Examining Gender Differences in Language: A Computational Analysis of Emotion in the Dutch Veteran Institute Oral History Archive

Bachelor's Thesis
Artificial Intelligence
Utrecht University
7.5 ects

Author:
Yanthe Roumen
6218245

Supervisor:
Francisca Pessanha

Second Reader:
David Terburg

15 August 2021

Abstract
Gender differences in post-traumatic stress symptoms suggest that expressing trauma is influenced by gendered socialization, as gender stereotypes influence the way people express their emotions. In this research, emotional language in 17 interviews of the Dutch Veteran Institute oral history archive was analyzed using applications of artificial intelligence. In these interviews, conflicting and traumatic events were recalled. The interviews were transcribed using Automatic Speech Recognition technology, and analyzed using a keyword-based approach using the Dutch LIWC2007 and LiLaH emotion lexicon. The objective of this research was to examine gender differences in the use of emotional language, and to see whether these differences fell in line with gender norms for emotional expression. It was found that women used more words related to joy, friends and humans, whereas men used more swear words and words describing positions in space. The results show that emotional gender stereotypes are slightly present in the language of this dataset.

Index terms:

gender differences, oral history archives, emotion, natural language processing, keyword-based approach

Table of Contents

1. Introduction

The existence and the nature of differences between men and women has been a popular topic in research over the last several decades. An interesting difference observed between the genders is the difference in the expression of post-traumatic stress symptoms (Green et al., 1997; Darvezs-Bornoz et al., 1998; Tolin et al., 2006; King et al., 2013). These differences suggest that processing and expressing trauma is influenced by gendered socialization. There is a pattern of externalizing behaviour in men after experiencing or witnessing trauma, which is consistent with the emotional stereotype that men are more aggressive (Plant et al., 2000). There is also a pattern of emotional suppression for men, consistent with the emotional stereotype that men are not emotional (Timmers et al,. 2003), that they are more likely to seek distraction (Strauss et al., 1997), and that depression is seen as something that is not masculine (Brody, 1999). With women, there is a pattern of internalizing behaviour, consistent with the stereotype that women use more rumination (Strauss et al., 1997), that they are believed to express more sadness and shame (Plant et al., 2000), and at the same time are discouraged from expressing anger (Brescoll & Uhlmann, 2008). Gender stereotypes influence the way people express emotions and trauma, as people feel the need to exert control over their emotions to align with gender stereotypes (Matsumoto et al., 1998; Simpson & Stroh, 2004).

Oral history archives, containing collections of video and audio interviews, provide insight into the way people experience, recall and process traumatic events. There have been several attempts to use applications of artificial intelligence, like Automatic Speech Recognition (ASR) and other Natural Language Processing (NLP) techniques, to automatically transcribe and analyze these collections of valuable subjective information. For example, interviews of the collection by The Survivors of the Shoah Visual History Foundation (VHF) have been automatically transcribed using ASR, and NLP techniques have been implemented to detect topic boundaries and to classify topics (Byrne et al., 2004; Olsson et al., 2007).

However, not a lot of work has been done on the automatic transcription and analysis of Dutch language, specifically for oral history. An example of a Dutch oral history archive is the archive from the Dutch Veteran Institute (Nederlands Veteranen Instituut, 2021), a collection of almost 1200 interviews. This collection contains audio files, mostly of stories by Dutch veterans who have participated in wars, peace missions or other missions from 1940 onwards. Examining language patterns in these interviews, specifically patterns in the use of words associated with emotion, will give information about the way in which the interviewees recall traumatic events. The words people use provide important information on their thoughts, emotional states, and how they organize, process and analyze the world (Tausczik & Pennebaker, 2010). Finding emotional language patterns that are consistent with emotion gender norms could suggest that processing traumatic events is influenced by gender norms. Finding these gendered language patterns could prove to be useful for algorithms that predict or diagnose mental health statuses.

The objective of the present research was to answer the question whether there are gender differences in emotion in the language of the oral history archive of the Dutch Veteran

Institute, using computational methods. The audio interviews were transcribed using an Automatic Speech Recognition program that was used earlier on this dataset (Veldema, 2021), and a keyword-based approach was used to extract emotional information from the transcripts. Two Dutch emotion lexicons, the Dutch LIWC2007 lexicon (Boot et al., 2017) and the LiLaH lexicon (Ljubesic et al., 2020), were used to detect words in the transcripts that belonged to psychologically meaningful categories, including a range of emotion categories. The research question was: *within which categories of the LIWC2007 and LiLaH lexicon are gender differences regarding emotional expression found in language used in the Dutch Veteran Institute oral history archive*? Based on existing gender norms and previous research on gender differences in language, it was expected that women would express more emotions, especially positive emotions, and that men would express more anger and swear more. The hypothesis was further specified as: women use more words from the Dutch LIWC2007 categories *positive emotion* and *affective processes* and the LiLaH categories *positive* and *joy* than men; men use more words from the Dutch LIWC 2007 & LiLaH categories *anger* and words from the Dutch LIWC2007 category *swear words* than women.

This research intersects language, gender, trauma, oral history and artificial intelligence. It takes small steps into the process of automatically analyzing emotion in oral history archives on a large scale and solving the gap in work on automatic analysis of Dutch language and oral history.

2. Background

2.1 Psychological Trauma

The Substance Abuse and Mental Health Services Administration defines individual psychological trauma as something that "results from an event, series of events, or set of circumstances that is experienced by an individual as physically or emotionally harmful or threatening and that has lasting adverse effects on the individual's functioning and physical, social, emotional, or spiritual well-being" (Substance Abuse and Mental Health Services Administration, 2014). In short, psychological trauma is damage to an individual's mind leading to negative long-term consequences, caused by experiencing or witnessing one or more events which caused harm exceeding the person's ability to cope with the emotional damage. Whether an event is experienced as traumatic depends on different factors. A few of these factors are: how the individual assigns meaning to the event, how they are disrupted by the event, the individual's cultural beliefs, the presence of social support, and the developmental stage of the individual (Substance Abuse and Mental Health Services Administration, 2014). Emotional reactions to trauma are influenced by an individual's sociocultural history. It is common that intense emotions, such as anger, fear, sadness and shame, surface after experiencing trauma. However, individuals may struggle with expressing or identifying surfacing emotions due to various reasons (Substance Abuse and Mental Health Services Administration, 2014).

2.2 Gender & Emotion

One of the reasons individuals encounter difficulties expressing or identifying emotions is due to gender norms. In this section, it is discussed how gender norms related to emotion can influence an individual's behaviour, as gender norms regulate their emotional expressions. To examine existing norms and stereotypes, different research discussed in the chapter Gender, Emotion and Socialization by Brody & Hall (2010) in the Handbook of Gender Research in Psychology will be presented.

The stereotype that women are more emotional is often believed (Timmers et al,. 2003), as well as that women are more emotionally expressive (Briton & Hall, 1995) and more likely to ruminate (thinking repeatedly about negative emotional experiences) than men, whereas men are believed to be more likely to cope using distraction (Strauss et al., 1997). It is believed that happiness, embarrassment, surprise, sadness, fear, shame and guilt are expressed more by women, while anger, contempt, disgust and pride are believed to be expressed more by men (Algoe et al., 2000; Hess et al., 2000; Plant et al. 2000; Parmley & Cunningham, 2008). These beliefs about emotional expression for each gender do not always take into account that social behaviour is often learnt instead of innate. The belief that a gender usually operates in a set way can influence the way one interprets a person of a certain gender (Brody & Hall, 2010). Most of all, it can influence the way one expresses themselves, and whether their feelings or their expressions are believed to be normal or socially appropriate. Violating gender norms for emotional expression is associated with anxiety and shame for men (Osherson & Krugman, 1990), for women expressing anger is associated with conflict and shame (Chrisler, 2008).

Moreover, violating emotional stereotypes can also result in negative social consequences, such as social exclusion, rejection and loss of status. For instance, depressed men were rated as “unmanly” and rated more negatively than depressed women (Brody, 1999). In research by Brescoll & Uhlmann (2008), women who expressed anger were rated as having lower status and competence compared to angry men, sad women and unemotional women. Negative social consequences can be a motivation to exert control over the expression of certain emotions. Women exerted more control over anger, contempt and disgust, and men exerted more control over fear and surprise (Matsumoto et al., 1998). In research by Simpson & Stroh (2004), women reported to feel more pressure to display positive emotions they did not feel and at the same time suppressed negative emotions more than men did, while men reported the opposite; they suppressed more positive emotions and displayed more negative emotions.

In conclusion, gender stereotypes regarding emotional expression are not just intangible ideas; they operate as social norms and have an effect on the material world. These stereotypes act as a baseline of what is socially acceptable and regulate the expression of emotion for each gender, as violating these rules can result in negative feelings or negative social consequences. Both men and women feel a need or pressure to act according to gender stereotypic rules when it comes to expressing emotions and tend to control their emotions to perform according to social norms, even when it does not fall in line with how they truly feel.

2.3 Gender differences in Post-traumatic Stress Symptoms

Gender norms also seem to be present in the expressions of post-traumatic stress. One striking example of gender differences in post-traumatic symptoms is that women are twice as likely to develop Post-traumatic Stress Disorder (PTSD) as men. 8.8% of Dutch women and 4.3% of Dutch men (RIVM, 2021) develop PTSD in their lifetime, whilst men are more likely to be exposed to a traumatic event compared to women (Tolin et al., 2006). Men report experiencing more disasters and nonsexual violence compared to women, while women report experiencing more sexual violence as children and adults than men. That women experience more sexual assault was thought to be an explanation for the higher prevalence in PTSD in women; however, when controlling the types of trauma, women are still more likely to meet PTSD criteria (Tolin et al., 2006). In the military, women and men had the same PTSD rates when the trauma type was controlled. However, women were more distressed than men by violent aspects of warfare specific to wounds and death (Hourani et al., 2014). In research on US veterans deployed in Afghanistan and Iraq, it was found that men and women experienced somewhat equal rates of PTSD, but female veterans reported more frequent concentration difficulties and distress from reminders, whereas men reported more frequent nightmares, emotional numbing and hypervigilance (King et al., 2013). These findings suggest that gender roles may influence women’s and men’s expressions of post-traumatic stress, as the rates of PTSD were similar, although the expressions varied.

Post-traumatic Stress Disorder is only one example of trauma leaving a mark on the human mind and body; other post-traumatic symptoms can arise after experiencing or witnessing trauma. Gender differences in symptom patterns are present in mental health disorders overall. Women may be more likely to report internalizing disorders like anxiety, depression and PTSD, while men are more likely to report externalizing disorders, like substance abuse disorders and conduct disorders (Kessler et al., 1995). Girls reported higher levels of anxiety and depression, whereas boys reported greater aggression after experiencing a dam collapse. Adult women also reported higher levels of anxiety and depression, whereas men reported more aggression and substance abuse (Green et al., 1997). Boys who experienced physical or sexual abuse reported more aggression and violent outbursts compared to girls (Darvezs-Bornoz et al., 1998), and showed symptoms of conduct disorder (Livingston et al., 1993). In research done on nonverbal behaviour indicators of depression and PTSD, women had a higher display and variation of emotion (Stratou et al., 2013). These findings, again, suggest that gender roles may influence expressions of post-traumatic stress across genders.

In conclusion, there is a pattern of externalizing behaviour in men after experiencing or witnessing trauma, such as conduct disorder, aggression and violence, which could possibly be caused by a significant sense of threat, just like the higher prevalence of hypervigilance. Another pattern found in men is emotional suppression, shown by the higher prevalence of emotional numbing and substance abuse. This behaviour falls in line with the emotional stereotype that men are more aggressive, less emotional and that depressed men are seen as unmanly. In women, there is a pattern of internalizing behaviour, shown by the higher prevalence of anxiety, depression and PTSD. This is consistent with the idea that women use more rumination, that they express more sadness, embarrassment, shame, fear and guilt, and that women are discouraged from expressing anger.

Experiencing a traumatic event may amplify pre-existing mechanisms, influenced by social norms, in response to - and in an attempt to process - trauma. Gender expectations might be more tolerant of certain symptoms and discourage other symptoms for each gender, influencing the expression of post-traumatic stress.

2.4 Oral History Archives

One way to gain insight into the thoughts and feelings of people who have experienced or witnessed trauma is through oral history archives. Oral history archives provide a collection of interviews, and focus on the subjective experience and memories of an individual or a group of people that took part in an event that is the subject of study. Digital technology has transformed oral history as a discipline, as oral history can now be recorded, preserved, and curated digitally on a large scale (Pessanha & Akdah Salah, 2021). An example of a large digital oral history archive is the collection by The Survivors of the Shoah Visual History Foundation (VHF), which consists of 52,000 interviews, created to preserve stories of survivors and witnesses of the Holocaust. But even with digital audio and video files, watching, listening or manually transcribing often long interviews is very time consuming. New computational methods, such as Automatic Speech Recognition (ASR) and other Natural Language Processing (NLP) techniques, could provide a more efficient way to curate and analyze oral history archives (Gustman et al., 2002). Natural Language Processing is a branch of artificial intelligence that is concerned with making machines “understand” and produce human language. Automatic Speech Recognition, a technique that recognizes and translates spoken language into text, could be used to automatically transcribe oral history archives. Other NLP techniques could be used to analyze and extract data from oral history, such as topics, emotions and entities. Pessanha & Akdah Salah (2021) explored various computational technologies useful for analyzing both visual (body language, facial expression) and non-visual (paralinguistics, breathing, heart rate) cues present in oral history collections. Attempts have been made to automatically transcribe oral history archives by Byrne et al. (2004), who manually annotated 10,000 hours of speech from the VHF archives to train a ASR algorithm for English and Czech oral history archives. They also used NLP techniques to detect topic boundaries within the transcripts and label them with corresponding semantic categories. Olsson et al. (2007), experimented with ASR transcriptions of interviews from the VHF collection, and applied NLP techniques for automatic topic classification. When it comes to Dutch oral history archives, the amount of automation done is very limited. Ordelman et al. (2005) used an ASR system designed for broadcast news to transcribe an oral history collection on a Dutch novelist, which resulted in transcripts with high error rates.

2.5 Emotion Recognition in Text

Analyzing language and the presence of emotion in transcripts of oral history interviews is one way to extract information on how people process traumatic events. The words we use reflect who we are and sometimes reveal social psychological processes. Tracking language use can provide information on where people lay their attention or how people process a situation or event (Tausczik & Pennebaker, 2010).

In order to recognize emotion in text, a model is needed which defines different emotions and distinguishes them from one another. Two fundamental ways to categorize emotions are on a discrete basis and on a dimensional basis. Discrete emotion models assume that there are a number of universal emotions and place these emotions into distinct classes or categories (Acheampong et al., 2020). An example of a discrete model is the Robert Plutchik (1980) model, which provides 8 emotions: anger, fear, sadness, disgust, surprise, anticipation, trust, and joy. Dimensional emotion models assume that emotions are not independent and that there is a relation between them (Acheampong et al., 2020). An example of a dimensional emotion model is one named by Gunes et al. (2010), which covers emotions in three dimensions: valence - how positive or negative an emotion is; arousal - how excited or apathetic an emotion is; and power - the degree of power an emotion has.

There are different levels of text emotion recognition: document level, paragraph level, sentence level and word level. The complexity of recognizing emotion increases with each level. Difficulties start to arise at the sentence level, where emotions are expressed through multiple words, their meanings and their relations, next to the use of idioms, metaphors and sarcasm (Alswaidan et al., 2020). There are different approaches to recognizing emotions in text. In the following sections, the keyword-based approach will be discussed in depth, and the learning-based approach will be discussed shortly.

2.5.1 Keyword-based approach

Keyword-based approach, often used for explicit emotion recognition (Alswaidan et al., 2020), relies on the presence of keywords in text and emotion lexicons. These lexicons contain emotional information about certain words, in the form of keywords that are mapped to certain emotion categories. For example, the keyword ‘wonderful’ in an emotion lexicon could be mapped to the emotion ‘happy’. Whenever one of these keywords is found in a body of text, something happens depending on the goal of the approach. For example, the sentence in which the keyword is detected gets a label that indicates the corresponding emotion category, or the corresponding emotion category gets incremented for the whole text. With the previous keyword example ‘wonderful’, the sentence “That’s wonderful.” gets the label ‘happy’, or the score of the category ‘happy’ for the whole text gets incremented by one point.

However, the keyword-based approach has some drawbacks. It does not take negators (e.g. ‘not’ in ‘not bad’) and intensifiers (e.g. ‘very’ in ‘very bad’) nor the context of a piece of text into account when detecting keywords, which can cause some problems. For example, ‘break a leg’ is a phrase often said to wish someone good luck, but it is possible that the word ‘break’ alone maps to a negative emotion label and that the sentence gets labeled as negative, while it is a positive phrase. The sentence “That’s not good” may get a positive label because of the word ‘good’ because it does not take the negator into account.

Several lexicons have been created to classify certain words in classes of emotional states or categories. Dutch emotion lexicons are rather scarce, only some are available. In the following subsections, the words in a body of text which is the subject of analysis will be referred to as *target words*, and the words listed in an emotion lexicon will be referred to as *keywords*. A description will be given of some emotion lexicons, their Dutch translations and the differences between the Dutch translations.

2.5.1.1 Linguistic Inquiry and Word Count

The Linguistic Inquiry and Word Count (LIWC) was created by Pennebaker et al., initially in 1993. It was developed to gain insight into the psychological state of the author of a text, especially those of traumatized patients. LIWC is a software program that calculates the percentage of target words of a body of text that match the keywords and their corresponding psychologically meaningful categories, such as affect words, social words and cognitive processes. It also contains linguistic categories, such as articles, verbs and pronouns (Pennebaker et al., 2001). LIWC provides a lexicon with English words, which each are mapped to one or several categories. For example, the keyword ‘cried’ is mapped to the following categories: sadness, negative emotion, overall affect, verb and past tense verb. Over the years, several versions of LIWC have been developed. The 2007 version also captures word stems; these words end with an asterisk. For example, the keyword ‘hungr*’ also counts ‘hungry’, ‘hungrier’ and ‘hungriest’ when it is detected in a text. LIWC2007 contains 4500 keywords (Pennebaker et al., 2007).

2.5.1.2 Dutch LIWC2007

The lexicon of the 2007 version of Linguistic Inquiry and Word Count has been translated into Dutch by Boot et al. (2017). This translation was done manually and resulted in a lexicon of 11091 words in 66 categories. Two new categories were added to the existing categories of the English LIWC lexicon, namely one for singular and plural third person pronouns and one for pronominal adverbs, such as ‘waarmee’ or ‘daarin’. All the emotion categories of the Dutch LIWC2007 lexicon with examples of words belonging to those categories are displayed in table 9.1 in the appendix.

2.5.1.3 NRC Word-Emotion Association Lexicon

Another large English emotion lexicon is the NRC Word-Emotion Association Lexicon, also known as EmoLex (Mohammad & Turney, 2013). It provides keywords and their associations with 10 categories, namely 8 basic emotions: *anger, fear, anticipation, trust, surprise, sadness, joy* and *disgust*, together with two sentiments: *positive* and *negative*. For example, the keyword ‘abhorrent’ has a negative sentiment and the associations *anger*, *disgust* and *fear*, the keyword ‘spa’ has a positive sentiment and the associations *anticipation, joy, surprise* and *trust*. It contains not only monograms (terms containing one word) but also bigrams (terms containing two words). The lexicon uses the Plutchik emotion model, since these classes can be paired in opposites, and potentially used for automatically computing word-pair antonymy. The lexicon was created through Amazon’s Mechanical Turk, an online crowdsourcing platform.

2.5.1.4 LiLaH

The LiLaH lexicon is a translation of the English NRC Emotion Lexicon, with translations into Croatian, Slovene and Dutch (Ljubesic et al., 2020). It uses the same discrete model of emotion as the English NRC Emotion Lexicon. It was first automatically translated, and then manually corrected. The Dutch version has 6468 keywords. All emotion categories of LiLaH and examples of Dutch words belonging to those categories are displayed in table 9.2 in the appendix.

2.5.1.5 Comparing Dutch LIWC2007 and LiLaH

The biggest difference between the Dutch LIWC2007 lexicon and the LiLaH lexicon is that they use different emotion models. LiLaH only provides 10 categories: eight basic emotions and two sentiments; whereas the Dutch LIWC2007 lexicon provides in total 66 categories, including *social processes, affective processes, perceptual processes, biological processes, personal concerns* and *spoken categories*, as well as several linguistic dimensions. The emotion categories of Dutch LIWC2007 are named under *affective processes* and provide only one category for positive emotions, whereas LiLaH provides several positive emotion categories, such as *positive, joy* and *trust*.

The Dutch LIWC2007 lexicon has a different format than the LiLaH lexicon. In the Dutch LIWC2007 lexicon, keywords are mapped to one or several numeric values, which are in turn mapped to category labels. In the LiLaH lexicon, per keyword a binary score is given to each of the 10 categories, with a score of 1 meaning that the word belongs to this category, and a score of 0 meaning that it doesn’t.

The lexicons also vary in processing homonyms (words with the same spelling but a different meaning). In Dutch LIWC2007 a word that has multiple meanings only has one instance, mapped to all possible categories any interpretation of the word may have. In LiLaH, homonyms have separate instances, each with scores corresponding to that individual instance. For example, ‘zuivering’ has two instances, one corresponding with the categories *positive* and *trust,* one with the categories *negative* and *fear*.

2.5.2 Learning-based approach

The learning based approach, introduced to recognize implicit emotion in text (Alswaidan et al., 2020), uses learning algorithms to classify text input into pre-defined emotion categories. A text classifier is trained using a large dataset, called the training set, which consists of pieces of text that are often already labeled with corresponding emotion categories. Based on certain features of these texts and their emotion label, the classifier finds a pattern that predicts the emotion label for new input. Finding important features that provide a correlation between features and the emotion label is called feature extraction. Common features used in text classification are, but not limited to: bag-of-words (BOW), n-grams, part-of-speech (POS) tags and word positions in a sentence.

2.6 Automatic Text Analysis & Gender

Both the keyword and learning-based approach have been used to analyze text and explore differences in language use between men and women. This section discusses previous research. Mehl & Pennebaker (2003) researched language use in natural conversations of 52 psychology students. Their conversations were recorded, transcribed and then analyzed using Linguistic Inquiry and Word Count (LIWC). Men were found to use more big words (more than 6 letters long), more articles, fewer first person singular pronouns, and fewer discrepancy words than women. Men used four times as many swear words as women, less filler words and less references to positive emotions. Newman et al. (2008) analyzed the gender differences in language use in a large sample of 14,000 text documents using LIWC. Women used significantly more pronouns, social words, verbs and a variety of psychological process references. Women also used more negations and references to the home. Men scored higher on word length and the usage of numbers, articles and prepositions, and used more swear words than women. The language of women emphasized psychological processes, social processes and verbs, and was more likely to discuss people and internal processes. The language of men emphasized current concerns, and described more external events, objects and processes. Thelwall et al. (2010) examined gender differences in language use in 1000 MySpace comments. In this dataset, women sent and received comments with significantly more positive emotion than men. There was no significant gender difference found for negative emotion in comments, something which was found in research by Zhang et al. (2013). They examined gender differences in emotional expression on Web forums, using sentiment analysis. The results showed that in general, women were more likely to express positive and negative emotions compared to men, and that women were also more likely to express their opinions subjectively. Montero et al. (2014) investigated the role of emotion-based features in author gender classification in text. Female authors were found to use a wider range of emotion categories than male authors.

3. Current Research

This research intersects language, gender, trauma, oral history and artificial intelligence. It will examine the oral history archive provided by the Dutch Veteran Institute. This archive contains almost 1200 interviews in audio format, mostly of stories by Dutch veterans who have participated in wars, peace missions or other missions from 1940 onwards. The interviews provide emotional content as the interviewees talk about conflicting and traumatic events they have experienced. Analyzing language patterns in these interviews, specifically patterns in the use of words associated with emotion, will give information about the way in which the interviewees recall and process traumatic events. Finding emotional language patterns that are consistent with emotion gender norms would suggest that processing traumatic events is influenced by gender norms and that gender norms could provide a framework in which people process trauma. Hence, the research question is: *What gender differences regarding emotional expression can be found in the language used in the Dutch Veteran Institute oral history archive*?

The research question will be answered using a keyword based approach for emotion detection in text for several reasons. Firstly, there is only a small amount of female interviews in the collection, thus there is too little data available to sufficiently train and test a learning algorithm. Second, the keyword-based approach gives an overview of what word categories were used, opposed to the learning-based approach, which is mainly used for classification. Using computational methods to analyze language use in a Dutch oral history archive will take small steps towards filling the gap in research on automating Dutch oral history, and research on automatic (emotional) analysis of oral history archives.

Both the Dutch LIWC2007 lexicon and the LiLaH lexicon will be used for analysis. The Dutch LIWC2007 provides a wide range of categories and has been used earlier in research on gender differences in language. Its results will be compared to those of the LiLaH lexicon, a lexicon that does not have such an extensive range of categories, but which provides a few basic emotion categories. Using these lexicons, the research question will be further specified to: *Within which categories of the LIWC2007 and LiLaH lexicon are gender differences regarding emotional expression found in language used in the Dutch Veteran Institute oral history archive*?

Based on the literature discussed in the section Gender & Emotional expression, it is expected that women will express a wider range of emotions than men, express more positive emotion, and that men will express more anger or aggression (Briton & Hall, 1995; Timmers et al., 2003; Plant et al., 2000). This would also be consistent with earlier research named in the section Automatic Text Analysis & Gender, where women made more references to positive emotion and expressed a higher range of emotions (Mehl & Pennebaker, 2003; Thelwall et al., 2010; Zhang et al., 2013; Montero et al., 2014). Although men did not necessarily use more words associated with anger in previous research, they did use more swear words (Mehl & Pennebaker, 2003; Newman et al., 2008) which can be associated with anger or aggression. In terms of the Dutch LIWC2007 lexicon, it is hypothesized that women will use more words from the categories positive emotion and affective processes than men, and that men will use more words from the categories *anger* and *swear words* than women. In terms of the LiLaH lexicon, it is hypothesized that women will use more words from the categories *positive* and *joy* than men, and that men will use more words from the category *anger* than women.

4. Methodology

To see in what way emotional gender differences are present in the transcripts of the Dutch Veteran Institute oral history collection, and in which emotion categories these are present, interviews were selected and transcribed using an Automatic Speech Recognition (ASR) program. For each interview transcript, it was counted how many words from the transcript mapped to each emotion lexicon category from the Dutch lexicon of LIWC2007 and the LiLah Lexicon. This section describes the data, the ASR program and how the emotion lexicons were used.

#	Gender	Mission	Deployment	Time	Number of Words
209	F	WW2	Messenger, resistance	1943 - 1945	26,689
220	F	KM	Navy, Sergeant Major	1946 - 1949	11,666
357	F	UNPROFOR/UNPF	Land forces, therapist	1992 - 1993	31,902
427	F	UNPROFOR/UNPF	Air forces, F16 pilot	1992 - 1995	27,950
649	F	Miscellaneous	N/A, secretary	1948 - 1950	10,679
1284	F	KM	Navy, nurse	1947 - 1950	15,128
1297	F	KNIL	Nurse	1944 - 1949	5235
1713	F	UNPROFOR/UNPF	Land forces, soldier	1994, 1996	20,041
278	M	KM	Navy, shooter	1947 - 1949	26,573
339	M	UNPROFOR/UNPF	Navy, explorer	1993	18,169
540	M	KM	Land forces, commander	1947 - 1949	9064
581	M	KM	Navy, marine	1945	15,663
602	M	KM	Navy, marine	1948 - 1950	17,563
608	M	UNPROFOR/UNPF	Land forces, communication	1995	21,566
664	M	IFOR SFOR EUFOR	Land forces	1997	21,061
809	M	UNPROFOR/UNPF	Marechaussee, group leader	1992	22,287
1374	M	KM	Navy, medic	1945	9147

Table 4.1 - **Overview Interviews**
Interview number, gender, mission, position of deployment time of deployment and number of words in the transcript.

4.1 Data

The data consists of 8 interviews of female interviewees and 9 interviews of male interviewees, in audio format obtained from the online interview collection of the Dutch

Veteran Institute. In these interviews, an interviewer asks questions about the interviewee's parttaken mission and life. The missions discussed in the interviews all took place between 1943 and 1997. All interviewees are veterans, except #649, who was a secretary during World War II in Batavia, the capital of the Dutch East Indies. In table 4.1 is an overview of each interview and its interviewee, the gender of the interviewee, which mission is discussed in the interview, position of deployment, years of deployment and how many words the transcript of the interview contains.

4.2 Automatic Speech Recognition

Transcripts were made of the interviews by an Automatic Speech Recognition (ASR) program created by Veldema (2021). This ASR program was created using VosK API (2021) and the Medium Dutch Model from Kaldi NL (2021), both speech recognition toolkits. Words that could not be recognized during speech recognition were transcribed as '<unk>'. After transcription, all '<unk>' tokens were removed from the transcripts. This resulted in transcripts in .txt format of various lengths, the average being 18257.82 words, the shortest transcript consisting of 5235 words, and the longest consisting of 31,902 words.

4.3 Automatic Text Analysis

The processing of the lexicons and text analysis were both done in python. In order to analyze the text, both the Dutch LIWC2007 lexicon and the LiLaH lexicon were used. Since they had different formats, they were used slightly differently.

4.3.1 Dutch LIWC2007

The Dutch LIWC2007 lexicon consisted of a .dic file basically containing two lists, one list of numeric values with each a corresponding category name, and one list of keywords with corresponding numeric values. Per transcript each target word was analyzed. If a target word was present in the list of Dutch LIWC2007 keywords, or if the list of Dutch LIWC2007 keywords provided a stem of the target word ending by '*', the target word was added to a list that contained all occurrences of the corresponding keyword. Then, all numeric values that belonged to that keyword were searched, and linked to the correct category names. The words in the list of words of the transcript that matched with that keyword were added to a list under the corresponding category names. This resulted in a list of all categories, mapped to all target words in the text that corresponded with that category. This allowed the amount of target words that matched each category to be counted.

4.3.2 LiLaH

The LiLaH lexicon consisted of a .tsv file, containing a table with columns for the Croatian, Dutch and Slovenian translation of a keyword, together with columns for each category. Per word, the categories had a value of either 1 - meaning that the word belonged to this category, or 0 - meaning that it didn't. To use this lexicon, a list of only the Dutch keywords was created, with a corresponding binary vector to indicate the associated categories. Because of the way LiLaH deals with homonyms, the list of Dutch keywords was modified so that words with multiple meanings also only had one instance, and that their corresponding categories were all the categories that corresponded with the homonyms. The example 'zuivering' from the section Emotion Recognition in Text, which first had two instances with each different corresponding emotion categories (*positive* and *trust* - *negative* and *fear*) would now get one instance with the categories *positive, negative, fear* and *trust*. Per transcript each target word was analyzed. If a target word was present in the list of Dutch keywords, it was added to a list containing all occurrences in the transcript of that keyword. Then, for all present keywords in the transcript, the words in the list of all occurrences would be added to the categories that had a score of 1 for that keyword. This resulted in a list of all categories, mapped to all target words in the text that corresponded to that category. This allowed the amount of target words that matched each category to be counted.

4.4 Scaling

Because the length of the transcripts varied a lot and the sample size was small, the data was scaled. Per transcript, an estimate was calculated of the results if each transcript would have a length of 10,000 words. This was done using a scaling factor, calculated for each transcript by *scaling factor = (10,000 / total amount of words in the transcript)*. Using this scaling factor, the estimated amount of matching words per 10,000 words for each lexicon category was calculated for each transcript, by multiplying the result scores for each lexicon category by this scaling factor.

5. Results

To answer the research question, whether there are gender differences in emotion in the language used in the oral history archive of the Dutch Veteran Institute, a Welch two sample t-test was performed on the scaled results of the individual lexicon categories with gender as the independent variable. The results from the LiLaH emotion lexicon are displayed in table 5.1. The results from the Dutch LIWC2007 lexicon are partially displayed in table 5.2, the full results are displayed in table 9.3 in the appendix.

	Total		Women		Men					
Category	**M**	**SD**	**M**	**SD**	**M**	**SD**	**t**	**df**	**p**	**d**
Positive	530.36	50.18	542.58	51.10	519.51	49.70	0.9412	15	0.362	0.46
Negative	261.64	36.70	263.67	33.54	259.84	41.26	0.2109	15	0.836	0.10
Trust	211.41	19.66	218.84	20.49	204.81	17.36	1.5129	14	0.153	0.74
Fear	205.97	32.92	207.46	31.57	204.64	35.94	0.1721	15	0.866	0.08
Anticipation	192.64	17.66	194.37	17.47	191.11	18.74	0.3718	15	0.715	0.18
Sadness	191.31	18.38	195.71	16.98	187.39	19.66	0.9362	15	0.364	0.45
Anger	150.20	24.70	146.07	19.47	153.87	29.26	-0.6532	14	0.524	-0.31
Joy	114.64	17.77	128.08	12.28	102.70	12.59	4.2017	15	0.001	2.04
Surprise	84.35	11.10	82.03	9.98	86.42	12.20	-0.8150	15	0.428	-0.39
Disgust	77.35	13.39	75.49	13.95	79.01	13.49	-0.5267	15	0.606	-0.26

Table 5.1 - **Results LiLaH**

M = mean, SD = standard deviation, t = t-value,
df = degrees of freedom, p = p-value, d = Cohen's d-value.

When we take a significance level of 0.05, there is only a significant difference in the use of words of the LiLaH *joy* category, where female interviewees were the more frequent users compared to the male interviewees. For the LIWC2007 lexicon, there were no significant differences in categories that fell under *affective processes,* including *affective processes* itself. When it came to other categories, the use of words in the *humans* category from the Dutch LIWC2007 lexicon were significantly greater for female interviewees than for male interviewees, the use of words in the *friends* category was significantly greater for female interviewees than for male interviewees and the use of words in the *space* category were significantly greater for male interviewees than for female interviewees. Use of words in the *swear words* category were also significantly greater for male interviewees than for female interviewees.

	Total		Women		Men					
Category	**M**	**SD**	**M**	**SD**	**M**	**SD**	**t**	**df**	**p**	**d**
Humans	86.33	26.56	107.10	23.78	67.86	10.06	4.3348	9	0.002	2.20
Swear	2.27	1.66	1.14	0.69	3.27	1.63	-3.5688	11	0.004	-1.66
Space	733.19	54.65	700.21	38.45	762.50	51.25	-2.8532	15	0.012	-1.36
Friend	12.12	6.24	15.78	6.20	8.87	4.37	2.6230	12	0.022	1.30
Anger	40.25	13.16	33.95	12.04	45.85	12.02	-2.0348	15	0.060	-0.99
Future	54.57	13.18	48.39	11.54	60.06	12.59	-1.9926	15	0.065	-0.96
Positive emotions	171.79	24.03	182.03	26.05	162.68	19.04	1.7295	13	0.108	0.86
Affective processes	285.08	32.18	296.78	20.51	274.68	38.01	1.5134	13	0.155	0.71
Negative emotions	96.31	21.76	94.15	19.72	98.23	24.45	-0.3798	15	0.709	-0.18
Function words	6094.12	294.57	6125.62	388.99	6066.13	198.17	0.3899	10	0.705	0.20
Cognitive processes	1923.09	93.51	1919.21	96.80	1926.53	96.24	-0.1562	15	0.878	-0.08
Relativity	1522.27	100.71	1485.89	90.41	1554.61	103.11	-1.4642	15	0.164	-0.71
Common verbs	1471.84	103.85	1468.72	111.58	1474.62	103.23	-0.1127	14	0.912	-0.06
Pronouns	1238.62	172.45	1292.30	219.70	1190.90	108.63	1.1831	10	0.264	0.60
Social processes	829.00	110.14	866.79	107.14	795.41	107.33	1.3699	15	0.191	0.67
Inclusive	631.21	63.48	633.21	38.66	629.44	82.12	0.1233	12	0.904	0.06
Time	570.50	53.76	573.49	29.03	567.84	70.89	0.2195	11	0.830	0.10
Religion	12.84	7.31	14.71	9.93	11.18	3.75	0.9469	9	0.369	0.48
Friend	12.12	6.24	15.78	6.20	8.87	4.37	2.6230	12	0.022	1.30
Sexual	3.39	1.96	3.34	1.47	3.44	2.40	-0.1040	13	0.919	-0.05
Filler words	0	0	0	0	0	0	N/A	N/A	N/A	N/A

Table 5.2 - **Partial Results Dutch LIWC2007**
M = mean, SD = standard deviation, t = t-value,
df = degrees of freedom, p = p-value, d = Cohen's d-value.

Results that neared significance, were that for the Dutch LIWC2007 lexicon men used more words in the category *anger*, more words in the category *future* and that women used more words in the category *family* and the category *3rd person singular* ($p < .1$).

On average, the Dutch LIWC2007 word categories that were used the most were *function words, cognitive processes, relativity, common verbs* and *pronouns*. When categories from linguistic dimensions were excluded, the word categories that were used the most were *cognitive processes, social processes, space, inclusive* and *time*. The categories that were used the least were *filler* (which had zero instances), *swear words, sexual, friend,* and *religion*. The LiLaH word categories that were used the most on average were *positive* and *negative*, when the sentiment categories were excluded this resulted in the categories *trust* and *fear*. The least used categories were *disgust* and *surprise*.

6. Discussion

Comparing the language men and women use in this small sample of oral history transcripts shows a few significant gender differences in language use. The research question was: *within which categories of the LIWC2007 and LiLaH lexicon are gender differences regarding emotional expression found in language used in the Dutch Veteran Institute oral history archive*? It was hypothesized that women will use more words from the Dutch LIWC2007 categories *positive emotion* and *affective processes* than men, and that men will use more words from the categories *anger* and *swear words* than women. For the LiLaH lexicon, it was hypothesized that women will use more words from the categories *positive* and *joy* than men, and that men will use more words from the category *anger* than women.

The only significant difference found in emotion categories was that women used more words in the category *joy* from the LiLaH lexicon. This result is consistent with the hypothesis and falls in line with previous research suggesting that women feel more pressured to express positive emotions and that men are socialized to express less positive emotions. What is interesting is that for the Dutch LIWC2007 category *positive emotions* , no significant difference was found. This was the case as well for the LiLaH sentiment category *positive*, the LiLaH category that is used the most by both men and women. The category *joy* is almost a subcategory of the category *positive*. A possible explanation for the significant difference in the use of the category *joy* but not *positive*, is that when women describe positive experiences, they choose words that are more intense in expressing a positive emotion.

Something that could be expected based on the theory on gender and emotion but was not found in any previous research on text analysis and gender, is that men used significantly more words in the category *anger*. This was also not found in the results for both Dutch LIWC2007 & LiLaH. In Dutch LIWC2007, the gender difference of using words from the category *anger* was near the significance level, but for LiLaH this wasn't the case at all. This difference is due to the fact the lexicons contain different keywords with different labels, probably because they were created for different purposes. The NRC Word-Emotion Lexicon, of which LiLaH is the Dutch translation, was initially created for sentiment and emotion analysis and its keywords are often used as a feature in the learning-based approach (Mohammad, n.d.), while LIWC is a text analysis program that uses the keyword-based approach. That men did not use significantly more words from the category *anger* does not necessarily mean that they did not express anger during the interviews. In this research, only the transcripts of the interview are analyzed. Transcribing what is said during interviews results in a loss of emotional information, like tone of voice and body language, which could also indicate someone is expressing anger or aggression, or other emotions in general.

Something that was found that can possibly be related to expressing more anger or aggression is that men used significantly more words from the category *swear words* than women. This finding is consistent with previous research on automatic text analysis and gender where it was found that men also used more swear words than women. Using swear words is often associated with expressing frustration or aggression, which is consistent with gender stereotypes, as men are seen as more aggressive and women are discouraged to express themselves aggressively.

Women did not use significantly more words from the Dutch LIWC2007 category *affective processes* and thus did not seem to use more words related to emotions than men. A possible explanation could be that both genders expressed a lot of emotion, since the subjects discussed during the interviews were very personal.

Women used more words of the Dutch LIWC2007 *human* and *friends* categories compared to men, both being categories that fall under social processes and are used to describe other people. This is consistent with earlier research that women use more words from the category social processes and use language more often to describe other people and what they are doing (Newman et al, 2008).

Men used significantly more words from the Dutch LIWC2007 category *space*, a category which includes words to describe locations or to indicate where something is located in space. This was not found in earlier research and a likely explanation for this result in this sample is the deployment position of the male interviewees. All of them had positions working on the ground, with one exception. Their position on the ground, requiring them to move around a lot and having to visit and process different locations and landmarks, could be an explanation to why references to space were very present as opposed to the positions filled by the female interviewees, which include (but are not limited to) positions as a nurse, secretary and therapist - which are not as heavily focussed on agility, mobility and the environment.

6.1 Limitations and Improvements

The first and most obvious limitation of this research is that a very small sample size was used, as only a small sample of interviews with women were available. To obtain more reliable results, a bigger sample needs to be used, since the current results might be more dependent on the individual experiences of the interviewees describing their life and mission in the interviews. Also, the results have been scaled, and only an estimation of the amount of words in each category per 10,000 words is presented. When a bigger sample is used, scaling won't be necessary.

The biggest downside of the keyword-based approach is that it does not take context into account. A lot of important linguistic data is lost when only individual words are subject to categorization. When detecting the keywords in the transcripts, negations and intensifiers were not processed which also could result in an incorrect categorization, as well as the program not being able to detect any sarcasm, idioms, metaphors, and so on. The lexicons used were created using different datasets than the dataset used for this research, so the provided associations with the keywords could possibly not be relevant for the context of military life and war. Not only a lot of emotional data in the transcripts is lost using the chosen method, but also other emotional cues, like tone of voice, body language, silence and facial expressions ,are not present when analyzing the transcripts, whilst they could provide relevant information when researching gender differences in emotional expression.

6.2 Future Research

For further research on gender differences in emotion in text, an advanced sentiment analysis system should be created that also takes negations and intensifiers into account. Sentiment analysis makes use of a sentiment library, similar to an emotion lexicon, consisting of adjectives and phrases that have been manually given a suitable score. It evaluates the sentiment towards an entity in a body of text, based on the score of its neighbouring words or words that are in relation to that word (Lexalytics, n.d.). Instead of counting every single instance of keywords, this system could return the sentiment or emotion per sentence so that it takes more context into account when analyzing text.

One factor that hasn't been taken into account in this work is the age of the interviewees. It would be interesting to examine the effect age has on gender differences in language use, as it could give clues on how not only language but also gender norms change over time.

7. Conclusion

This work examined whether there were any gender differences present regarding emotional expression in language used in the Dutch Veteran Institute oral history archive by using an application of artificial intelligence. It was expected that women would express more emotions - positive emotions especially - and that men would express more anger and swear more. The hypothesis that women would use more words from the categories *positive emotion*, *affective processes* (Dutch LIWC2007), *positive* and *joy* (LiLaH) than men, and that men would use more words from the categories *anger* (Dutch LIWC 2007 & LiLaH) and *swear words* (Dutch LIWC2007) than women, was not proven to be true. Analyzing transcripts of the Dutch Veteran Institute interviews using the Dutch LIWC2007 and LiLaH emotion lexicon showed that women used more words associated with joy and that women used language to describe other people more often, whereas men used more swear words and more often words to indicate or describe where something took place or where something is located in space. Although the full hypothesis is not proven to be true, the results show that there are gender differences present in language of the data set, as well as that gender stereotypes are slightly present in the language. When the results are viewed in the context of trauma and processing trauma, this research does not provide much evidence on the way people actually process trauma. Although the choice of words in the interviews say something about how an individual experienced or processed an event, the interviews being largely about conflicting or traumatic events, it is not a given that the interviewees were still processing trauma at the moment the interviews took place. Furthermore, a single interview is not a complete representation of how an individual processes trauma or any event in general. What the current results can tell is that for this dataset, emotional gender stereotypes are slightly present in language when recalling traumatic events. This research presented an approach to analyzing emotion in oral history archives which could be applied on a larger scale. Analyzing human emotions using computational methods is a complex matter, and something that is not easily done using a keyword-based approach. Emotional expression can not be fully captured by text, let alone individual words.

8. Bibliography

Acheampong, F. A., Wenyu, C., & Nunoo-Mensah, H. (2020). *Text-based emotion detection: Advances, challenges, and opportunities.* Engineering Reports, 2(7), e12189.

Algoe, S. B., Buswell, B. N., & DeLamater, J. D. (2000). *Gender and job status as contextual cues for the interpretation of facial expression of emotion.* Sex Roles, 42, 183–208.

Alpha Cephei Ink. (2021). VOSK API. https://github.com/alphacep/vosk-api

Alswaidan, N., & Menai, M. E. B. (2020). *A survey of state-of-the-art approaches for emotion recognition in text.* Knowledge & Information Systems, 62(8).

Boot, P., Zijlstra, H., & Geenen, R. (2017). *The Dutch translation of the linguistic inquiry and word count (LIWC) 2007 dictionary*. Dutch Journal of Applied Linguistics, 6(1), 65-76.

Brescoll, V. L., & Uhlmann, E. L. (2008). *Can an angry woman get ahead? Status conferral, gender, and expression of emotion in the workplace.* Psychological Science, 9, 268–275.

Briton, N. J., & Hall, J. A. (1995). *Beliefs about female and male nonverbal communication*. Sex Roles, 32, 79–90. Brody, L. R. (1997). Beyond stereotypes: Gender and emotion. Journal of Social Issues, 53, 369–393.

Brody, L. R., & Hall, J. A. (2010). *Gender, emotion, and socialization*. In: Handbook of gender research in psychology(pp. 429-454). Springer, New York, NY.

Byrne, W., Doermann, D., Franz, M., Gustman, S., Hajic, J., Oard, D. & Zhu, W. J. (2004). *Automatic recognition of spontaneous speech for access to multilingual oral history archives.* IEEE Transactions on Speech and Audio Processing, 12(4), 420-435.

Chrisler, J. C. (2008). *Fear of losing control: Power, perfectionism, and the psychology of women.* Psychology of Women Quarterly, 32, 1–12.

Darves-Bornoz, J. M., Choquet, M., Ledoux, S., Gasquet, I., & Manfredi, R. (1998). *Gender differences in symptoms of adolescents reporting sexual assault*. Social Psychiatry and Psychiatric Epidemiology, 33(3), 111-117.

Green, B. L., Kramer, T. L., Grace, M. C., Gleser, G. C., Leonard, A. C., Vary, M. G., & Lindy, J. D. (1997). *Traumatic events over the life span: Survivors of the Buffalo Creek disaster.*

Gunes, H., & Pantic, M. (2010). *Automatic, dimensional and continuous emotion recognition.* International Journal of Synthetic Emotions (IJSE), 1(1), 68-99.

Gustman, S., Soergel, D., Oard, D., Byrne, W., Picheny, M., Ramabhadran, B., & Greenberg, D. (2002). *Supporting access to large digital oral history archives.* In Proceedings of the 2nd ACM/IEEE-CS joint conference on Digital libraries (pp. 18-27).

Hess, U., Senecal, S., Kirouac, G., Herrera, P., Philippot, P., & Kleck, R. E. (2000). *Emotional expressivity in men and women: Stereotypes and self-perceptions*. Cognition & Emotion, 14, 609–642.

Hourani, L. L., Williams, J., Bray, R. M., & Kandel, D. B. (2014). *Posttraumatic stress disorder, substance abuse, and other behavioral health indicators among active duty military men and women*. J Trauma Stress Disord Treat, 3(3), 2.

Hourani, L. L., Williams, J., Bray, R. M., Wilk, J. E., & Hoge, C. W. (2016). *Gender differences in posttraumatic stress disorder and help seeking in the US Army*. Journal of women's health, 25(1), 22-31.

Huijbregts, M., Ordelman, R., & de Jong, F. M. (2005). *A spoken document retrieval application in the oral history domain.* In: Proceedings of 10th international conference Speech and Computer, Patras, Greece (SPECOM 2005) (Vol. 2, pp. 699-702). University of Patras/WCL Moscow State Linguistics Uni.

Irish, L. A., Fischer, B., Fallon, W., Spoonster, E., Sledjeski, E. M., & Delahanty, D. L. (2011). *Gender differences in PTSD symptoms: an exploration of peritraumatic mechanisms*. Journal of Anxiety Disorders, 25(2), 209-216.

Kaldi NL. (2021) https://github.com/opensource-spraakherkenning-nl/Kaldi_NL

Kessler, R. C., Sonnega, A., Bromet, E., Hughes, M., & Nelson, C. B. (1995). *Posttraumatic stress disorder in the National Comorbidity Survey*. Archives of general psychiatry, 52(12), 1048-1060.

King, M. W., Street, A. E., Gradus, J. L., Vogt, D. S., & Resick, P. A. (2013). *Gender differences in posttraumatic stress symptoms among OEF/OIF veterans: An item response theory analysis*. Journal of Traumatic Stress, 26(2), 175-183.

Lexalytics (n.d.) *Sentiment Analysis Explained*. https://www.lexalytics.com/technology/sentiment-analysis

Livingston, R., Lawson, L., & Jones, J. G. (1993). *Predictors of self-reported psychopathology in children abused repeatedly by a parent*. Journal of the American Academy of Child & Adolescent Psychiatry, 32(5), 948-953.

Ljubešić, N., Markov, I., Fišer, D., & Daelemans, W. (2020). *The LiLaH emotion lexicon of Croatian, Dutch and Slovene.* In: Proceedings of the Third Workshop on Computational Modeling of People's Opinions, Personality, and Emotions in Social Media (pp. 153-157).

Maguen, S., Ren, L., Bosch, J. O., Marmar, C. R., & Seal, K. H. (2010). *Gender differences in mental health diagnoses among Iraq and Afghanistan veterans enrolled in veterans affairs health care*. American journal of public health, 100(12), 2450-2456.

Mehl, M. R., & Pennebaker, J. W. (2003). *The sounds of social life: a psychometric analysis of students' daily social environments and natural conversations.* Journal of personality and social psychology, 84(4), 857.

Mohammad, S. M. (n.d.) *NRC Word-Emotion Association Lexicon (aka EmoLex).* https://saifmohammad.com/WebPages/NRC-Emotion-Lexicon.htm

Mohammad, S. M., & Turney, P. D. (2013). *Nrc emotion lexicon*. National Research Council, Canada, 2.

Mohammad, S. M., & Turney, P. D. (2010). *Emotions evoked by common words and phrases: Using mechanical turk to create an emotion lexicon.* In Proceedings of the NAACL HLT 2010 workshop on computational approaches to analysis and generation of emotion in text (pp. 26-34).

Montero, C. S., Munezero, M., & Kakkonen, T. (2014). *Investigating the role of emotion-based features in author gender classification of text*. In: International Conference on Intelligent Text Processing and Computational Linguistics (pp. 98-114). Springer, Berlin, Heidelberg.

Nederlands Veteranen Instituut. (2021). *Interviewcollectie Nederlandse Veteranen.* https://www.veteranenvertellen.nl

Newman, M. L., Groom, C. J., Handelman, L. D., & Pennebaker, J. W. (2008). *Gender differences in language use: An analysis of 14,000 text samples*. *Discourse processes*, *45*(3), 211-236.

Olsson, J. S., & Oard, D. W. (2007). *Improving text classification for oral history archives with temporal domain knowledge.* In: Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval (pp. 623-630).

Ordelman, R., de Jong, F., Huijbregts, M., & van Leeuwen, D. (2005). *Robust audio indexing for Dutch spoken word collections.* Humanities, Computers and Cultural Heritage, 215.

Osherson, S., & Krugman, S. (1990). *Men, shame, and psychotherapy.* Psychotherapy, 27, 327–339.

Parmley, M., & Cunningham, J. G. (2008). *Children's gender-emotion stereotypes in the relationship of anger to sadness and fear.* Sex Roles, 58, 358–370.

Pennebaker, J. W. (1993). *Putting stress into words: Health, linguistic, and therapeutic implications.* Behaviour research and therapy, 31(6), 539-548.

Pennebaker, J. W., Chung, C. K., Ireland, M. E., Gonzales, A., & Booth, R. J. (2007). *The development and psychometric properties of LIWC2007.* University of Texas at Austin and The University of Auckland, New Zealand. Development, 1(2) 1-22.

Pennebaker, J. W., Francis, M. E., & Booth, R. J. (2001). *Linguistic inquiry and word count: LIWC 2001.* Mahway: Lawrence Erlbaum Associates, 71, 2001.

Pennebaker, J. W., Mehl, M. R., & Niederhoffer, K. G. (2003). *Psychological aspects of natural language use: Our words, our selves.* Annual review of psychology, 54(1), 547-577.

Pessanha, F. & Akdag Salah, A. (2021) *A Computational Look at Oral History Archives.* ACM J. Comput. Cult. Herit. 1, 1 (february 2021), 16 pages.

Plant, E. A., Hyde, J. S., Keltner, D., & Devine, P. G. (2000). *The gender stereotyping of emotions.* Psychology of Women Quarterly, 24, 81–92.

Plutchik, R. (1980). *A general psychoevolutionary theory of emotion*. In: Theories of emotion (pp. 3-33). Academic press.

Richardson, L. K., Frueh, B. C. & Acierno, R. (2010). *Prevalence estimates of combat-related post-traumatic stress disorder: Critical review.* Australian & New Zealand Journal of Psychiatry, 44, 4-19.

RIVM. (2021). *Posttraumatische stressstoornis: cijfers & context.* https://www.volksgezondheidenzorg.info/onderwerp/posttraumatische-stressstoornis/cijfers-context/huidige-situatie. RIVM: Bilthoven, 2021.

Simpson, P. A., & Stroh, L. K. (2004). *Gender differences: Emotional expression and feelings of personal inauthenticity.* Journal of Applied Psychology, 89, 715–721.

Stratou, G., Scherer, S., Gratch, J., & Morency, L. P. (2013, September). *Automatic nonverbal behavior indicators of depression and PTSD: Exploring gender differences.* In: 2013 Humaine Association Conference on Affective Computing and Intelligent Interaction (pp. 147-152). IEEE.

Strauss, J., Muday, T., McNall, K., & Wong, M. (1997). *Response style theory revisited: Gender differences and stereotypes in rumination and distraction.* Sex Roles, 36, 771–792.

Substance Abuse and Mental Health Services Administration. (2014). *SAMHSA's Concept of Trauma and Guidance for a Trauma-Informed Approach*. HHS Publication No. (SMA) 14-4884.

Substance Abuse and Mental Health Services Administration. (2014). *Understanding the Impact of Trauma*. In: Trauma-Informed Care In Behavioral Health Services. Treatment Improvement Protocol (TIP) Series, No. 57. Rockville, MD. https://www.ncbi.nlm.nih.gov/books/NBK207191/

Survivors of the Shoah Visual History Foundation (VHF). http://www.vhf.org

Tausczik, Y. R., & Pennebaker, J. W. (2010). *The psychological meaning of words: LIWC and computerized text analysis methods.* Journal of language and social psychology, 29(1), 24-54.

Thelwall, M., Wilkinson, D., & Uppal, S. (2010). *Data mining emotion in social network communication: Gender differences in MySpace*. Journal of the American Society for Information Science and Technology, 61(1), 190-199.

Timmers, M., Fischer, A. H., & Manstead, A. S. R. (2003). *Ability versus vulnerability: Beliefs about men's and women's emotional behavior*. Cognition & Emotion, 17, 41–63.

Tolin, D. F., & Foa, E. B. (2006). *Sex differences in trauma and posttraumatic stress disorder: a quantitative review of 25 years of research.*

Van Wissen, L., & Boot, P. (2017). *An electronic translation of the LIWC Dictionary into Dutch*. In: Electronic lexicography in the 21st century: Proceedings of eLex 2017 conference (pp. 703-715). Lexical Computing.

Veldema, M. (2021). *PTSD Recognition using Computational Paralinguistic Analysis.*

Veteraneninstituut. (2015). *Posttraumatische stressstoornis na uitzending.* https://www.nlveteraneninstituut.nl/content/uploads/2015/02/FS6-Posttraumatische-stressstoornis-na-uitzending1.pdf

Vingerhoets, A. J., Bylsma, L. M., & De Vlam, C. (2013). *Swearing: A biopsychosocial perspective*. Psihologijske teme, 22(2), 287-304.

Zhang, Y., Dang, Y., & Chen, H. (2013). *Examining gender emotional differences in Web forum communication*. Decision Support Systems, 55(3), 851-860.

9. Appendix

Table 9.1 - **Dutch LIWC2007 Emotion Lexicon**

Categories, abbreviations of categories and examples of Dutch words belonging to listed categories.

Category	**Abbreviation**	**Examples**
Linguistic dimensions		
Function words	funct	*Aan, al, boven*
Pronouns	pronoun	*Zij, hen, haar*
Personal pronouns	ppron	*Ik, jij, onze*
1st pers singular	i	*Ik, mij, mijn*
1st pers plural	we	*Wij, onze, ons*
2nd person	you	*Jij, jullie, jouw*
3rd pers singular	shehe	*Zij, hij*
3rd pers plural	they	*Zij, hen*
Impersonal pronouns	ipron	*Allemaal, alles, sommig**
Articles	article	*De, het, een*
Common verbs	verb	*Aaien, scheer, zwom*
Auxiliary verbs	auxverb	*Behoort, bent*
Past tense	past	*Deelde, discussieërde, domineerde*
Present tense	present	*Afleren, bewaren, gebeuren*
Future tense	future	*Zal, wens, wil*
Adverbs	adverb	*Absoluut, heel, ogenblikkelijk*
Prepositions	preps	*Om, op, sinds*
Conjunctions	conj	*Hoewel, inzover*, naarmate*
Negations	negate	*Zonder, evenmin, buiten*
Quantifiers	quant	*Omtrek, item*, bosje*
Numbers	number	*Één, dertig, miljoen*
Swear words	swear	*Opsodemieter*, pik, rotvent*
Psychological Processes		
Social processes	social	*Communiceren, delen, helpen*
Family	family	*Moeder, ouders, neef*

Friends	friends	*Vriend, vriendschap, vriendin*
Humans	humans	*Meisje, mens, volwassenen*
Affective processes	affect	*Blij, verdrietig, somber*
Positive emotion	posemo	*Gelukkig, dankbaar*
Negative emotion	negemo	*Bedroefd, vijandig, wanhoop*
Anxiety	anx	*Zenuwachtig, bang, gespannen*
Anger	anger	*Boos, dreigen, ergeren*
Sadness	sad	*Huilen, somber, teleurstelling*
Cognitive processes	cogmech	*Oorzaak, weten, denken*
Insight	insight	*Overwegen, ophelderen, realiseren*
Causation	cause	*Omdat, waarom, vandaar*
Discrepancy	discrep	*Zou, behoren, conflict*
Tentative	tentat	*Misschien, waarschijnlijk, voorlopig*
Certainty	certain	*Volstrekt, absoluut, vastbesloten*
Inhibition	inhib	*Blokkeren, hinderen, inhouden*
Inclusive	incl	*Omvat*, optelling, toegevoegd**
Exclusive	excl	*Onderscheid*, buitengesloten, tenzij*
Perceptual processes	percept	*Zien, voelen, horen*
See	see	*Kijken, zicht, blik*
Hear	hear	*Geluid, luisteren, klank*
Feel	feel	*Aanraken, aftasten, vasthouden*
Biological processes	bio	*Ogen, proeven, uitput**
Body	body	*Oksel*, naakt*, tiet**
Health	health	*Cyst*, gebraakt, gynaecol**
Sexual	sexual	*Flirten, zoen, beminnen*
Ingestion	ingest	*Drinken, honger, voeding*
Relativity	relativ	*Ogenblik, oud, vlakbij*
Motion	motion	*Kar, omgegooid, liepen*
Space	space	*Vlakbij, plaats, noord*
Time	time	*Zomer, vroeger, zodra*

Personal Concerns		
Work	work	*Carrière, collega, concurreren*
Achievement	achieve	*Behalen, bekroning, beloning*
Leisure	leisure	*Fietsen, fitness, trainen*
Home	home	*Keuken, huis, tuin*
Money	money	*Winst, factuur, failliet*
Religion	relig	*Bidden, vereren, zegen*
Death	death	*Dood, treuren, sterfbed*
Spoken categories		
Assent	assent	*Oke, prima, haha*
Nonfluencies	nonfl	*Pff, uh*, zzz**
Fillers	filler	*Rr*, nouja, blabla*
Dutch lexicon only		
Pronominal adverbs	pronadv	*Omheen, tegenaan, vantussen*
Single and plural third person pronouns	shehethey	*Z'n, ze, haarzelf*

Table 9.2 - **LiLaH Emotion Lexicon**

Categories and examples of Dutch words belonging to listed categories.

Category	**Examples**
Positive	*Mogelijkheid, absoluut, prestatie*
Negative	*Gewapend, opscheppen, medeplichtigheid*
Anger	*Onstuimig, heersen, afdwingen*
Anticipation	*Traktatie, voorzichtigheid, samenzweerder*
Disgust	*Verdorven, duivel, epidemie*
Fear	*Veroordeling, dakloos, grim*
Joy	*Opvoeden, krachtig, romantiek*
Sadness	*Uitbarsting, herdenken, priesterschap*
Surprise	*Bestemming, musical, terugdeinzen*
Trust	*Winnend, zusterschap, onderdompelen*

Table 9.3 - **Results of Dutch LIWC2007**

M = mean, SD = standard deviation, t = t-value, df = degrees of freedom, p = p-value, d = Cohen's d-value.

	Total		Women		Men					
Category	**M**	**SD**	**M**	**SD**	**M**	**SD**	**t**	**df**	**p**	**d**
humans	86.33	26.56	107.10	23.78	67.86	10.06	4.3348	9	0.002	2.20
swear	2.27	1.66	1.14	0.69	3.27	1.63	-3.5688	11	0.004	-1.66
space	733.19	54.65	700.21	38.45	762.50	51.25	-2.8532	15	0.012	-1.36
friend	12.12	6.24	15.78	6.20	8.87	4.37	2.6230	12	0.022	1.30
anger	40.25	13.16	33.95	12.04	45.85	12.02	-2.0348	15	0.060	-0.99
future	54.57	13.18	48.39	11.54	60.06	12.59	-1.9926	15	0.065	-0.96
family	48.83	29.19	62.77	34.12	36.44	17.85	1.9577	10	0.078	0.99
shehe	68.91	15.18	75.69	13.56	62.88	14.59	1.8759	15	0.080	0.91
posemo	171.79	24.03	182.03	26.05	162.68	19.04	1.7295	13	0.108	0.86
see	65.95	22.43	56.80	14.07	74.07	25.98	-1.7292	13	0.108	-0.81
negate	207.64	43.11	224.61	50.72	192.56	30.35	1.5563	11	0.147	0.78
affect	285.08	32.18	296.78	20.51	274.68	38.01	1.5134	13	0.155	0.71
article	668.62	43.24	651.45	57.69	683.87	16.43	-1.5351	8	0.163	-0.79
relativ	1522.27	100.71	1485.89	90.41	1554.61	103.11	-1.4642	15	0.164	-0.71
leisure	56.06	14.97	61.42	12.49	51.30	16.05	1.4599	15	0.165	0.70
pronadv	56.22	16.73	50.47	9.20	61.33	20.58	-1.4294	11	0.180	-0.67
social	829.00	110.14	866.79	107.14	795.41	107.33	1.3699	15	0.191	0.67
nonfl	310.43	118.02	271.17	62.48	345.32	146.74	-1.3816	11	0.194	-0.64
i	301.89	70.05	325.14	71.56	281.22	65.67	1.3127	14	0.210	0.64
quant	366.00	60.69	386.04	73.92	348.19	42.75	1.2716	11	0.230	0.64
health	23.50	15.44	28.41	16.94	19.14	13.42	1.2394	13	0.237	0.61
tentat	295.12	49.20	280.38	14.91	308.23	65.08	-1.2477	9	0.244	-0.57
pronoun	1238.62	172.45	1292.30	219.70	1190.90	108.63	1.1831	10	0.264	0.60
preps	889.30	71.20	867.80	82.76	908.41	57.21	-1.1627	12	0.267	-0.58
ipron	695.39	107.30	726.82	110.11	667.46	102.72	1.1450	14	0.271	0.56
shehethey	139.93	34.50	150.24	39.97	130.76	27.97	1.1504	12	0.272	0.57
percept	218.14	44.70	205.70	20.11	229.19	57.87	-1.1427	10	0.280	-0.53
anx	16.65	5.00	18.08	6.14	15.37	3.61	1.0945	11	0.297	0.55
discrep	167.33	27.31	160.03	18.72	173.81	32.93	-1.0747	13	0.302	-0.51
past	604.21	128.86	639.16	137.41	573.15	119.92	1.0493	14	0.312	0.51
present	716.66	136.65	680.78	119.51	748.55	149.73	-1.0363	15	0.317	-0.50
insight	162.18	27.21	169.35	32.11	155.80	21.94	1.0037	12	0.335	0.50
motion	285.74	47.94	273.88	31.39	296.28	58.89	-0.9933	12	0.339	-0.47
relig	12.84	7.31	14.71	9.93	11.18	3.75	0.9469	9	0.369	0.48

bio	78.83	26.31	84.85	26.88	73.47	26.15	0.8824	15	0.392	0.43
certain	298.62	57.54	311.03	44.49	287.58	67.82	0.8514	14	0.409	0.40
hear	80.37	19.39	76.27	18.16	84.02	20.78	-0.8201	15	0.425	-0.40
ppron	772.85	113.90	797.24	138.79	751.18	89.26	0.8026	12	0.438	0.40
adverb	748.08	96.04	766.72	96.56	731.50	98.14	0.7449	15	0.468	0.36
achieve	80.65	13.97	83.44	16.54	78.18	11.67	0.7484	12	0.468	0.37
ingest	27.37	11.32	29.48	11.57	25.49	11.43	0.7147	15	0.486	0.35
conj	732.54	67.90	720.94	45.30	742.85	84.69	-0.6749	12	0.512	-0.32
we	117.52	40.26	110.57	42.57	123.70	39.54	-0.6562	14	0.522	-0.32
they	86.68	27.04	91.27	30.75	82.61	24.39	0.6380	13	0.534	0.31
death	15.23	8.16	16.55	9.44	14.05	7.22	0.6069	13	0.554	0.30
home	34.80	14.34	36.88	17.64	32.95	11.44	0.5379	12	0.601	0.27
sad	27.07	8.11	28.19	8.75	26.07	7.88	0.5228	14	0.609	0.26
auxverb	838.18	77.83	848.52	74.21	828.98	84.23	0.5084	15	0.619	0.25
funct	6094.12	294.57	6125.62	388.99	6066.13	198.17	0.3899	10	0.705	0.20
negemo	96.31	21.76	94.15	19.72	98.23	24.45	-0.3798	15	0.709	-0.18
excl	406.21	57.08	400.81	57.12	411.01	60.06	-0.3585	15	0.725	-0.17
body	23.67	7.89	23.04	5.37	24.23	9.92	-0.3108	13	0.761	-0.15
work	110.36	27.07	112.56	32.80	108.40	22.69	0.2999	12	0.769	0.15
assent	179.78	56.63	183.93	54.75	176.09	61.30	0.2783	15	0.785	0.13
time	570.50	53.76	573.49	29.03	567.84	70.89	0.2195	11	0.830	0.10
money	24.37	8.07	24.83	9.23	23.96	7.43	0.2104	13	0.837	0.10
cogmech	1923.09	93.51	1919.21	96.80	1926.53	96.24	-0.1562	15	0.878	-0.08
feel	41.67	9.80	41.31	7.61	41.99	11.89	-0.1425	14	0.889	-0.07
you	199.49	62.53	197.19	53.98	201.53	72.53	-0.1409	15	0.890	-0.07
cause	150.23	46.92	151.78	41.61	148.86	53.69	0.1263	15	0.901	0.06
incl	631.21	63.48	633.21	38.66	629.44	82.12	0.1233	12	0.904	0.06
inhib	20.52	5.85	20.32	7.82	20.69	3.83	-0.1219	10	0.905	-0.06
verb	1471.84	103.85	1468.72	111.58	1474.62	103.23	-0.1127	14	0.912	-0.06
sexual	3.39	1.96	3.34	1.47	3.44	2.40	-0.1040	13	0.919	-0.05
number	170.47	42.11	171.14	51.17	169.87	35.42	0.0587	12	0.954	0.03
filler	0	0	0	0	0	0	N/A	N/A	N/A	N/A