Retrieved from:

<http://www.depts.ttu.edu/psy/lusi/files/liwc2015transcriptioncleaning.pdf>

Espeset, E. M. S., Gulliksen, K. S., Nordbo, R. H. S., Skarderud, F., & Holte, A. (2012). The link between negative emotions and eating disorder behavior in patients with anorexia nervosa. *European Eating Disorders Review*, *20*(6), 451–460.

<https://doi.org/10.1002/erv.2183>

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*, 1149–1160. <http://dx.doi.org/10.3758/BRM.41.4.1149>

Gaur, M., Kursuncu, U., Alambo, A., Sheth, A., Daniulaityte, R., Thirunarayan, K., & Pathak, J. (2018). “Let me tell you about your mental health!” Contextualized classification of Reddit posts to DSM-5 for web-based intervention. *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*, 753–762.

<https://doi.org/10.1145/3269206.3271732>

Greaves, M. M., & Dykeman, C. (2019, February 3). A corpus linguistic analysis of public Tumblr blog posts on non-suicidal self-injury. <https://doi.org/10.31234/osf.io/k4qt3>

Gulbas, L. E., Hausmann-Stabile, C., De Luca, S. M., Tyler, T. R., & Zayas, L. H. (2015). An exploratory study of non-suicidal self-injury and suicidal behaviors in adolescent Latinas. *American Journal of Orthopsychiatry*, *84*(4), 302–314.

<http://dx.doi.org/10.1037/ort0000073>

- Hull, K., Kim, J. K., & Stillwell, M. (2018): Fotos de béisbol: An examination of the Spanish-language Instagram accounts of Major League Baseball teams. *Howard Journal of Communications*, 30, 249–264. <https://doi.org/10.1080/10646175.2018.1471756>
- Joiner, T. E., Jr., Van Orden, K. A., Witte, T. K., & Rudd, M. D. (2009). *The interpersonal theory of suicide: Guidance for working with suicidal clients*.
<http://dx.doi.org/10.1037/11869-000>
- Kolar, D. R., Mejia Rodriguez, D. L., Mebarak Chams, M., & Hoek, H. W. (2016).
Epidemiology of eating disorders in Latin America: A systematic review and meta-analysis. *Current Opinion in Psychiatry*, 29(6), 363–371.
<http://dx.doi.org/10.1097/YCO.0000000000000279>
- Matsakis, L. (2018). How pro-eating disorder posts evade filters on social media. Retrieved from
<https://www.wired.com/story/how-pro-eating-disorder-posts-evade-social-media-filters/>
- McNamara, P., McLaren, D., & Durso, K. (2007). Representation of the self in REM and NREM dreams. *Dreaming*, 17(2), 113–126. <http://dx.doi.org/10.1037/1053-0797.17.2.113>
- Merikangas, K. R., He, J., Burstein, M., Swanson, S. A., Avenevoli, S., Cui, L., . . . Swendsen, J. (2010). Lifetime prevalence of mental disorders in U.S. adolescents: Results from the National Comorbidity Study-Adolescent Supplement (NCS-A). *Journal of the American Academy of Child and Adolescent Psychiatry*, 49(10), 980–989. <http://dx.doi.org/10.1016/j.jaac.2010.05.017>
- Moreno, M. A., Ton, A., Selkie, E., & Evans, Y. (2016) Secret society 123: Understanding the language of self-harm on Instagram. *Journal of Adolescent Health*, 58(1), 78–84.
<http://dx.doi.org/10.1016/j.jadohealth.2015.09.015>

- O'Dea, B., Larsen, M. E., Batterham, P. J., Calear, A. L., & Christensen, H. (2017). A linguistic analysis of suicide-related Twitter posts. *Crisis: The Journal of Crisis Intervention and Suicide Prevention*, 38, 319–329. <https://doi.org/10.1027/0227-5910/a000443>
- Pennebaker, J. W. (2011). *The secret life of pronouns: What our words say about us*. New York, NY: Bloomsbury Press.
- Pennebaker, J. W., Booth, R. J., Boyd, R. L., & Francis, M. E. (2015). *Linguistic Inquiry and Word Count: LIWC2015* (Operator's Manual). Austin, TX: Pennebaker Conglomerates, Inc.
- Pennebaker, J. W., Boyd, R. L., Jordan, K., & Blackburn, K. (2015). *The development and psychometric properties of LIWC2015*. Austin, TX: Pennebaker Conglomerates, Inc.
- Perez, M., Ohrt, T. K., & Hoek, H. W. (2016). Prevalence and treatment of eating disorders among Hispanics/Latino Americans in the United States. *Current Opinion in Psychiatry*, 29(6), 378–382. <https://doi.org/10.1097/YCO.0000000000000277>
- Pew Hispanic Center (2018). Hispanic trends. Retrieved from <http://www.pewhispanic.org>
- Pew Research Center (2015). Demographics of Tumblr. Retrieved from http://www.pewinternet.org/2015/08/19/mobile-messaging-and-social-media-2015/2015-08-19_social-media-update_05
- Pew Research Center (2018). Social media use in 2018. Retrieved from <http://www.pewinternet.org/2018/03/01/social-media-use-in-2018>
- Prochaska, J. O., & DiClemente, C. C. (1983). Stages and processes of self-change of smoking: Toward an integrative model of change. *Journal of Consulting and Clinical Psychology*, 51(3), 390–395.

- Rayson, P., & Garside, R. (2000). Comparing corpora using frequency profiling. *Proceedings of the Workshop on Comparing Corpora*, 9, 1–6.
<http://dx.doi.org/10.3115/1117729.1117730>
- Rosnow, R. L., & Rosenthal, R. (2003). Effect sizes for experimenting psychologists. *Canadian Journal of Experimental Psychology*, 57, 221–237. <http://dx.doi.org/10.1037/h0087427>
- Schott, N. D. (2017). Race, online space and the feminine: Unmapping “black girl thinspiration.” *Critical Sociology*, 43(7-8), 1029–1043. <http://dx.doi.org/10.1177/0896920516652456>
- Swanson, S. A., Crow, S. J., Le Grange, D., Swendsen, J., & Merikangas, K. R. (2011). *Archives of General Psychiatry*, 68(7), 714–23.
<http://dx.doi.org/10.1001/archgenpsychiatry.2011.22>
- Tausczik, Y. R., & Pennebaker, J. W. (2009). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology*, 29, 24–54. <https://dx.doi.org/10.1177/0261927x09351676>
- Wallace, K. (2015, November 3). Teens spend a ‘mind boggling’ 9 hours a day using media. Retrieved from <https://www.cnn.com/2015/11/03/health/teens-tweens-media-screen-use-report/index.html>
- Wang, S., Eccleston, C., & Keogh, E. (2015). The role of spatial frequency information in the recognition of facial expressions of pain. *Pain*, 156, 1670–1682.
<https://doi.org/10.1097/j.pain.0000000000000226>
- Wang, T., Markus, B., Ianni, A., & Mentzakis, E. (2018). Social interactions in online eating disorder communities: A network perspective. *PLoS One* 13(7), 1–17.
<http://dx.doi.org/10.5258/SOTON/D0551>

- Wardecker, B. M., Edelstein, R. S., Quas, J. A., Cordo, I. M., & Goodman, G. S. (2017). Emotion language in trauma narratives is associated with better psychological adjustment among survivors of childhood sexual abuse. *Journal of Language and Social Psychology, 36*(6), 628–653. <http://dx.doi.org/10.1177/0261927X17706940>
- Weisser, M. (2016). *Practical corpus linguistics: An introduction to corpus-based language analysis*. Hoboken, NJ: John Wiley & Sons, Inc.
- Wong, C. A., Merchant, R. M., & Moreno, M. A. (2014). Using social media to engage adolescents and young adults with their health. *Healthcare, 2*(4), 220–224. <https://doi.org/10.1016/j.hjdsi.2014.10.005>

Table 1

Anorexia Item Word Count

	Pro- Ana English %	Pro- Ana English Ct.	Pro- Ana Spanish %	Pro- Ana Spanish Ct.	Pro- Ana Mixed %	Pro- Ana Mixed Ct.
Reasons for Anorexia	2.80	1314	4.02	130	20.38	108
Pro-Ana Terms	.81	380	3.7	120	2.28	12
Methods of Anorexia	0	0	.06	2	.14	1
Hidden- Hashtag Terms	.01	5	.01	0	.05	0

Table 2

English Pronoun Word Count

	Pro- Ana English %	Pro- Ana English Ct.	Blog Norms English %	Blog Norms English Ct.	Total %	Total Ct.
All Words	93.60	46921	6.40	3206	100	50127
First- Person Singular	7.42	3482	6.26	201	7.35	3682
Third- Person Singular	.41	192	1.50	48	.48	240
Third- Person Plural	.29	136	0.68	22	.32	158
First- Person Plural	.16	75	0.91	29	.21	104

Table 3

English Pronoun Test Results

Pronoun	Pro- Ana Eng. Obs. Ct.	Pro- Ana Eng. Exp. Ct.	Blog Norms Eng. Obs. Ct.	Blog Norms Eng. Exp. Ct.	Family- Wise Error Rate	Adj. Error Rate	<i>df</i>	<i>cv</i>	<i>LL</i>	<i>BIC</i>
Third- Person Singular	192	224.65	48	15.35	0.05	0.0033	1	8.62	49.13	38.31
First- Person Plural	75	97.35	29	6.65	0.05	0.0033	1	8.62	46.27	35.45
Third- Person Plural	136	147.89	22	10.11	0.05	0.0033	1	8.62	11.42	0.6
First- Person Singular	3482	3446.48	201	235.52	0.05	0.0033	1	8.62	7.69	-3.13

Table 4

English Psychological Process Category Test Results

Category	Pro- Ana Eng. Obs. Ct.	Pro- Ana Eng. Exp. Ct.	Blog Norms Eng. Obs. Ct.	Blog Norms Exp. Ct.	Family- Wise Error Rate	Adj. Error Rate	<i>df</i>	<i>cv</i>	<i>LL</i>	<i>BIC</i>
Ingestion	4664	4281.58	17	299.42	0.05	0.0033	1	8.62	485.14	474.32
Health	2876	2710.75	20	185.25	0.05	0.0033	1	8.62	251.33	240.51
Body	1140	1089.54	24	74.46	0.05	0.0033	1	8.62	48.87	38.05
Sadness	713	680.5	14	46.5	0.05	0.0033	1	8.62	32.92	22.1
Neg. E.	1760	1709.2	66	116.8	0.05	0.0033	1	8.62	27.75	16.93
Insight	774	792.82	73	54.18	0.05	0.0033	1	8.62	6.34	-4.48
Anger	488	477.38	22	32.62	0.05	0.0033	1	8.62	4.15	-6.68
Death	127	122.62	5	8.38	0.05	0.0033	1	8.62	.0528	-7.07
Anxiety	211	205.93	9	14.07	0.05	0.0033	1	8.62	2.22	-8.6
Cog. Pro.	5171	5187.5	371	354.5	0.05	0.0033	1	8.62	0.81	-10.01

Table 5

Spanish Psychological Process Category Test Results

Category	Pro- Ana Span. Obs. Ct.	Pro- Ana Span. Exp. Ct.	Blog Norms Eng. Obs. Ct.	Blog Norms Exp. Ct.	Family- Wise Error Rate	Adj. Error Rate	<i>df</i>	<i>cv</i>	<i>LL</i>	<i>BIC</i>
Ingestion	163	90.38	17	89.62	0.05	0.0036	1	8.49	135.71	126.94
Cog. Pro.	672	523.73	371	519.27	0.05	0.0036	1	8.49	86.56	76.79
Body	123	73.81	24	73.19	0.05	0.0036	1	8.49	72.1	63.33
Health	62	41.18	20	40.82	0.05	0.0036	1	8.49	22.21	13.44
Neg. E.	115	90.89	66	90.11	0.05	0.0036	1	8.49	13.02	4.25
Death	23	14.06	5	13.94	0.05	0.0036	1	8.49	12.39	3.62
Insight	109	91.39	73	90.61	0.05	0.0036	1	8.49	6.86	-1.91
Anger	40	31.13	22	30.87	0.05	0.0036	1	8.49	5.15	-3.62
Sadness	28	21.09	14	20.91	0.05	0.0036	1	8.49	4.64	-4.13
Anxiety	16	12.55	9	12.45	0.05	0.0036	1	8.49	1.93	-6.84

Table 6

Mixed Psychological Process Category Test Results

Category	Pro- Ana Mixed Obs. Ct.	Pro- Ana Mixed Exp. Ct.	Blog Norms Eng. Obs. Ct.	Blog Norms Exp. Ct.	Family- Wise Error Rate	Adj. Error Rate	<i>df</i>	<i>cv</i>	<i>LL</i>	<i>BIC</i>
Health	108	18.19	20	109.81	0.05	0.0036	1	8.49	316.68	308.46
Ingestion	101	16.76	17	101.24	0.05	0.0036	1	8.49	302.09	293.87
Sadness	75	12.64	14	76.36	0.05	0.0036	1	8.49	219.54	211.31
Neg. E.	106	24.44	66	147.56	0.05	0.0036	1	8.49	204.87	196.64
Death	55	8.52	5	51.48	0.05	0.0036	1	8.49	181.77	173.54
Anger	37	8.38	22	50.62	0.05	0.0036	1	8.49	73.21	64.98
Body	35	8.38	24	50.62	0.05	0.0036	1	8.49	64.22	56
Cog. Pro.	27	56.55	371	341.45	0.05	0.0036	1	8.49	21.66	13.43
Anxiety	9	2.56	9	15.44	0.05	0.0036	1	8.49	12.93	4.7
Insight	5	11.08	73	66.92	0.05	0.0036	1	8.49	4.74	-3.48

Chapter 4: A General Conclusion

Overview

To better inform research and clinical practice, this dissertation examined the language of public pro-nonsuicidal self-injury (NSSI) and pro-anorexia (pro-ana) posts written in Spanish and appearing on Tumblr, an online microblog. The population at the heart of this research, adolescents and young adults, are those most likely to engage in self-harming and anorexia-related behaviors; and those for whom Spanish is their dominant language are more likely to express themselves in their native tongue. This study explored the language processes of these individuals both to better understand what they have been communicating to others and to provide insight into their cognitive processes.

Tumblr is an immensely popular microblogging social media platform with worldwide appeal, claiming 345 million registered users, 738 million unique visitors during its history as a website, and over 148 billion posts as of 2017, with the majority of users under age 25 (Clement, 2019). Tumblr's users are highly engaged, meaning the site's public posts were readily available for inclusion in this study. Considering Tumblr's younger and active user base and given that the age of onset for anorexia and NSSI behaviors coincided with the age of users, Tumblr was an ideal platform to rely on for this study.

The remainder of this chapter will, first, summarize the two research manuscripts, Chapter 2, which focused on pro-NSSI language written in Spanish and appearing on Tumblr and Chapter 3, which also attended to posts written in Spanish and presented on Tumblr, but which focused, instead, on pro-ana language. These summaries will be followed by a discussion of the thematic links between the two works, with attention to the numerous contributions this dissertation makes to the literature in this research domain. Finally, this dissertation will conclude with an explication of an agenda for future work that builds from these manuscripts.

Summary of the Research Manuscripts

Manuscript 1

The first research manuscript explored the language used in public pro-NSSI Tumblr posts written in Spanish, comparing such language to the words arising within blogs overall. This study addressed four research questions. The first question involved a scrutiny of the usage of pro-NSSI-specific Spanish words. The second question opened an exploration of the linguistic processes of any English portion of the posts under review in the study. The third question offered an examination of whether the linguistic processes of these posts differed from posts occurring on blogs, overall. Finally, the fourth question focused on the psychological processes underway within the Spanish portions of these posts.

In reference to the first question, *reasons for NSSI* was the most commonly occurring category of pro-NSSI-specific words. This result was consistent with research on English-language posts, and it supports the generally accepted belief that those who engage in NSSI behaviors do so because they are experiencing intense distress. These individuals, therefore, turn to such activities as a way of alleviating or, at a minimum, modulating their pain. Even more, the results from this study showed that, when those posting on Tumblr were writing in their dominant language (Spanish), the words used in this category were prevalent as individuals retrieved traumatic memories.

Data generated in response to the second and third research questions showed that pronoun use differed from English Twitter norms most frequently and significantly in the forms of first-person singular, followed by third-person plural and first-person plural. Using the first-person singular form can indicate depression and the perception of lower social status (Pennebaker, 2011). The English sub-corpus constituted language taken from posts including

Spanish, which assumed both that the user switched languages in the post and that the individual was bilingual. When bilingual individuals use their non-dominant language to discuss emotional issues, it can be a way for them to create distance from or entirely avoid emotional experiences and vulnerability.

The results from the fourth research question revealed the psychological process word-use patterns across all three pro-NSSI sub-corpora. The most striking result was for *negative emotion*, a category that ranked first in effect size within the sub-corpora containing Spanish (Spanish and Mixed). In the case of the English-language sub-corpora, *health* was the category first in effect size, with *negative emotion* a distant second. These results suggest that expression in the non-dominant language could be more concrete and less emotionally charged (Bailey, McIntyre, Arreola, & Venta, 2019), explaining why *health* was significantly more prevalent than *negative emotion* in the second-language-only sub-corpus.

These results are consistent with prior research suggesting that people who engage in self-harm are in immense pain and would be expected to use a negative emotional tone to express their distress. Research supports a correlation between feelings of low self-esteem and the use of words in the category *negative emotion* (Pennebaker, Mehl, & Niederhoffer, 2003). People who engage in NSSI are likely to experience low self-esteem along with their emotional pain.

This study was one of the first to examine NSSI-related language written in Spanish and appearing on Tumblr (or any other social media website), successfully identifying language markers demonstrating a level of mental illness and emotional pain in public posts on a microblogging social media website. Still, while these results are robust and inviting of additional research attention, this study does admit of a few limitations. First, words can have

multiple meanings, and words in LIWC software can exist in more than one category, skewing these data to a degree. To control for this potential, before these data were included, they were analyzed by LIWC software. Second, this study focused on public Tumblr posts emanating from North, Central, and South America. Individuals posting on Tumblr could have done so through a private account, or they could have been Spanish-speaking individuals who did not reside in the countries included within these three regions. This constraint could have presented a less-than-whole picture of the relevant population. Third, this study began by relying on a Spanish-language corpus but ended up using Spanish, English, and Mixed sub-corpora because of the phenomenon of code-switching. Spanish and mixed-language blog norms were not available for comparison in this study. In sum, while the effect size for this study was large, inferences from these data should be drawn with care.

Practitioners can learn from this research. For one thing, this study suggests what to inquire about when encountering a client's NSSI behaviors or ideation; this work indicates the lines of inquiry about specific NSSI behaviors and online activities that counselors should pursue. Asking clients about NSSI and online behaviors is a necessary skill for counselors to have, especially considering these topics might not be ones that an adolescent client would be likely to raise. And nor would such issues be covered by standard assessment questions. This is especially important for bilingual clients whose dominant language is Spanish.

For another thing, this research points to the need to consider bilingualism with clients with NSSI issues. These clients often feel isolated, and they could be reaching out to online communities where they are not shamed and where they might feel better understood. The language of these clients' distress is evident in the frequency of their negative emotional tone and sadness. Distress of this kind can be overwhelming to counselors; and, to this point, there

has been no evidence-based treatment of NSSI (Washburn et al., 2012). Even more, considering the dearth of Spanish-speaking or bilingual counselors, narration in trauma-focused therapies could be less effective, leaving practitioners feeling stuck in the treatment of their clients.

In terms of the research implications of this study, these results demonstrate the need for a more robust and refined understanding of the relationships and interactions between language and NSSI behaviors. It is vitally important that scholars initiate and sustain lines of research that examine NSSI behavior and ideation in Spanish and Spanish-English code-switching client narratives. So, too, does the present study underscore the fact that researchers must develop Spanish blog norms as norms for code-switching corpora. Existing options are demonstrably insufficient.

Manuscript 2

The second research manuscript explored the language of public pro-ana Tumblr posts written in Spanish, comparing the use of language on pro-ana blogs to blogs overall. Four research questions were addressed in this study. The first question focused on the use of specific pro-ana words; the second and third questions explored the linguistic processes of the posts, comparing them to posts in blogs overall; and, the fourth question looked at the use of psychological processes in the posts.

The findings from the first research question revealed that the most common category of word usage was *reasons for anorexia*. The words in this category (e.g., hate, control, and depression) were associated with negative emotions, which are more prominent in the dominant language (Spanish) during the retrieval of traumatic memories and which could offer potential explanations for or justifications of such behaviors as expressed in these posts. Specifically,

anorexia behaviors are a way for individuals to manage their feelings of anger and to avoid other social forms of expression of anger (Espeset, Gulliksen, Nordbo, Skarderud, & Holte, 2012).

The results from the second and third research questions revealed the pronoun use rate differed from English Twitter norms most frequently and significantly in third-person singular, followed by first-person plural and then third-person plural. In a pro-ana community, we can expect language to reflect a greater focus on people and objects that have been linked to diminished prospects for recovery. Combined with the overall negative emotional tone of these posts, the focus on others and objects signals feelings of low self-esteem and diminished inclination to change (Pennebaker et al., 2003).

The results from the fourth research question revealed that the most common and statistically significant psychological process category in the English and Spanish sub-corpora was *ingestion*. The most common and statistically significant psychological process category in the Mixed sub-corpus was *health* followed by *ingestion*. These results are consistent with previous research on English dominant pro-ana groups online, and the data are similar in their focus on body, health, and ingestion. These results are also compatible with prior research findings that pro-ana communities exacerbate and reinforce harmful ideals about weight and weight-loss practices.

While this study successfully identified language markers that demonstrated a level of mental illness and emotional pain in public posts on a microblogging social media website, there are some noteworthy limitations to the results. First, when using LIWC software, words can arise in more than one category and can have multiple meanings. This could somewhat skew the results. To reduce confusion over the meaning of words, the data were reviewed before being analyzed by LIWC software. The second limitation to the study is that private posts, and posts

made outside South, North, and Central America were not included. Because of banning hashtags and content blocking on social media sites, it is possible that users have private accounts where they screen followers to ensure that their content won't be reported by other users. This limits the data available to collect and could have affected the results. The third limitation is that this study attempted to use a Spanish-language corpus, initially, but it ended up including Spanish, English, and Mixed sub-corpora due to code-switching. There were no Spanish-language or mixed-language blog norms available for comparison, thus limiting the external validity of the study. Given these limitations, inferences into data should be undertaken with careful consideration.

Two implications for practitioners and one implication for researchers emerged from this study. First, this research suggests areas of additional inquiry when encountering a Spanish-speaking or bilingual client engaging in anorexia behaviors. Asking clients about such behaviors is a necessary skill for counselors to have because it is unlikely adolescent clients would bring up the topic on their own. Gathering more information about online behaviors, trauma history, and engagement with emotions affords additional opportunities for assessment with these clients.

The second implication would be for counselors to include information about linguistic markers in pro-ana and pro-recovery posts in their implementation of the Stages of Change (or Transtheoretical) Model (Prochaska & DiClemente, 1983). As research continues to develop linguistic markers among groups, the goal remains for counselors to help facilitate health and wellness. This study demonstrates that social media usage can often be a maladaptive way for these individuals to cope. Individuals are using their writing to reach out to people in a community of others with similar challenges as a means of managing their emotions, but they are not doing so effectively. As Pennebaker (2013) found, writing with words of a negative

emotional nature and a negative tone does not help to relieve distressing feelings. By monitoring clients' language use to indicate stages of change and by more accurately targeting interventions for progression, counselors can provide more tailored and individualized treatment to increase recovery rates and lengths of wellness. Finally, an important implication for this research is the continued call for a deeper understanding of the relationship between language and anorexia behaviors, particularly in Spanish-speaking and bilingual individuals. For research to continue to expand and become more inclusive, Spanish blog norms and norms for code-switching corpora must be developed.

Thematic Links Between Manuscripts and Contributions to the Literature

The studies in this dissertation have much in common. The first manuscript explored the language, linguistic processes, and psychological processes of public pro-NSSI Tumblr posts written in Spanish, while the second manuscript looked at the same phenomena but with attention to pro-ana rather than pro-NSSI posts. While these communities are discrete populations, with their own distinctive qualities and concerns, those displaying pro-NSSI and pro-ana behaviors also overlap in significant ways (Cipriano, Cella, & Cotrufo, 2017). For example, both studies revealed similar findings for the first research question, where the *reasons* item (i.e., reasons for the behavior) was the most common word category appearing in members' posts. This indicates that those engaging in NSSI and anorexia behaviors are seeking engagement and trying to cope with their considerable emotional distress.

So, too, did each study reveal a negative emotional tone, demonstrating that both populations struggle with mental illness and maladaptive behaviors. While attempting to cope with their feelings, members reach out to others about their behaviors, even as their efforts are often, ultimately, ineffective—especially considering that repeated expression cast in a negative

emotional tone encourages a negative emotional climate as opposed to relieving distress (Pennebaker, 2013). At the same time, research has also shown that social sharing can improve interpersonal relationships and that social support can reduce stress (Rime, Paez, Basabe, & Martinez, 2010). Given the paradoxical nature of these phenomena and to better evaluate their clients' readiness for change, counselors treating clients with either NSSI or anorexia behaviors must carefully consider the nature of the words being expressed, particularly those that might be of a negative emotional quality.

Overall, as NSSI behaviors have proliferated, and following the identification of NSSI as an area for future research in the fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5), these concerns have garnered greater attention from the public. Since the expansion of the category of eating disorders in the DSM-5, there has also been more research on anorexia behaviors and other eating disorders; however, there has still been too little study of the language use and language processes employed by those engaging in NSSI and anorexia behaviors and later posting about their activities online. This is even more the case for those depicting their activities in Spanish.

This relative lack of research attention is unfortunate considering the sizable percentage of the global population that claims Spanish as its dominant language, as well as the prevalence of both NSSI and anorexia behaviors among adolescents and young adults around the world. To meet the needs of these communities, thus far underrepresented in and underserved by scholarship in the field, this dissertation focused on understanding language and psychological markers that might shed insight into the mental health, behaviors, and cognitive processes of these at-risk populations. As this work has shown, these populations increasingly rely on the

interconnectedness of the Internet, even as their deployment of such resources can exacerbate underlying problems.

One of the challenges confronted by researchers working in this area, which to a degree explains the comparative shortage of attention to Spanish-speakers demonstrating disordered behavior, is the inconsistent use of language, in English and Spanish, regarding behaviors and symptoms associated with NSSI and anorexia. Even key words and abbreviations used in the research can vary (e.g., compare *NSSI* to *ANS*), making it even more difficult to successfully search databases. By focusing on the Spanish language, coordinating bilingual titles and key words in manuscripts for publication, and reinforcing the call for Spanish and code-switching blog norms, this dissertation addressed a significant gap in the literature and provided a foundation for future study.

Finally, in situations where a trauma history will likely lead to presenting issues, the manuscripts in this dissertation emphasized the importance of discretion and prudence when selecting treatment modalities for bilingual or dominantly Spanish-speaking individuals. Recall that, in the posts of both pro-NSSI and pro-ana populations, the most frequent word category associated with their behaviors was *reasons*—involving intense emotion words offered in the dominant language of Spanish. This suggests that emotion words are more easily accessed and used in the dominant language, a finding consistent with previous literature (Pavlenko, 2008).

Counselors might not realize that interventions around trauma could be less effective when conducted in a non-dominant language; this concern might be overlooked, especially when the proximate goals for this population could be oriented more toward immediate intervention—stopping self-harming and anorexic behaviors—rather than processing the trauma narrative. Following the present studies then, counselors should assess for NSSI and anorexia behaviors

while also scrutinizing the online behaviors of their clients. It is imperative that counselors are aware of these findings and that they incorporate them into their regular assessment and practice procedures. The research in this dissertation provides guidance and specific markers to be aware of when doing so.

Future Research Agenda

This research opens many avenues for future investigation. The study of language and its application to mental health, specifically as it relates to online behaviors, is still a new arena of inquiry in the field of counseling. A future study that could be conducted with the same data generated by and analyzed in this dissertation might focus, instead, on gendered language, allowing the researcher to determine which gender is posting more about their behaviors online, while also testing for significant differences between groups. Such a study could be conducted by novice researchers because it could be carried out with minimal resources. Because there is currently so little information available using Spanish corpora, the possibilities in this direction are vast and inviting.

Another study flowing from the present research could be an intervention-based assessment of clients' writing habits and tendencies, an inquiry looking for either indications of a positive emotional tone or markers for recovery following specific interventions. Such a study has not yet been conducted with Spanish-speaking individuals involved in NSSI or anorexia behaviors. This sort of work is, however, vital: As this dissertation has shown, it is imperative that researchers better understand this growing segment of the population.

I would like to expand my research by using additional corpus linguistic techniques to explore synchronous and asynchronous interactions between people in online communities and in sessions with therapists. This would involve examining transcripts of therapy sessions to test

for a match in linguistic style between bilingual therapists and clients, affording the chance to evaluate the working alliance and to, more specifically, explore the treatment of trauma. I would also like to continue to use data from online sources such as social media platforms to look for what are known as *collocations*, or words frequently appearing with a specific word of interest. There is overlap between pro-NSSI and pro-ana populations, but the degree of this coexistence is unclear. To that end, I would like to explore other co-occurring disorders to better understand how language is associated with specific target words across populations.

References

- Bailey, C., McIntyre, E., Arreola, A., & Venta, A. (2019). What are we missing? How language impacts trauma narratives. *Journal of Child & Adolescent Trauma*, 1–9.
<https://doi.org/10.1007/s40653-019-00263-3>
- Cipriano, A., Cella, S., & Cotrufo, P. (2017). Nonsuicidal self-injury: A systematic review. *Frontiers in Psychology*, 8(1946), 1–14.
- Clement, J. (2019). Cumulative total of Tumblr blogs from May 2011 to July 2019 (in millions) [Blog post]. Retrieved from <https://www.statista.com/statistics/256235/total-cumulative-number-of-tumblr-blogs/>
- Espeset, E. M. S., Gulliksen, K. S., Nordbo, R. H. S., Skarderud, F., & Holte, A. (2012). The link between negative emotions and eating disorder behavior in patients with anorexia nervosa. *European Eating Disorders Review*, 20(6), 451–460.
<https://doi.org/10.1002/erv.2183>
- Pavlenko, A. (2008). Emotion and emotion-laden words in the bilingual lexicon. *Bilingualism: Language and cognition*, 11(2), 147–164.
<https://psycnet.apa.org/doi/10.1017/S1366728908003283>
- Pennebaker, J. W. (2013). *The secret life of pronouns: What our words say about us*. New York, NY: Bloomsbury Press.
- Pennebaker, J. W., Mehl, M. R., & Niederhoffer, K. G. (2003). Psychological aspects of natural language use: Our words, our selves. *Annual Review of Psychology*, 31, 547.
- Prochaska, J. O., & DiClemente, C. C. (1983). Stages and processes of self-change of smoking: Toward an integrative model of change. *Journal of Consulting and Clinical Psychology*, 51(3), 390–395.

Rime, B., Paez, D., Besaba, N., & Martinez, F. (2010). Social sharing of emotion, post-traumatic growth and emotional climate: Follow-up of Spanish citizen's response to the collective trauma of March 11th terrorist attacks in Madrid. *European Journal of Social Psychology*, *40*, 1029–1045. <https://doi.org/10.1002/ejsp.700>

Washburn, J. J., Richardt, S. L., Styer, D. M., Gebhardt, M., Juzwin, K. R., Yourek, A., & Aldridge, D. (2012). Psychotherapeutic approaches to non-suicidal self-injury in adolescents. *Child and Adolescent Psychiatry and Mental Health*, *6*(14). <https://doi.org/10.1186/1753-2000-6-14>

Bibliography

- Aalai, A. (2018, June 8). The social contagion of suicide [Blog post]. Retrieved from <https://www.psychologytoday.com/us/blog/the-first-impression/201806/the-social-contagion-suicide>
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.
- Arcelus, J., Mitchell, A. J., Wales, J., & Neilson, S. (2011). Mortality rates in patients with anorexia nervosa and other eating disorders: A meta-analysis of 36 studies. *Archives of General Psychiatry*, *68*(7), 724–731. <https://doi.org/10.1001/archgenpsychiatry.2011.74>
- Arseniev-Koehler, A., Lee, H., McCormick, T., & Moreno, M. A. (2016). #Proana: Pro-eating disorder socialization on Twitter. *Journal of Adolescent Health*, *58*, 659–664. <https://doi.org/10.1016/j.jadohealth.2016.02.012>
- Ashton, M. C., Lee, K., Perugini, M., Szarota, P., de Vries, R. E., Di Blas, L., . . . De Raad, B. (2004). A six-factor structure of personality-descriptive adjectives: Solutions from psycholexical studies in seven languages. *Journal of Personality and Social Psychology*, *86*, 356–366. <http://dx.doi.org/10.1037/0022-3514.86.2.356>
- Bailey, C., McIntyre, E., Arreola, A., & Venta, A. (2019). What are we missing? How language impacts trauma narratives. *Journal of Child & Adolescent Trauma*, 1–9. <https://doi.org/10.1007/s40653-019-00263-3>
- Baron, N. (2008). *Always on: Language in an online and mobile world*. New York, NY: Oxford University Press.
- Benjet, C., Gonzalez-Herrera, I., Castro-Silva, E., Mendez, E., Borges, G., Casanova, L., & Medina-Mora, M. E. (2017). Non-suicidal self-injury in Mexican young adults:

- Prevalence, associations with suicidal behavior and psychiatric disorders, and DSM-5 proposed diagnostic criteria. *Journal of Affective Disorders*, 215, 1–8.
<https://doi.org/10.1016/j.jad.2017.03.025>
- Bentley, K. H., Nock, M. K., & Barlow, D. H. (2014). The four-function model of nonsuicidal self-injury: Key directions for future research. *Clinical Psychological Science*, 2(5), 638–656. <https://psycnet.apa.org/doi/10.1177/2167702613514563>
- Biber, D. E., Reppen, R., & Fringinal, E. (2012). Research in corpus linguistics. In *The Oxford Handbook of Applied Linguistics*, (2nd ed.) Oxford University Press.
<https://doi.org/10.1093/osfordhb/9780195384253.013.0038>
- Bjekić, J., Lazarević, L. B., Živanović, M., & Knežević, G. (2014). Psychometric evaluation of the Serbian dictionary for automatic text analysis-LIWCser. *Psihologija*, 47, 5–32.
<https://doi.org/10.2298/PSI1401005B>
- Brezina, V. (2018). *Statistics in corpus linguistics: A practical guide*. Cambridge, UK: Cambridge University Press.
- Brown, R. C., Fischer, T., Goldwich, A. D., Keller, F., Young, R., & Plener, P. L. (2018). #cutting: Non-suicidal self-injury (NSSI) on Instagram. *Psychological Medicine*, 48(2), 337–346. <https://doi.org/10.1017/S0033291717001751>
- Campbell, K., & Peebles, R. (2014). Eating disorders in children and adolescents: State of the art review. *Pediatrics*, 134(3), 582–592. <https://doi.org/10.1542/peds.2014-0194>
- Castro, T. S., & Osorio, A. J. (2013). “I love my bones!”¹ - self-harm and dangerous eating youth behaviors in Portuguese written blogs. *Young Consumers*, 14(4), 321–330.
<https://doi.org/10.1108/YC-03-2013-00351>

- Cervantes, R. C., Goldback, J. T., Varela, A., & Santisteban, D. A. (2014). Self-harm among Hispanic adolescents: Investigating the role of culture-related stressors. *Journal of Adolescent Health, 55*, 633–639. <https://doi.org/10.1016/j.jadohealth.2014.05.017>
- Chancellor, S., Kalantidis, Y., Pater, J., De Choudhury, M., & Shamma, D. A. (2017). *Multimodal classification of moderated pro-eating disorder content*. Paper presented at the 2017 CHI Conference. <http://dx.doi.org/10.1145/3025453.3025985>
- Chancellor, S., Lin, Z., Goodman, E. L., Zerwas, S., & Choudhury, M. D. (2016). Quantifying and predicting mental illness severity in online pro-eating disorder communities. *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, 1171–1184. <https://dx.doi.org/10.1145/2818048.2819973>
- Chancellor, S., Mitra, T., & De Choudhury, M. (2016). Recovery amid pro-anorexia: Analysis of recovery in social media. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 2111–2123. <https://doi.org/10.1145/2858036.2858246>
- Chancellor, S., Pater, J., Clear, T., Gilbert, E., & De Choudhury, M. (2016). #thyghgapp: Instagram content moderation and lexical variation in pro-eating disorder communities. *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, 1201–1213. <https://doi.org/10.1145/2818048.2819963>
- Choudhury, M. D., Counts, S., & Horvitz, E. (2013). Major life changes and behavioral markers in social media. *Proceedings of the 2013 Conference on Computer-Supported Cooperative Work*, 1431–1442. <https://dx.doi.org/10.1145/2441776.2441937>
- Cipriano, A., Cella, S., & Cotrufo, P. (2017). Nonsuicidal self-injury: A systematic review. *Frontiers in Psychology, 8*(1946), 1–14. <https://doi.org/10.3389/fpsyg.2017.01946>

- Croyle, K. L., (2007). Self-harm experiences among Hispanic and Non-Hispanic White young adults. *Hispanic Journal of Behavioral Sciences*, 29(2), 242–253.
- De Muth, J. E. (2014). *Basic statistics and pharmaceutical statistical applications*. Danvers, MA: Chapman and Hall/CRC.
- Department of Psychological Services. (2015). *Reformatting transcripts for LIWC 2015*. Lubbock, TX: Author. Retrieved from:
<http://www.depts.ttu.edu/psy/lusi/files/liwc2015transcriptioncleaning.pdf>
- Espeset, E. M. S., Gulliksen, K. S., Nordbo, R. H. S., Skarderud, F., & Holte, A. (2012). The link between negative emotions and eating disorder behavior in patients with anorexia nervosa. *European Eating Disorders Review*, 20(6), 451–460.
<https://doi.org/10.1002/erv.2183>
- Explainer: What is Tumblr? (2018, August 08). Retrieved from
<https://www.webwise.ie/parents/explainer-what-is-tumblr-2/>
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, 41, 1149–1160. <http://dx.doi.org/10.3758/BRM.41.4.1149>
- Favazza, A. R. (1998). The coming of age of self-mutilation. *The Journal of Nervous & Mental Disease*, 186, 259–268. <https://dx.doi.org/10.1097/00005053-199805000-00001>
- Feder, P. I. (1975). The log likelihood ratio in segmented regression. *The Annals of Statistics*, 3(1), 84–97. <https://www.jstor.org/stable/2958080>
- Fuertes, J. N. (2004). Supervision in bilingual counseling: Service delivery, training, and research considerations. *Journal of Multicultural Counseling and Development*, 32, 84–94. <https://doi.org/10.1002/j.2161-1912.2004.tb00363.x>

- Gaur, M., Kursuncu, U., Alambo, A., Sheth, A., Daniulaityte, R., Thirunarayan, K., & Pathak, J. (2018). "Let me tell you about your mental health!" Contextualized classification of Reddit posts to DSM-5 for web-based intervention. *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*, 753–762. <https://doi.org/10.1145/3269206.3271732>
- Gonzalez-Barrera, A., & Lopez, M. H. (2013). Spanish is the most spoken non-English language in U.S. homes, even among non-Hispanics. Retrieved from <http://pewrsr.ch/11S8uek>
- Greaves, M. M., & Dykeman, C. (2019, February 3). A corpus linguistic analysis of public Tumblr blog posts on non-suicidal self-injury. <https://doi.org/10.31234/osf.io/k4qt3>
- Gulbas, L. E., Hausmann-Stabile, C., De Luca, S. M., Tyler, T. R., & Zayas, L. H. (2015). An exploratory study of non-suicidal self-injury and suicidal behaviors in adolescent Latinas. *American Journal of Orthopsychiatry*, 84(4), 302–314. <http://dx.doi.org/10.1037/ort0000073>
- Guttfreund, D. G. (1990). Effects of language usage on the emotional experience of Spanish-English and English-Spanish bilinguals. *Journal of Consulting and Clinical Psychology*, 58(5), 604–607. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/2254507>
- Harned, M. S., Najavits, L. M., & Weiss, R. D. (2006). Self-harm and suicidal behavior in women with comorbid PTSD and substance dependence. *The American Journal on Addictions*, 15, 392–395. <https://doi.org/10.1080/10550490600860387>
- Hashtag. (n.d.). Retrieved from <https://www.merriam-webster.com/dictionary/hashtag>
- Hull, K., Kim, J. K., & Stillwell, M. (2018): Fotos de béisbol: An examination of the Spanish-language Instagram accounts of Major League Baseball teams. *Howard Journal of Communications*, 30, 249–264. <https://doi.org/10.1080/10646175.2018.1471756>

- Ilakkuvan, V., Johnson, A., Villanti, A. C., Evans, W. D., & Turner, M. (2019). Patterns of social media use and their relationship to health risks among young adults. *Journal of Adolescent Health, 64*(2), 158–164. <https://doi.org/10.1016/j.jadohealth.2018.06.025>
- Jarvi, S., Jackson, B., Swenson, L., & Crawford, H. (2013). The impact of social contagion on non-suicidal self-injury: A review of the literature. *Archives of Suicide Research, 17*, 1–19. <https://doi.org/10.1080/13811118.2013.748404>
- Joiner, T. E., Jr., Van Orden, K. A., Witte, T. K., & Rudd, M. D. (2009). *The interpersonal theory of suicide: Guidance for working with suicidal clients*. <http://dx.doi.org/10.1037/11869-000>
- Karp, M., & Vogele, C. Does anyone still understand me? Psychotherapy and multilingualism. *Verhaltenstherapie, 26*(3), 156–157. <https://doi.org/10.1159/000448818>
- Kerr, P., Muehlenkamp, J., & Turner, J. (2010). Nonsuicidal self-injury: A review of current research for family medicine and primary care physicians. *Journal of the American Board of Family Medicine 23*(2), 240–259. <https://doi.org/10.3122/jabfm.2010.02.090110>
- Klonsky, D. E., Victor, S. E., & Saffer, B. Y. (2014). Nonsuicidal self-injury: What we know, and what we need to know. *Canadian Journal of Psychiatry, 59*(11), 565–568. <https://dx.doi.org/10.1177%2F070674371405901101>
- Kolar, D. R., Mejia Rodriguez, D. L., Mebarak Chams, M., & Hoek, H. W. (2016). Epidemiology of eating disorders in Latin America: A systematic review and meta-analysis. *Current Opinion in Psychiatry, 29*(6), 363–371. <http://dx.doi.org/10.1097/YCO.0000000000000279>

- Koutek, J., Docourkova, J., & Dudovva, I. (2016). Suicidal behavior and self-harm in girls with eating disorders. *Neuropsychiatric Disease and Treatment, 12*, 787–793.
<http://dx.doi.org/10.2147/NDT.S103015>
- Lewis, S. P., Heath, N. L., Michal, N. J., & Duggan, J. M. (2012). Non-suicidal self-injury, youth, and the internet: What mental health professionals need to know. *Child and Adolescent Psychiatry and Mental Health, 6*(13), 1–9. <https://doi.org/10.1186/1753-2000-6-13>
- Li, F., & Du, T. C. (2014). Listen to me — Evaluating the influence of micro-blogs. *Decision Support Systems, 62*, 119–130. <https://dx.doi.org/10.1016/j.dss.2014.03.008>
- Matsakis, L. (2018). How pro-eating disorder posts evade filters on social media. Retrieved from <https://www.wired.com/story/how-pro-eating-disorder-posts-evade-social-media-filters/>
- McEnery, T., & Hardie, A. (2012). *Corpus linguistics: Method, theory and practice*. Cambridge, UK: Cambridge University Press.
- McNamara, P., McLaren, D., & Durso, K. (2007). Representation of the self in REM and NREM dreams. *Dreaming, 17*(2), 113–126. <http://dx.doi.org/10.1037/1053-0797.17.2.113>
- Merikangas, K. R., He, J., Burstein, M., Swanson, S. A., Avenevoli, S., Cui, L., . . . Swendsen, J. (2010). Lifetime prevalence of mental disorders in U.S. adolescents: Results from the National Comorbidity Study-Adolescent Supplement (NCS-A). *Journal of the American Academy of Child and Adolescent Psychiatry, 49*(10), 980–989. <http://dx.doi.org/10.1016/j.jaac.2010.05.017>
- Moessner, M. M., Feldhege, J., Wolf, M., & Bauer, S. (2018). Analyzing big data in social media: Text and network analyses of an eating disorder forum. *International Journal of Eating Disorders, 51*, 656–667. <https://doi.org/10.1002/eat.22878>

- Moreno, M. A., Ton, A., Selkie, E., & Evans, Y. (2016). Secret society 123: Understanding the language of self-harm on Instagram. *Journal of Adolescent Health, 58*(1), 78–84.
<http://dx.doi.org/10.1016/j.jadohealth.2015.09.015>
- Nock, M. K., & Favazza, A. R. (2009). Non-suicidal self-injury: Definition and classification. In M. K. Nock (Ed.), *Understanding non-suicidal self-injury: Origins, assessment, and treatment* (pp. 9–18). Washington DC: American Psychological Association.
- Nock, M. K., & Prinstein, M. J. (2005). Contextual features and behavioral functions of self-mutilation among adolescents. *Journal of Abnormal Psychology, 114*(1), 140–146.
<https://psycnet.apa.org/doi/10.1037/0021-843X.114.1.140>
- O’Dea, B., Larsen, M. E., Batterham, P. J., Calex, A. L., & Christensen, H. (2017). A linguistic analysis of suicide-related Twitter posts. *Crisis: The Journal of Crisis Intervention and Suicide Prevention, 38*, 319–329. <https://doi.org/10.1027/0227-5910/a000443>
- Olatunji, B. O., Cox, R., Ebesutani, C., & Wall, D. (2015). Self-harm history predicts resistance to inpatient treatment of body shape aversion in women with eating disorders: The role of negative affect. *Journal of Psychiatric Research, 65*, 37–46.
<https://doi.org/10.1016/j.jpsychires.2015.03.016>
- O’Reilly, M., Dogra, N., Whiteman, N., Hughes, J., Eruyar, S., & Reilly, P. (2018). Is social media bad for mental health and wellbeing? Exploring the perspectives of adolescents. *Clinical Child Psychology and Psychiatry, 23*, 601–613.
<https://dx.doi.org/10.1177/1359104518775154>
- Pavlenko, A. (2008). Emotion and emotion-laden words in the bilingual lexicon. *Bilingualism: Language and cognition, 11*(2), 147–164.
<https://psycnet.apa.org/doi/10.1017/S1366728908003283>

- Pennebaker, J. W. (2011). *The secret life of pronouns: What our words say about us*. New York, NY: Bloomsbury Press.
- Pennebaker, J. W., Booth, R. J., Boyd, R. L., & Francis, M. E. (2015). *Linguistic Inquiry and Word Count: LIWC2015 (Operator's Manual)*. Austin, TX: Pennebaker Conglomerates, Inc.
- Pennebaker, J. W., Boyd, R. L., Jordan, K., & Blackburn, K. (2015). *The development and psychometric properties of LIWC2015*. Austin, TX: Pennebaker Conglomerates, Inc.
- Perez, M., Ohrt, T. K., & Hoek, H. W. (2016). Prevalence and treatment of eating disorders among Hispanics/Latino Americans in the United States. *Current Opinion in Psychiatry*, 29(6), 378–382. <https://doi.org/10.1097/YCO.0000000000000277>
- Perry, S. (2013). The bilingual brain. Retrieved from <http://www.brainfacts.org/Archives/2008/The-Bilingual-Brain>
- Pew Hispanic Center (2018). Hispanic trends. Retrieved from <http://www.pewhispanic.org>
- Pew Research Center (2015). Demographics of Tumblr. Retrieved from http://www.pewinternet.org/2015/08/19/mobile-messaging-and-social-media-2015/2015-08-19_social-media-update_05
- Pew Research Center (2018). Social media use in 2018. Retrieved from <http://www.pewinternet.org/2018/03/01/social-media-use-in-2018>
- Ramos-Sanchez, L. (2007). Language switching and Mexican Americans' emotional expression. *Journal of Multicultural Counseling and Development*, 35, 154–168. <https://psycnet.apa.org/doi/10.1002/j.2161-1912.2007.tb00057.x>

- Rayson, P., & Garside, R. (2000). Comparing corpora using frequency profiling. *Proceedings of the Workshop on Comparing Corpora*, 9, 1–6.
<http://dx.doi.org/10.3115/1117729.1117730>
- Rime, B., Paez, D., Besaba, N., & Martinez, F. (2010). Social sharing of emotion, post-traumatic growth and emotional climate: Follow-up of Spanish citizen's response to the collective trauma of March 11th terrorist attacks in Madrid. *European Journal of Social Psychology*, 40, 1029–1045. <https://doi.org/10.1002/ejsp.700>
- Rodrigues, M. (2017). Do Hispanic girls develop eating disorders? A critical review of the literature. *Hispanic Health Care International*, 15(4), 189–196.
<https://doi.org/10.1177%2F1540415317744500>
- Rosnow, R. L., & Rosenthal, R. (2003). Effect sizes for experimenting psychologists. *Canadian Journal of Experimental Psychology*, 57, 221–237. <http://dx.doi.org/10.1037/h0087427>
- Santiago-Rivera, A. L., & Altarriba, J. (2002). The role of language in therapy with the Spanish-English bilingual client. *Professional Psychology: Research and Practice*, 33(1), 30–38.
<http://dx.doi.org/10.1037/0735-7028.33.1.30>
- Schott, N. D. (2017). Race, online space and the feminine: Unmapping “black girl thinspiration.” *Critical Sociology*, 43(7-8), 1029–1043. <http://dx.doi.org/10.1177/0896920516652456>
- Self-injurious behavior, self-injury treatment. (2010, February 22). Retrieved from
<https://www.healthyplace.com/abuse/articles/self-injurious-behavior-self-injury-treatment>
- Swanson, S. A., Crow, S. J., Le Grange, D., Swendsen, J., & Merikangas, K. R. (2011). Prevalence and correlates of eating disorders in adolescents: Results from the National Comorbidity Survey Replication Adolescent Supplement. *Archives of General Psychiatry*, 68(7), 714–23. <http://dx.doi.org/10.1001/archgenpsychiatry.2011.22>

- Syed-Abdul, S., Fernandez-Luque, L., Jian, W., Li, Y., Crain, S., Hsu, M., . . . Liou, D. (2013). Misleading health-related information promoted through video based social media: Anorexia on YouTube. *Journal of Medical Internet Research*, *15*(2), e30. <https://doi.org/10.2196/jmir.2237>
- Tausczik, Y. R., & Pennebaker, J. W. (2009). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology*, *29*, 24–54. <https://dx.doi.org/10.1177/0261927x09351676>
- Textese. (n.d.). Retrieved from <https://www.dictionary.com/browse/textese>
- Thinspiration (n.d). Retrieved from www.dictionary.com
- Thyssen, L. S., & van Camp, I. (2014). Non-suicidal self-injury in Latin America. *Salud Mental*, *4*(37), 153–157. <https://psycnet.apa.org/doi/10.17711/SM.0185-3325.2014.019>
- Vega, D., Sintes, A., Fernadez, M., Punti, J., Soler, J., Santamarina, P., . . . Pascual, J. C. (2018). Revision y actualizacion de la autolesion no suicida: Quien, como y por que? *Actas Espanolas de Psiquiatria*, *46*(4), 146–155.
- Velazquez, I. (2017). Reported literacy, media consumption and social media use as measures of relevance of Spanish as a heritage language. *International Journal of Bilingualism*, *21*(1), 21–33. <https://doi.org/10.1177%2F1367006915596377>
- Vilares, D., Alonso, M. A., & Gomez-Rodriguez, C. (2016). *EN-ES-CS: An English-Spanish code-switching Twitter corpus for multilingual sentiment analysis*. Paper presented at the Tenth International Conference on Language Resources and Evaluation, Paris, France. Abstract retrieved from https://www.researchgate.net/publication/303666796_EN-ES-CS_An_English-Spanish_Code-Switching_Twitter_Corpus_for_Multilingual_Sentiment_Analysis

- Wallace, K. (2015, November 3). Teens spend a ‘mind boggling’ 9 hours a day using media. Retrieved from <https://www.cnn.com/2015/11/03/health/teens-tweens-media-screen-use-report/index.html>
- Wang, S., Eccleston, C., & Keogh, E. (2015). The role of spatial frequency information in the recognition of facial expressions of pain. *Pain, 156*, 1670–1682. <https://doi.org/10.1097/j.pain.0000000000000226>
- Wang, T., Markus, B., Ianni, A., & Mentzakis, E. (2018). Social interactions in online eating disorder communities: A network perspective. *PLoS One 13*(7), 1–17. <http://dx.doi.org/10.5258/SOTON/D0551>
- Wardecker, B. M, Edelstein, R. S., Quas, J. A., Cordo, I. M, & Goodman, G. S. (2017). Emotion language in trauma narratives is associated with better psychological adjustment among survivors of childhood sexual abuse. *Journal of Language and Social Psychology, 36*(6), 628–653. <http://dx.doi.org/10.1177/0261927X17706940>
- Warren, C. S., Schoen, A., & Schafer, K. J. (2010). Media internalization and social comparison as predictors of eating pathology among Latino adolescents: The moderating effect of gender and generational status. *Sex Roles, 63*, 712–724. <https://doi.org/10.1007/s11199-010-9876-1>
- Washburn, J. J., Richardt, S. L., Styer, D. M., Gebhardt, M., Juzwin, K. R., Yourek, A., & Aldridge, D. (2012). Psychotherapeutic approaches to non-suicidal self-injury in adolescents. *Child and Adolescent Psychiatry and Mental Health, 6*(14). <https://doi.org/10.1186/1753-2000-6-14>
- Weisser, M. (2016). *Practical corpus linguistics: An introduction to corpus-based language analysis*. Hoboken, NJ: John Wiley & Sons, Inc.

Wong, C. A., Merchant, R. M., & Moreno, M. A. (2014). Using social media to engage adolescents and young adults with their health. *Healthcare*, 2(4), 220–224.

<https://doi.org/10.1016/j.hjdsi.2014.10.005>

Zeng, M. L. (2005). Semantic relationships used in controlled vocabularies. Retrieved from

<http://marciazeng.slis.kent.edu/Z3919/42equivalency.htm>

Appendix A: IRB Document 1



Oregon State University Research Office

Human Research Protection Program
& Institutional Review Board
B308 Kerr Administration Bldg, Corvallis OR 97331
(541) 737-8008
IRB@oregonstate.edu
<http://research.oregonstate.edu/irb>

Date of Notification	07/13/2018	Study Number	8720
Notification Type	Oversight Determination		
Principal Investigator	Cass Dykeman		
Study Team Members	Kristeen Elrod		
Study Title	Un estudio de lingüística de corpus de publicaciones públicas de Instagram y Tumblr a las pró-anorexia (pró-ana) en español. [A corpus linguistic analysis of public Instagram and Tumblr posts on pro anorexia (pro-ana) behaviors.]		
Funding Source	None	Cayuse Number	N/A

DETERMINATION: RESEARCH, BUT NO HUMAN SUBJECTS

It has been determined that your project, as submitted, does meet the definition of research but **does not** involve human subjects under the regulations set forth by the Department of Health and Human Services 45 CFR 46.

Additional review is not required for this study.

Please do not include HRPP contact information on any of your study materials.

Note that amendments to this project may impact this determination. Please submit a new request if there are changes (e.g., funding, data sources, access to individual identifiers, interaction with research subjects, etc.).

The federal definitions and guidance used to make this determination may be found at the following link: [Human Subject](#)

Appendix B: IRB Document 2



Oregon State University
Research Office

Human Research Protection Program
 & Institutional Review Board
 B308 Kerr Administration Bldg, Corvallis OR 97331
 (541) 737-8008
IRB@oregonstate.edu
<http://research.oregonstate.edu/irb>

Date of Notification	07/13/2018	Study Number	8721
Notification Type	Oversight Determination		
Principal Investigator	Cass Dykeman		
Study Team Members	Kristeen Elrod		
Study Title	Un estudio de lingüística de corpus de publicaciones públicas de Instagram y Tumblr a las pró Autolesiones No Suicidas (ANS) en español. [A corpus linguistic analysis of public Instagram and Tumblr posts on pro non-suicidal self-injury (NSSI) behaviors.]		
Funding Source	None	Cayuse Number	N/A

DETERMINATION: RESEARCH, BUT NO HUMAN SUBJECTS

It has been determined that your project, as submitted, does meet the definition of research but **does not** involve human subjects under the regulations set forth by the Department of Health and Human Services 45 CFR 46.

Additional review is not required for this study.

Please do not include HRPP contact information on any of your study materials.

Note that amendments to this project may impact this determination. Please submit a *new* request if there are changes (e.g., funding, data sources, access to individual identifiers, interaction with research subjects, etc.).

The federal definitions and guidance used to make this determination may be found at the following link: [Human Subject](#)