Emotion-Conveying Words in Polish Social Media*

Anna Prince, [1] C. Anton Rytting, [2] and Ewa Golonka [2]
[1] Georgetown University, [2] University of Maryland College Park**
map407@georgetown.edu, crytting@umd.edu, egolonka@umd.edu

A growing body of research has attempted to categorize emotions in social media text. However, emphasis on macro-scale trends does not provide a nuanced view of how those classifications are drawn. This article builds on Oberländer's work on semantic role labeling in sentiment analysis, using their 2020 schema of *cue word*, *target*, *cause*, and *experiencer* to examine semantic roles in social media posts. Using a corpus of geopolitical Polish-language Facebook data annotated for the presence and intensity of 23 distinct emotions, we generate three hypotheses regarding the actors and emotions in our data. We use two subcorpora of posts containing contempt and admiration, emotions that are roughly bivalent and under-researched in the current literature. Our findings suggest that part-of-speech is not a relevant consideration, and that emotion-conveying words are monovalent–that is, they do not signal multiple emotions in different contexts. We also find differences in the semantic roles towards which our two bivalent emotions are directed, as well as the relative intensity with which they are expressed. We hope this exploratory study can inform future research on the integration of semantic role labeling and sentiment analysis.
Keywords: sentiment analysis; semantic role labeling; emotion; social media

1 Background

A standard definition of *emotion* is hard to pin down, in part because scholarly understandings have evolved separately across a variety of disciplines. Early social-psychological research includes Plutchik's (1980) typology, which included eight basic emotions thought to be universal, followed by Ekman's basic emotions derived from facial expressions (Ekman, 1992), and modern

*We are grateful to Agata Bieniek, Anna Kostrzewa, Agata Kuzia, Klaudia Kuźnicka, Małgorzata Perczak-Partyka, Rafał Rosiak, Ewa Szczepska, and Marta Urbańska-Łaba for their annotation work, and Susannah Paletz for her input on this project.
**This material is based on work supported, in whole or in part, with funding from the United States Government Minerva Research Initiative / Office of Naval Research Grant #N00014-19-1-2506 and internal University of Maryland Applied Research Laboratory for Intelligence and Security funds. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the University of Maryland, College Park or the United States Government.

approaches embracing a large number of emotions that are usually extracted from text using computational methods (e.g., Cowen & Keltner, 2021). Though criticized for not being based on modern psychological theory or definitions of emotion (Ortony, 2022), Plutchik's work has informed more recent research within computational social science communities on emotional typologies, including that of Mohammad (2018), which quantified the affective *intensity* of individual words associated with Plutchik's emotional classifications. This research has informed more recent typologies such as that of Paletz et al. (2022), whose 23 emotions are the drawing board of our current study.

For consistency, we defer to a definition provided by our parent project, *Emotions in Social Media* at the University of Maryland. This project examines emotion as "feelings that arise in the interpretation of events, are culturally expressed and understood, have underlying universal bases, and involve some kind of signals or expression" (Paletz et al., 2022, p. 2). This definition, which emphasizes the interaction between universal experiences and individual action, is consistent with recent scholarship defining emotion as "appraisals, experiences, expressive behavior, physiological response, influences upon ensuing thought and action, and language-based representations" (Cowen and Keltner, 2021, p. 125). These definitions are particularly useful because they emphasize a connection between emotion and language–a major means for expressing emotion, and the primary medium used to infer emotion within most computational social science research.

Other researchers have explored computational tools and linguistic resources attempting to narrow down specific relationships between the two, many of them open-source and foundational to this project (see WordNet (Fellbaum, 1998), WordNet Affect Lexicon (Strapparava & Valitutti, 2004), EmoLex (Mohhammad & Turney, 2013), or plWordNet (Maziarz et al., 2016), to name a few). Such tools generally fall under the umbrella of *sentiment analysis*, or the automatic detection of emotion within linguistic data. Yet, as useful as these resources are, there is still limited understanding as to *how* emotion is configured through language. This gap is articulated by Masjid (2012), who, noting the dynamic and multi-tiered interactions between the two, argues that "emotional expression is finely tuned to language-specific structures" deserving of further empirical research. Prince (2022) notes that emotion lexicons often fall short in capturing topic-driven public discourse, such as geopolitical social media data. Mohammad and Turney (2013) also identify contextual ambiguity and sense-scoping issues as potential limitations of lexicons in general.

As such, other scholars have attempted to integrate semantic role labeling into sentiment analysis techniques. These approaches generally follow *Frame Semantics Theory* (Fillmore, 1976, 1982), which argues that the meaning of a word in a sentence depends on its semantic relationship to the words around it. Semantic role labeling attempts to uncover what words "do" in a sentence; in this context, how words convey emotion, by whom, and towards whom. One useful schema for understanding semantic roles in emotion detection is that of Oberländer et al. (2020), who develop the terms *cue word*, *target*, and *experiencer*. These terms describe words that evoke an emotion in the reader, the person or object at which the emotion is directed, and the person or object who experiences the emotion, respectively. As an example, take the following sentence, extracted from a social media post in our corpus:

> [Polish]: *Matoły w sejmie, pajace w koalicji. Tylko w ludziach ostatnia nadzieja.*
>
> [English]: *Dummies in the parliament, clowns in the coalition. Only in [ordinary] people is the last hope.*

In this sentence, *[m]atoły* (unintelligent people) and *pajace* (clowns) were identified through annotation as "cue words," or words triggering a specific emotion–in this case, contempt. This contempt is felt by the author of the post (the "experiencer"), and directed towards politicians (the "target"). We use this schema to develop the following exploratory research questions (RQs):

RQ1. What parts of speech are cue words? How does their grammatical distribution vary across context and emotion?
RQ2. Is there a difference between the prevalence of cue words in describing targets and experiencers?
RQ3. Is there a difference between emotions describing targets and experiencers?

2 Methodology

This work closely follows the *Emotions in Social Media* project at the University of Maryland, which curated a corpus of social media texts taken from Facebook pages of Polish sociopolitical influencers and annotated a sample of 3,649 of these posts for 23 distinct emotions on a 0-100 intensity scale using the Social Media Emotions (SMEmo) annotation guide (Paletz et al., 2022). We refer here to the portion of the Polish Facebook corpus annotated for emotions as the SMEmo Polish corpus. This project examines one pair of antithetical emotions–*contempt* and *admiration*–because they are both well-represented in the SMEmo Polish corpus (Paletz et al., 2023), but not in the greater literature. The SMEmo annotation guide, adapting research from Ekman (1992) and Ekman and Corado (2011), defines *contempt* as "disregard, condescension, disdain, looking down on someone, feeling superior to someone or something, or having no respect for the other party and what they are doing" (p. 10). Conversely, adapting Cowen and Keltner's (2017) definition, the SMEmo guide defines *admiration* as "respect and appreciation for a person or thing in a way that is distinct from love or sexual attraction [… a] positive emotion associated with a specific person, object, or group that does not entail a long-lasting, mutual bond, but feeling impressed and amazed at another's traits or actions" (p. 12). While some studies (e.g., Lunando & Purwarianti, 2013; Maynard & Greenwood, 2014) have explored contempt in the context of sarcasm, little research has examined either of these emotions in their own right. This is surprising, as our research showed that almost a quarter of all posts in the SMEmo Polish corpus (923 out of 3,649) conveyed contempt, and an even greater proportion (1,051 out of 3,649) conveyed admiration.

Our dataset selected for this study is a subsample of the posts annotated for admiration and/or contempt, consisting of 591 posts, with 300 posts conveying admiration and 300 conveying contempt. Nine posts conveyed substantial levels of both admiration and contempt, and thus counted for both subcorpora. While these posts were randomly selected from the SMEmo Polish corpus, we narrowed selection criteria to account for length, intensity, geopolitical variance, and access to attached multimedia. We excluded posts with fewer than 20 or more than 160 words, since these posts had too much or too little content to identify specific thematic roles. Intensity of the relevant emotion (contempt or admiration), using SMEmo's 0-100 scale, was set at a minimum of 20. We set this threshold to avoid low-level or covert instances of emotion.

In order to minimize selection bias, we also drew from a range of different political events. The posts selected for annotation and inclusion in the SMEmo Polish corpus were sampled within a defined temporal proximity to one of four political events occurring during the period over which it was collected (2015-2020), which include–briefly summarized–two political elections, the Czarny women's strike, and a COVID-19 lockdown. As such, we automatically included ten posts from each event-centered subcorpus conveying the most contempt and admiration, respectively, for a total of 80 pre-selected posts. We find that the remaining posts, which were randomly generated, do not over-represent any of the four events.

Using the semantic role framework established by Oberländer[1] et al. (2020), the following prompts, adapted from Bostan et al. (2020) and translated to Polish were used to guide annotation:

[1] Bostan and Oberländer appear to be the same scholar.

1. Which words helped you in identifying the given emotion?
2. Is the experiencer of the emotion mentioned in text, media, both, or neither?
    a. If yes, who are they?
    b. If there are words describing the experiencer, please list them.
3. Who or what is the emotion directed at?
    a. If there are words describing the experiencer, please list them.
4. Select the words that explain what happened that caused the expressed emotion.

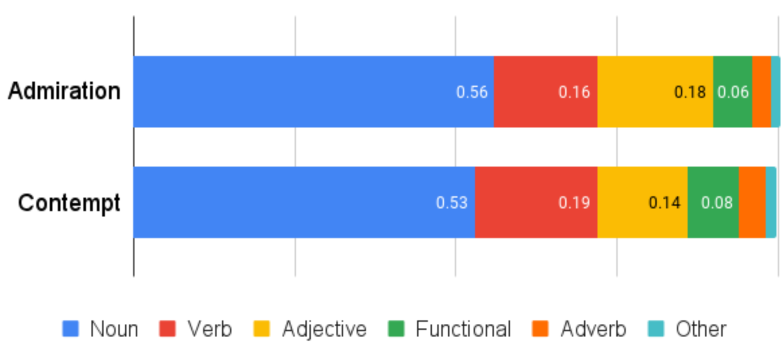
We asked our annotator, a native Polish speaker familiar with the geopolitical context of the data, to respond to these prompts on a shared virtual spreadsheet. Responses were open-ended except for Prompt 2, which required one of four pre-configured responses. For RQ1 and RQ2, we used the SpaCy natural language processing library. This is consistent with the emotion lexicon used for RQ3, which was developed by Prince (2022) using the SpaCy python library and the full SMEmo Polish corpus. We removed stop words, or semantically insignificant yet commonly occurring words, using a Polish-language list provided by Paletz et al. (2022). This list mirrors, but is larger than, the Polish-language stopword list available in the SpaCy python library.

3 Results

3.1 RQ1: Part-of-Speech Distribution of Cue Words

Figure 1 shows the breakdown of parts-of-speech of words identified in Prompt 1 (hereafter, "cue words"). Note that there were slightly more admiration cue words than contempt cue words, with roughly 2,400 cue words in each corpus.

Figure 1. *Part-of-speech Comparison from Our Corpus*

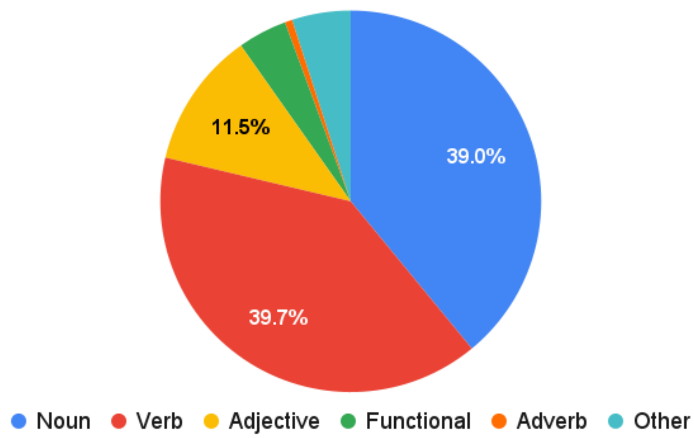


Note: we manually aggregated these words into broader types.

While adjectives are better-represented in admiration cue words, and verbs better-represented in words conveying contempt, an unpaired t-test shows that these differences are not statistically significant.

While we could not find other Polish-language corpora annotated for cue words, the English-language *GoodNewsEveryone* corpus (Bostan et al. 2020), which informs our study, is publicly available for comparison. Using the English-language SpaCy pipeline, the composition of the 6,690 annotator-identified cue words in *GoodNewsEveryone* is seen below. Stopwords were removed using the default English-language SpaCy list.

Figure 2. *Parts-of-speech of All Cue Words in GoodNewsEveryone Corpus*



Note: GoodNewsEveryone uses an emotional typology that does not include contempt and admiration. Direct comparisons are therefore not possible.

As in our corpus, the major part-of-speech categories are nouns, verbs, and adjectives. However, verbs are far better-represented in *GoodNewsEveryone* than in our corpus. By extension, nouns are proportionately less.

3.1.1 Discussion

Manual review of both corpora revealed complications with participles, which were sometimes misclassified as verbs or adjectives (e.g., the difference between “a wrinkled shirt” and “he wrinkled the shirt”). Still, the similarity between results in our two subcorpora are notable, especially in contrast with results from the *GoodNewsEveryone* corpus. It is possible that bivalent cue words have semantically similar roles, or that Polish cue words *in general* assume a similar part-of-speech distribution, regardless of the emotion they convey. However, these findings do suggest that part-of-speech analysis between bivalent emotion-conveying words is not a useful direction for further research.

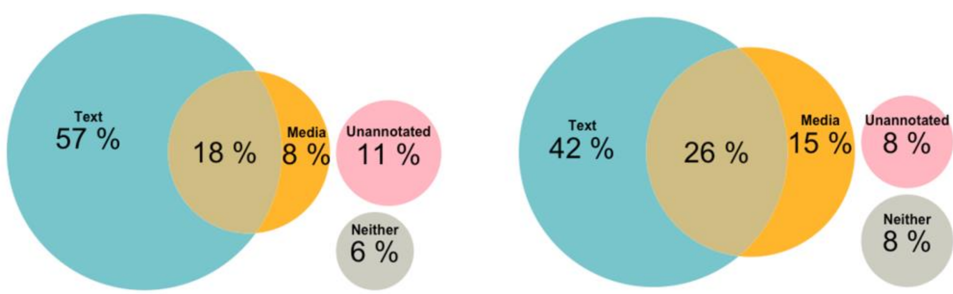
The differences between our identified cue words and those of Bostan et al. (2020) could indicate a few areas of exploration. *GoodNewsEveryone* is a collection of newspaper headlines, which assume a different linguistic style and audience than personal Facebook posts. These differences could implicate different thematic roles, if newspapers tend towards impersonal and verb-heavy descriptions of events. They could also implicate reader-side expectations, if people *perceive* less emotion in an "objective" newspaper article than an equivalent Facebook post. People may also engage with social media and newspapers for different purposes, or in different emotional states, which could influence their emotional experience in consuming a given piece of content. However, cross-language comparisons are inherently difficult, even with multi-language tools such as SpaCy, and further research would benefit from comparisons across the same language.

3.2 RQ2: Representations of Semantic Entities

3.2.1 Identification of Experiencers in Text and Media

Figure 3 shows the representation of experiencers in text and media (Prompt 2). In posts conveying both admiration and contempt, roughly 10% of posts were not annotated due to the unavailability of attached multimedia, and roughly 7% of posts did not mention an experiencer anywhere in the post (in either text or other media).

Figure 3. *Identification of Experiencers*



Posts evoking admiration

Posts evoking contempt

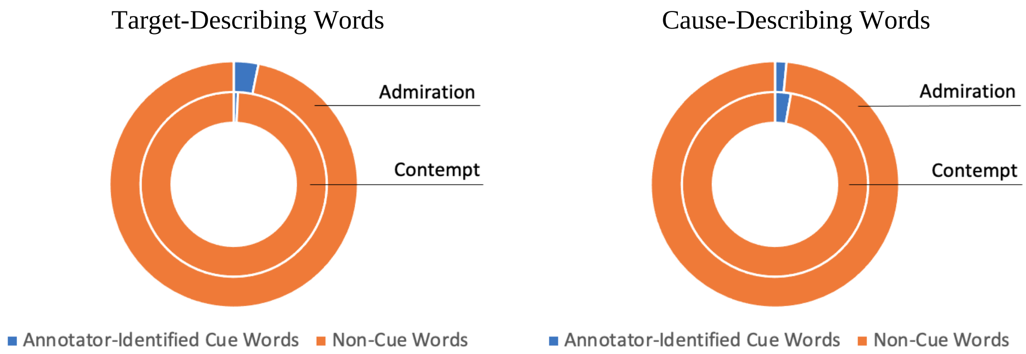
Within the roughly 83% of posts that do mention an experiencer, posts conveying admiration and contempt both tend to rely on text rather than media to identify them, and are about twice as likely to mention an experiencer in text *and* media than media alone. However, posts evoking contempt appear to more often identify experiencers with multimedia than posts evoking admiration. This includes experiencer references made exclusively through multimedia, as well as references made alongside text.

3.2.2 Prevalence of Cue Words in Descriptions of Targets and Causes

We did note differences in the prevalence of cue words as descriptors of both targets and causes. Words that describe a target more often conveyed admiration than contempt. Conversely, words describing a cause more often conveyed contempt than admiration. Out of the 677 words that contemptuously describe a target (Prompt 3a), only five had been previously identified as contempt cue words (Prompt 1). Yet out of the 651 words that admiratively describe a target, 21 had been identified as cue words. Using a two-proportion z-test, the calculated z-value is -3.2705 (p = .00108), indicating a statistically significant difference. Conversely, of the 5177 words that contemptuously describe a cause (Prompt 4), 143 were contempt cue words, whereas the 3148 admirative cause-describing words only contained 47 admiration cue words. This difference is also significant with a z-value of 3.7603 (p = .00016).

The number of cause-describing words is substantially higher than the number of target-describing words. We speculate that targets may have been described with single adjectives or nouns, whereas causes required complex grammatical constructions (e.g., “that happened yesterday,” “in the warehouse,” et cetera). While we excluded stop words, it is possible that our collection of cause-describing words is inflated by words that are normally semantically significant, but not in this context.

Figure 4. *Cue Words as Target and Cause Descriptors*



3.2.3 Discussion

From the observed differences in where specific cue words are found–in words describing targets of admiration, but causes for contempt–it seems that the locus of the emotion differed between the two emotions studied. The exact reasons for this are not clear from the statistics alone. It is possible that words describing a cause trigger contempt more often than admiration, since causes are often events or circumstances rather than people. It is harder to admire an impersonal event, such as a hurricane or election, than it is to feel frustrated or bitter about it. It is also possible that the sampled posts in our corpus praised targets of admiration for who they *were* (i.e., with words describing stable, admirable character traits) but criticized targets of contempt for what they *did* (i.e., how they contributed to or (re)acted during events annotated as causes for contempt).

However, it is surprising that words evoking contempt rarely described targets, given the sheer amount of contemptuous content identified in the broader corpus. One explanation could be the use of multimedia identified discussed above, which suggested that posts containing contempt made greater use of multimedia than posts containing admiration. It is possible that contempt towards targets was conveyed visually, such as through critical pictures and videos, rather than through words. This is consistent with the fact that within Prompt 1, the subcorpus of words conveying contempt was smaller than the subcorpus of words conveying admiration. It is also possible that contempt was more frequently conveyed indirectly–e.g., through irony, sarcasm, humor, innuendo, or other rhetorical devices (cf. Lunando & Purwarianti, 2013; Maynard & Greenwood, 2014)–whereas that admiration was more frequently conveyed directly and straightforwardly in our corpus.

3.3 RQ3: Range and Intensity of Emotions

3.3.1 Valence

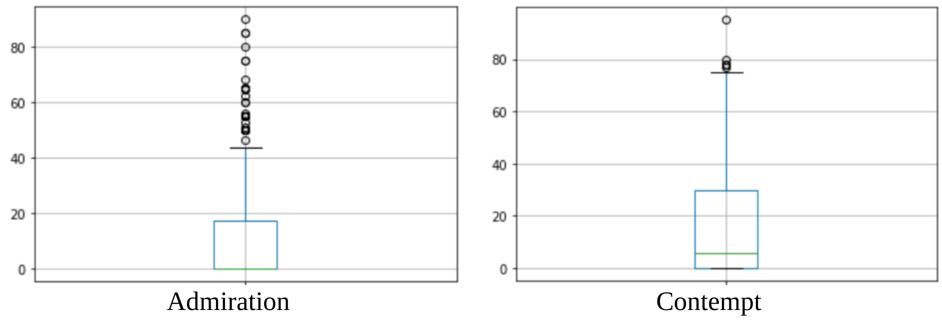
Among all words identified in Prompt 1, we find only a 6% overlap between words in the contempt subcorpus, and words in the admiration subcorpus. Within that small percentage, almost all words were proper nouns or functional words that were not included in the stop word list. This percentage also includes cue words from the nine ambivalent posts, which necessarily overlap as they appeared in both subcorpora. These findings suggest that cue words are monovalent; that is, words that convey emotions tend to convey the same emotion consistently, regardless of context–at least for this pair of emotions.[2] Proper nouns do create an interesting exception, since it would appear that the target of the emotion can itself evoke the emotion. We speculate that in such instances, there may have been visual or other cues to contextualize the emotion, or perhaps that these proper nouns name polarizing figures in contemporary Poland, who evoke admiration among some Poles and knee-jerk contempt in others.

3.3.2 Intensity

Figure 5 shows the emotional intensities of cue words identified in Prompt 1. Following the SMEmo 0-100 emotional intensity scale, these graphs represent the emotional intensities of individual words, according to an emotional lexicon developed by Prince (2022).

[2] Naturally, since we only looked at contempt and admiration, there is nothing in these findings that preclude the same word conveying similar emotions not examined here: e.g., the same word conveying both admiration and love, or conversely, a single word conveying contempt, hate, and anger. However, these results suggest that none of the words that were annotated as cueing admiration were used elsewhere ironically to cue contempt–or if they were, our annotator found more compelling words to mark as cues for contempt.

Figure 5. *Word-Level Intensities of Annotator-Identified Cue Words*



The intensity of individual words identified in Prompt 1 is rather low. Among words identified as conveying admiration, the median admirative content is zero, with the third quartile still falling below 20–the baseline intensity we required for posts. Words conveying contempt appeared slightly more emotive; while the median still fell below 20, a majority of words were individually associated with some level of contempt, with the upper quartile scored as 30 or higher. This is notable given that, as established in RQ2, posts conveying contempt made greater use of multimedia and used fewer textual descriptors overall. It could be the case that multimedia content is somehow related to emotion in text. Perhaps intense emotional states lead people to use multimedia more, or multimedia intensifies emotion already present in a text. Alternatively, it may be an algorithmic issue, in the sense that larger corpora dilute the intensity of words within them.

3.2.3 Discussion

While we found that individual *words* in our corpus consistently conveyed discrete emotions–either contempt or admiration, but not both–in nine instances the *post* did not. This raises the question of how ambivalence can exist within a text, if the words within it convey specific emotions. The general pattern we observed in our corpus is that these posts specify different targets for the respective emotions. Furthermore, in some instances the juxtaposition of opposing emotions, such as contempt and admiration, intensified one another via contrast. In all nine ambivalent posts, we identified at least one target of each respective emotion. One of these ambivalent posts, taken from a high-profile public figure in our dataset, provides a particularly salient example:

> *[Polish]: Prawicowi fanatycy przegięli. Dziesiątki tysięcy kobiet, które pokazały już na ulicach, że stawią im opór, to dopiero początek. Wywołali lawinę. Teraz ta lawina spadnie na ich głowy.*

[English]: Right-wing fanatics have gone too far. The tens of thousands of women who have already shown in the streets that they will resist them is only the beginning. They caused an avalanche. Now this avalanche will fall on their heads.

Here, the author's admiration for the marchers ("tens of thousands of women") is paired with contempt for their political opponents ("right-wing fanatics"). Yet, the identified cause is the same, with the Czarny protests identified as the cause of both emotions. While this contrast is clear enough to an attentive human reader, for a machine performing an emotion inference task, this will only be clear with a semantic labeling approach to ground-truth annotation, which takes a more granular look at the emotional dynamics within the entire post.

Notably, these findings are not inconsistent with our earlier point on monovalence. Whether words convey emotion, and which emotions those words convey, are different questions; that is, a word that is discreetly associated with one emotion may not signal it in all contexts. This would suggest that there is not a linear relationship between the amount of emotion-signaling words in a text, and the emotional intensity of the text at large. Thus, while word-level emotional analyses are useful, they do not go far enough in explaining how those words become "activated" in context. This point again stresses the utility of semantic roles in sentiment classification tasks.

4 Conclusion

This exploratory study provides several useful directions for further research. Primarily, we have argued that current sentiment analysis techniques may overlook emotional dynamics within individual posts. Given that we did find differences in emotional descriptions of different actors, such as targets and experiencers, we argue that integrating semantic role labeling into current research could benefit future sentiment analysis algorithms. Additionally, given our findings that cue words are monovalent while posts may convey multiple and conflicting emotions, we raise the possibilities of multiple layers of referent within a text, as well as interplay between text and visual media. Thus, we argue that focusing on specific words and word-types, as well as their semantic role within a sentence, could be an efficient way of capturing some of that nuance and compensating for visual information.

4.1 Limitations

We do note a few limitations in this exploratory study. First, our sample size of 600 posts is somewhat small. While larger samples may be useful for future research, we believe that this exploratory study can still be useful in guiding that research. Second, since the SMEmo Polish corpus was collected several years ago, some multimedia were no longer available at the time of annotation (roughly 10% of posts, as noted in RQ2). Since we chose posts randomly, within certain parameters, we hope that bias coming from unavailability of multimedia is limited. And finally, we limited our analysis to responses from a single annotator, as open-ended response schemas lend themselves to inconsistent annotations. Difficulty ensuring inter-annotator agreement was also noted by Bostan et al. (2020), who pointed to the inherent subjectivity of emotion in textual content. We hope that the refinement of these annotation methodologies, as well as heavier investment into annotator training, can make these tasks easier for future researchers. Nevertheless, despite issues with annotator training and consistency, we believe this study provides support for the general feasibility of the annotation framework they propose, by extending it to another language and text type.

4.2 Implications and Future Research

More research into the linguistic mechanics of emotion would greatly benefit sentiment analysts across a variety of domains. Adjacent research projects under the *Emotions in Social Media* grant find clear connections between emotion and willingness to share information, which presents clear applications to disinformation studies. Golonka et al. (2023), for instance, partially address our question regarding the interaction between emotion conveyed in text and media by implicating "cute" images as a vector for social media influence. Analyzing semantic roles in conjunction with non-textual data could provide a more complete picture of the social media [dis]information environment, providing better intelligence to social media companies about malicious activity on their platforms. Large-scale data about specific political grievances, or which actors feel what emotion towards which targets, could be particularly useful to human rights advocates or law enforcement.

The medical field may also benefit from further research. Recent studies have used sentiment analysis techniques to better understand the emotional experiences of people with chronic diseases, such as cancer (Edara et al., 2023), and their perceptions of specific treatments, such as anticonvulsant medications (Mathieson et al., 2022). A number of researchers have also used sentiment analysis of social media data to identify psychiatric disorders and assess suicide risk (see Bittar et al., 2021; Sawalha et al., 2022), which could facilitate more effective interventions at both the individual and community levels. Though these tools carry a number of complex ethical considerations, they have clear applications for saving and improving lives, particularly if classification algorithms can effectively cut through the noise to extract specific sources of patient discomfort. These are fundamentally semantic role-labeling tasks and should be treated as such.

5 About the Authors

We hope this project shows, by example, that community-building across disciplines can facilitate innovative research. Anna Prince is a former research intern at the University of Maryland Applied Research Laboratory for Intelligence and Security (ARLIS), and a current undergraduate student at Georgetown University. She is majoring in linguistics and government, and minoring in tech ethics. C. Anton Rytting and Ewa Golonka are Associate Research Scientists at the University of Maryland ARLIS. Dr. Rytting holds a PhD in Linguistics from the Ohio State University, with a specialization in computational linguistics. His current research focuses on how authors' emotions, personality traits, values, and identity are revealed in texts of various genres. Dr. Golonka's research has focused on instructed second language acquisition, multilingualism, cognitive aptitude, and the analysis of social media corpora in various languages. She holds a PhD in Russian and Second Language Acquisition from Bryn Mawr College.

References

Bittar, A., Velupillai, S., Roberts, A., and Dutta, R. (2021). Using general-purpose sentiment lexicons for suicide risk assessment in electronic health records: Corpus-based analysis. *JMIR Med Inform*, 13(9). doi: 10.2196/22397

Bostan, L., E. Kim, & R. Klinger. (2020). GoodNewsEveryone: A corpus of news headlines annotated with emotions, semantic roles, and reader perception. In *Proceedings of the 12th Language Resources and Evaluation Conference*, 1554–1566. https://aclanthology.org/2020.lrec-1.194

Cowen, A.S., & Keltner, D. (2021). Semantic space theory: A computational approach to emotion. *Trends in Cog. Sciences* 25(2), 124–136. https://doi.org/10.1016/j.tics.2020.11.004.

Edara, D.C., Vanukuri, L.P., Sistla, V., and Kolli, V.K.K. (2023). Sentiment analysis and text categorization of cancer medical records with LSTM. *Journal of Ambient Intelligence and Humanized Computing 14*, 5309–5325. https://doi.org/10.1007/s12652-019-01399-8

Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion*, 6(3–4), 169–200. https://doi.org/10.1080/02699939208411068.

Ekman, P., and Cordaro, D. (2011). What is meant by calling emotions basic? *Emotion Review*, *3*(4), 364–370. https://doi.org/10.1177/1754073911410740

Fellbaum, C. (1998, ed.) *WordNet: An electronic lexical database.* Cambridge, MA: MIT Press.

Fillmore, C. (1982). Frame semantics. In *Linguistics in the morning calm: Selected papers from SICOL-1981*, 111–137. Hanshin Pub. Co. http://brenocon.com/Fillmore%201982_2up.pdf

Fillmore, C. (1976). Frame semantics and the nature of language. In *Annals of the New York Academy of Sciences: Conference on the Origin and Development of Language and Speech* 280, 20–32. doi:10.1111/j.1749-6632.1976.tb25467.x

Golonka, E.M., Jones, K.M., Sheehan, P., Pandza, N.B., Paletz, S., Rytting, C.A, and Johns, M., (2023). The construct of cuteness: A validity study for measuring content and evoked emotions on social media. *Frontiers in Psychology*, 14:1068373. doi: 10.3389/fpsyg.2023.1068373

Lunando, E., & Purwarianti, A. (2013). Indonesian social media sentiment analysis with sarcasm detection. *2013 International Conference on Advanced Computer Science and Information Systems (ICACSIS)*, 195–198. Doi: 10.1109/icacsis.2013.6761575

Masjid, A. (2012). Current emotion research in the language sciences. *Emotion Review*, 4(4), 432–443. https://doi.org/10.1177/1754073912445827

Mathieson, S., O'Keefe, M., Traeger, A.C., Ferreira, G.E., & Shaheed, C.A. (2022). Content and sentiment analysis of gabapentinoid-related tweets: An infodemiology study. *Drug and Alcohol Review*, 1–11. https://doi.org/10.1111/dar.13590

Maynard, D., & Greenwood, M. (2014). Who cares about sarcastic tweets? Investigating the impact of sarcasm on sentiment analysis. In *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*, 4238–4243, Reykjavik, Iceland. European Language Resources Association (ELRA).

Maziarz, M., Piasecki, M., Rudnicka, E., Szpakowicz, S., and Kędzia, P. (2016). plWordNet 3.0 – A comprehensive lexical-semantic resource. In *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers*, 2259–2268, Osaka, Japan. The COLING 2016 Organizing Committee.

Mohammad, S. (2018). Word affect intensities. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*, Miyazaki, Japan. European Language Resources Association (ELRA).

Mohammad, S., and Turney, P. (2013). Crowdsourcing a word-emotion association lexicon. *Computational Intelligence*, 29(3), 436–446. https://doi.org/10.48550/arXiv.1308.6297

Oberländer, L. A. M., Reich, K., & Klinger, R. (2020). Experiencers, stimuli, or targets: Which semantic roles enable machine learning to infer the emotions?. In *Proceedings of the Third Workshop on Computational Modeling of People's Opinions, Personality, and Emotions in Social Media*. (pp. 119–128). https://doi.org/10.48550/arXiv.2011.01599

Ortony, A. (2022). Are all "basic emotions" emotions? A problem for the (basic) emotions construct. *Perspectives on Psychological Science*, 17, 41–61. https://doi.org/10.1177/1745691620985415

Paletz, S. B. F., Rytting, C. A., Johns, M. A., Pandža, N. B., Golonka, E. M., Murauskaite, E. E., & Buntain, C. (2023). *Emotions in Polish and Lithuanian social media* [dataset]. Digital Repository at the University of Maryland. https://doi.org/10.13016/b4lh-j5tg

Paletz, S. B. F., Golonka, E. M., Stanton, G., Murauskaite, E., Ryan, D., Rytting, C. A., & Bradley, P. (2022). Social media emotion annotation guide (SMEmo), Version 4.0. University of Maryland College Park Applied Research Laboratory for Intelligence and Security.

Plutchik, R. (1980). *Emotion: A psychoevolutionary synthesis.* Harper & Row.

Prince, M.A. (2022, April 24). *Emotional intensity lexicons for event-based corpora*. Poster presented at the 16th Annual Cornell Undergraduate Linguistics Colloquium (CULC16), Cornell University, Ithaca, NY.

Sawalha, J., Yousefnezhad, M., Shah, Z., Brown, M.R.G., Greenshaw, A.J., & Greiner, R. (2022). Detecting presence of PTSD using sentiment analysis from text data. *Frontiers in Psychiatry,* 12:811392. doi: 10.3389/fpsyt.2021.8113

Strapparava, C., & Valitutti, A. (2004). Wordnet-affect: An affective extension of WordNet. In *Proceedings of the 4th International Conference on Language Resources and Evaluation (LREC-2004)*,1083–1086, Lisbon, Portugal. http://www.lrec-conf.org/proceedings/lrec2004/pdf/369.pd