A Word-use Model for Depressed Thai Teenagers

Received : Nov 2, 2021
Revised : Jan 27, 2022
Accepted : Apr 7, 2022

Phanintra Teeranon
School of Liberal Arts
University of Phayao

Abstract

This study aims to introduce a word-use model for depressed Thai teenagers aged between 19–24 years old. Ten depressed Thai teenagers were asked to give a speech about their family life. Then, the words present in their speech were assigned to 11 word categories based on Linguistic Inquiry and Word Count (LIWC). Exploratory factor analysis (EFA) was utilized to determine the word categories. It was found that the social processes category was the most used by the depressed Thai teenagers. The word-use model suggested that three groups of words were used by the depressed Thai teenagers: "Emotion words," "Self-focus," and "Driving force." The results confirm Beck's cognitive model in terms of verbalizing thoughts as the signal for depression. It is suggested that most of the word categories used by depressed individuals tend to signal emotive meaning. In addition, driving force word-use category also indicates help-seeking of the depressed Thai teenagers.

Keywords: Thai, Depressed, Word

1. Introduction

Word categories serve as the foundation for semantic categories; words appear in certain locations in sentences and can thus be categorized syntactically (Rauh, 2016). The study of words can lead to the scientific conclusion that human language is made up of mental mechanisms (Pinker, 1998). In word studies, especially those on word categories, Chen and Jin (2017), Jarrold et al. (2011), and Zimmermann et al. (2017), among others, found that words can convey cognitive processes and two types of meaning: emotive meaning and conceptual meaning (Angkapanichkit et al., 2019). Emotive meaning signals the feelings or attitudes of a person, while conceptual meaning is mostly a basic propositional meaning. Based on these findings, not only negative perspectives are expressed by depressed individuals, as noted in Beck's cognitive model (Beck & Clark, 1997; Beck & Haigh, 2014), but also other conceptual word meanings are plausibly ventilated by the depressed ones.

Ramirez-Esparza et al. (2009), Wolohan et al. (2018), and Yoo and Ha (2019) are examples of studies that found that negative words are consistently used by depressed individuals in online platforms. In Thailand, where the prevalence of depression in Thai teenagers was found to be 14.9% (Chaveepojnkamjorn et al., 2017), Angkapanichkit et al. (2019) and Sathientharadol (2020) reported comparable results regarding the speech of Thai depressed teenagers. It can be inferred that these results confirm Beck's cognitive model.

Language, as a means of communication, can express the insights, feelings, and thoughts of depressed speakers (Evans, 2012). Currently, medical research and psychological research have undoubtedly accepted language as one indicator that can be used for the diagnosis of depression (O'Dea et al., 2021; Zapata-Vega et al., 2010).

The behavioural measure of self-focused attention correlates to Pyszczynski and Greenberg's (1987) model according to which the use of first-person singular pronouns, such as "I," "me," and myself," in spoken language is considered to be a well-established as a sign of depression (Rude et al., 2004). Edwards and Holtzman (2017) performed a meta-analysis of correlations between depression and first-person singular pronoun use and reported that "I-self-focus" was significantly related to depression. Later, Angkapanichkit et al. (2019) found that not only 'I-self-focus' but also first-person plural pronouns, such as "we," were often used by depressed Thai teenagers. In the context of inclusive storytelling, depressed Thai teenagers tend to include themselves in the story, which is an additional type of self-focused communication.

Recently, multiple studies have empirically extended Beck's cognitive model and Pyszczynski and Greenberg's model to include more word categories to signal various aspects emotional and mental health; one example is the Linguistic Inquiry and Word Count (LIWC) program by Francis and Pennebaker (1993). LIWC, as a linguistics installation, provides affirmed word-use categories associated with psychological thought. The word categories in LIWC have been widely employed by numerous studies (Lieberman & Goldstein, 2006; Ramirez-Esparza et al., 2009; Wolohan et al., 2018; Yoo & Ha, 2019) to examine more categories of language patterns used by depressed individuals.

Previous studies have indicated two main categories of words expressed by depressed individuals: linguistic dimension and psychological processes. The linguistic dimension includes function words, such as pronouns, prepositions, questions, and numbers; while psychological processes include affective processes, social processes, cognitive processes, perceptual processes, biological processes, drives, time orientations, relativity, personal concerns, and informal language.

Using LIWC and corpus-based analysis, Wolohan et al. (2018) reported that the following were highly expressed among depressed individuals: affective processes signalling negative emotion words, social processes signalling family and friend relationships, and time orientations telling the past and the future. However, using an online data set posted by depressed individuals, Yoo and Ha (2019) suggested that informal words, such as slang and abbreviations, were used as much as storytelling. The results of Yoo and Ha (2018) are consistent with the research by Angkapanichkit et al. (2019), in which storytelling was also reported to be a linguistic feature of the speech of depressed Thai teenagers, such as "I don't have much support" or "I feel like I can't breathe." In consequence, previous research findings on the Thai language has generated results that are comparable to studies in other languages. De Choudhury et al. (2013) extended classifiers based on the LIWC categories to find words and language patterns not included in the LIWC categories that would help identify people with mental illnesses.

In Thai depression research, Angkapanichkit et al. (2019) tentatively summarized approximately six word categories used by depressed Thai teenagers (collected from diary writings and interviews): *deixis* (e.g., I, father, mother), *cognitive processes* (e.g., think,

disappear, decide), perception (e.g., hearing, interested in), *affective processes* (e.g., happy, rich, lovely, bad, terrible, tired, angry), *negation processes*, *metaphor* (e.g., left behind, trash, garbage, sponge), and *action word* (e.g., forget, wake up). Moreover, in the study, the use of a combination of words among these six word categories and the use of intensifiers, for example, very+disappointed, were reported to be a sign of early depression in Thai teenagers. Angkapanichkit et al. (2019) also primarily reported pragmatic strategies used in depressed participants. Another study of depressed Thai teenagers'utterances was that by Sathientharadol (2020). The study reported four categories of word use in the participants: self, disease, society, and hope. Hence, it can be inferred that the word categories used in the speech of depressed Thai individuals are not clear, as few studies have addressed this issue. A fundamental finding was that studying word categories among depressed individuals yielded 72% accuracy in the early detection of depression (De Choudhury et al., 2013). Moreover, linguistic issues appear to be more important in medical and psychological studies than in other studies; the results shed some light on the development of cognitive models and indicators of depression in the fields of psychology and medicine to save lives. To posit reliable word categories used by depressed Thai individuals, more linguistic studies on this issue are necessary to establish linguistic patterns associated with depression.

To answer the research questions of which LIWC word categories highly relate to depressed Thai teenagers' cognition and the extent to which LIWC word categories represent depressed Thai teenagers' cognition, exploratory factor analysis (EFA) is applied for quantitative analysis. There are a small number of word categories found in previous studies on language among depressed Thai individuals, and factor analysis is a multivariate technique (Fávero & Belfiore, 2019) suitable for identifying a comparatively small number of word categories that represent the cognition of depressed Thai teenagers. Therefore, this study aims to establish realistic linguistic patterns or word-use model associated with depressed Thai teenagers using LIWC and to quantify word categories using EFA.

2. Literature review

Few linguistic studies have addressed utterances among depressed Thai individuals to gain insights into depression; hence, to posit common ground of previous research, both the studies of depressed Thai contexts and non-depressed Thai contexts were reviewed and synthesized. In the context of Thai which is the main focus, Teeranon (2020) initially analyzed phonetic features among depressed Thai teenagers. Angkapanichkit et al. (2019) and Sathientharadol (2020) investigated utterances used by depressed Thai teenagers. Consequently, this research divides language use by depressed Thai teenagers into two linguistic levels: the phonetic and utterance levels.

2.1 Phonetic level

Depression was acoustically assessed using phonetic features, including pauses, loudness, pitch, intonation, and stress. Vicsi et al. (2012) reported the pause and speech rates in the speech of depressed individuals. This point was later confirmed by Yang et al. (2013) in a comparative study of the vocal characteristics of depressive and non-depressed adults. It was reported that pauses and loud speech were more likely to be used by depressed patients than non-depressed patients. The results of Vicsi et al. (2012) and Yang et al. (2013) were in line with those of Greden and Carroll (1980) and Greden et al. (1981), who found that frequent pauses in speech could indicate depression. Likewise, Vogel et al. (2011) found that people

with depression spoke more slowly than non-depressed people. However, Williamson et al. (2016) recently found that loudness was not an indicator of depression.

Some studies have found that in addition to pauses and loudness of speech, pitch is another clear indicator of depression. Breznitz (1992) and Nilsome (1998) found different fundamental frequency values before and after treatment for depression. Jiang et al. (2018), who examined this issue in the Chinese language, reported similar results. With reference to Thai language, Teeranon (2020) measured fundamental frequency and pause occurrences in the speech of depressed Thai individuals. The pause duration was approximately 1.4–1.5 times longer in those with depression than in those without depression. This finding is congruent with other previous research.

2.2 Utterance level

This assessment of depression was based on the words and utterances verbalized by the depressed individuals. Upon word level, the word categories tested in this study were compiled and synthesized from the research of Angkapanichkit et al. (2019), Francis and Pennebaker (1993), and Zimmermann et al. (2017); two main word categories were identified among depressed individuals: 1) linguistic dimension and 2) psychological dimensions. The utterance level was emphatically incorporated.

The linguistic dimension includes function words, such as pronouns; prepositions; and other grammatical aspects, such as questions and numbers.

Psychological processes include affective processes, social processes, cognitive processes, perceptual processes, biological processes, drives, time orientations, relativity, personal concerns, and informal language.

It is interesting that Ramirez-Esparza et al. (2009) found that word categories used by Spanish-speaking depressed people were more correlated with *social processes* than other types of word categories. A study by Wolohan et al. (2018) found that the words of depressed people had a high level of affective processes, such as anxiety and sadness. "I-self-focus" was also found to be an indicator of depression, consistent with the work of Zimmermann et al. (2017). Edwards and Holtzman (2017) performed a meta-analysis of the correlations between depression and first-person singular pronouns and confirmed that "I self-focus" was significantly associated with depression.

Regarding depression among Thai individuals, the research of Angkapanichkit et al. (2019) proposed a tentative word category and utterances for diagnosing depression. There are six word categories used by depressed Thai teenagers, including *deixis*, *cognitive processes*, *affective processes*, *negation processes*, *metaphor*, and *action words*. Recently, Sathientharadol (2020) found four word categories used by depressed Thai teenagers: self, disease, society, and hope.

It is worth mentioning that Angkapanichkit et al. (2019) tentatively formulated a word checklist in the Thai language for the early detection of depression. Modifiers to intensify the degree of feeling or stress are included in the checklist. Therefore, affection processes could co-occur with a modifier; for example, [kròt] 'angry/affection processes' + [ŋâai] 'easily/modifier' means [kròt.ŋâai] to 'get angry easily.' This study noted some other pragmatic strategies and pressure utterances used by depressed Thai individuals, such as 'I feel that can't

breathe.' Unlike previous research performed on Western languages, such as English, the first-person pronoun used in the utterances of depressed Thai individuals is [rau] 'we,' which is the first-person plural pronoun. Angkapanichkit et al. (2019) previously explained that depressed Thai teenagers tend to use 'we' as an inclusive pronoun and not 'self-focus' to include themselves into the story in order to signal that they are experiencing pressure and feel unworthy of being involved in the actions of others. Sathientharadol (2020) reported a similar finding that the 'hope' word category relates to subcategory of words signalling other people, such as family and friends; for example, 'friend'+'support' in 'I am lucky that my friend supports me' signals hope for depressed Thai teenagers to cope with negative feeling with the help of the others. In addition to the conceptual meaning of the words, the emotional meaning is also presented in the words used by depressed Thai teenagers.

It can be said that, at the utterance level among depressed Thai individuals, the first-person pronoun, i.e., the pronoun 'we,' and the utterances that express depression, loneliness, self-blame, and self-worthlessness could be the main clue to detect depression. It was also suggested that two or more linguistics features should be combined for screening depression.

It is clear that word categories in Thai associated with depressed thoughts need more linguistics study for solid evidence to establish a Thai depressed word-use model as an alternative method of Thai depression assessment. Additionally, more scientific and quantified research on the utterances of depressed Thai individuals will provide an explicit and well-adopted depressed word-use model.

3. Methodology

The studies of Litvinova et al. (2016) and Mairesse et al. (2007) noted that combining indexes from various studies is also crucial. Francis and Pennebaker (1993) provided early work on LIWC, while the Zimmermann et al. (2017) study is one of the most cited works in the field; Angkapanichkit et al.'s (2019) publication was the first linguistic study addressing Thai teenagers' depression. The synthesis of this research showed that there were two main categories of words: the linguistic dimension and psychological dimension.

3.1 Word categories derived from previous research on language use among depressed individuals

The linguistic dimension includes function words, such as pronouns; prepositions; and other grammatical aspects, such as questions and numbers. However, only first-person pronouns were selected to represent the linguistic dimension in the present paper, as Edwards and Holtzman (2017) reliably reported first-person pronouns to be significant well-established indicators of depression.

The psychological dimension or psychological processes include affective processes, social processes, cognitive processes, perceptual processes, biological processes, drives, time orientations, relativity, personal concerns, and informal language.

These aforementioned studies were ultimately enacted as the present research analysis framework.

3.2 Participants

The empirical data were from 10 depressed Thai volunteers attending Lampang Hospital, Thailand. All participants were diagnosed with an F.32 depressive disorder and were

between 19–24 years old. This participants' age range was assigned according to a World Health Organization (2018) statement noting that depression is most common among the 15–29-year-old age group. Moreover, it was found that 50% of patients with depression eventually attempt suicide (Cummins et al., 2015). The participants voluntarily participated in the project, and ethical approval was granted by the Lampang Hospital and the University of Phayao no. 3/019/2562.

3.3 Data collection

The 10 participants were interviewed for eight minutes using open-ended questions, such as "If you had eight minutes, what would you like to tell the doctor?" Face validity, which is appropriate for a short list of question (Connell et al., 2018; Crawford et al., 2011), was tested with three experts to justify the appropriateness of the questions. Two experts were psychiatrists working in the hospital for 20 years the other one was a depressed person using a service at the hospital. Connell et al. (2018) highlighted the value of service users' opinions on question acceptability and validity when creating a new metric. Then, the speech of the 10 depressed Thai teenagers was recorded using ICD-UX200 digital voice recorder. All speech was transcribed. Ramirez-Esparza et al. (2009) confirmed that a body of interview data of approximately 2,000 words is appropriate for analysis. This present study had a body of data of 13,537 words.

3.4 Data analysis

The 13,537 words of transcribed speech were categorized according to 11 main categories of words: first-person pronouns, affective processes, social processes, cognitive processes, perceptual processes, biological processes, drives, time orientations, relativity, personal concerns, and informal language.

First-person pronouns are words signalling the speaker, both singular and plural.

Affective processes are word categories signalling both negative and positive emotions (e.g., happy, fun, sad, and angry).

Social processes are words signalling relationships (e.g., family relationships or friendship).

Cognitive processes are words signalling words associated with thought, such as see and know.

Perceptual processes are words signalling the five human senses (e.g., hear and feel).

Biological processes are words signalling the body and consumption (e.g., body and rice).

Drives are words related to achievement (e.g., win and success).

Time orientations are words signalling time (e.g., present, past, or future statements).

Relativity is a word category that shows the direction (e.g., in, out, narrow, and end).

Personal concerns are word categories signalling personal matters (e.g., money, religion, and death).

Informal language is a word category signalling the informal use of language (e.g., abbreviation and slang).

There were two raters who are linguists in the analysis process to justify the classification of ambiguous words. Then, a table showing the word categories, frequency, and percentages was drawn.

3.5 Model validation

EFA was applied to analyze the relationships among the 11 word categories used by the 10 depressed teenagers and depression. EFA is a well-accepted statistical method that is the first step in finding correlation between variables (Marcoulides, 1998), and it is considered to be appropriate method for small samples.

The raw data of 13,537 words from word counts were integrated into the EFA. The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy and Bartlett's test of sphericity were used to test whether the raw data of 13,537 words from the 10 participants were adequate for EFA. The words from the 10 depressed teenagers were considered sufficient for EFA.

The KMO index (which ranges from 0 to 1) is considered appropriate (see Table 1) when it is around or greater than 0.50 (Chua, 2014). Bartlett's test of sphericity was shown to be significant (p =0.000, df = 55).

Table 1

KMO index and Bartlett's test of sphericity

Kaiser-Meyer-Olkin (KMO) Measure of Sampling Adequacy	**Bartlett's Test of Sphericity**		
	Chi-Square (χ^2)	**df**	**p-value**
0.5	8188.774	55	0.000

Then, AMOS software was applied for the statistical analysis of EFA. The cumulative variance was examined to confirm whether or not the reconfigured number of word categories was statistically acceptable for Thai depressed teenagers' utterances. Orthogonal rotation using the varimax method was applied to determine the factor loadings, communalities, eigenvalues, and percentage of variance for depressed Thai teenagers' utterances. Items with communalities values less than 0.50 were dropped from further analysis (Hair et al., 2006). An eigenvalue > 1 indicated a word category should be retained (Girden, 2001). A word category model was drawn. In labeling new word categories in the model, two linguists specialized in semantics were the raters of justification.

4. Results

The word frequency count (see Table 2) found that first-person pronouns accounted for 15.54% of the words used by the Thai depressed teenagers. Words signalling affective processes accounted for 5.59% of words used, social processes for 21.96%, cognitive processes for 17.13%, perceptual processes for 5.87%, biological processes for 3.46%, drives for 2.01%, time orientation for 10.50%, relativity for 3.55%, personal concerns for 7.62%, and informal language for 6.77%. The words in the social processes category were the most common among the depressed Thai teenagers, and the drive category was the least used.

Table 2
Word frequencies of the participating 10 depressed Thai teenagers

Words	**Word examples**	**Frequency**	%
Linguistic dimension first-person pronouns	[nŭu] 'I', [phûak.rau] 'we'	2,103	15.54
Psychological processes			
Affective processes	[sâu] 'sad', [nɨ̀əi] 'tired', [kròt] 'angry'	757	5.59
Social processes	[phɨ̂ən] 'friend', [khrɔ̂p.khrua] 'family', [phîi] 'brother'	2,973	21.96
Cognitive processes	[khít] 'think', [tɔ̂ŋ.kaan] 'want'	2,319	17.13
Perceptual processes	[rŭu.sɨ̀k] 'feel', [faŋ] 'listen', [hěn] 'see'	795	5.87
Biological processes	[hŭa] 'head', [khâu] 'rice'	468	3.46
Drives	[rák] 'love', [khwam.sùk] happiness	272	2.01
Time orientations	[lăew] 'already', [tɔ̀ɔ.nɨ̂əŋ] 'continuously'	1421	10.50
Relativity	[khâu] 'enter', [càak] 'from'	480	3.55
Personal concerns	[taai] 'dead', [nîi] 'debt'	1032	7.62
Informal language	[wít.lai] 'college', [mɔɔ] 'high school', [phɨ̌i.bâa] 'crazy person'	917	6.77
Total		**13,537**	**100.00**

** There are 5 tones in Thai: [ ] means mid tone, [̀] means low tone, [̂] means falling tone, [́] means high tone, and [̌] means rising tone.

The EFA identified three factors or three categories that can be derived from Table 2 with a cumulative variance of 99.9% (Table 3). From this, it could be inferred that three word categories were statistically acceptable for Thai depressed teenagers' utterances.

Table 3
Cumulative variance in depressed Thai teenagers' utterances

Factors	**Proportion variance**	**Cumulative variance**
1	99.8	99.8
2	00.1	99.9
3	00.1	99.9

After employing orthogonal rotation using the varimax method, the cumulative variance was 99%, indicating that the values calculated for EFA in Table 4 could explain a large proportion of the variation in the data.

Table 4

Factor loadings, communalities, and percentages of cumulative variance for depressed Thai teenagers utterances

Group	Word categories	Factor loadings			Communality	Eigenvalue	% of variance
		1	2	3			
1	Affective processes	0.592	0.577	0.562	.875	10.9799	99.8
	Social processes	0.627	0.544	0.557	.919		
	Cognitive processes	0.612	0.564	0.553	.930		
	Perceptual processes	0.599	0.568	0.564	.921		
	Biological processes	0.595	0.572	0.563	.886		
	Time orientations	0.618	0.587	0.523	.925		
	Personal concerns	0.604	0.565	0.562	.914		
	Informal language	0.614	0.573	0.542	.943		
2	First-person pronouns	0.565	0.625	0.538	.931	0.0068	00.1
3	Drives	0.562	0.581	0.589	.919	0.0059	00.1
	Relativity	0.585	0.554	0.592	.959		
Total							99.9

All factor loadings in Table 4 were found to be higher than 0.50, with three word categories according to the factor loadings. Next, the communalities of the items were all higher than 0.50; therefore, all word categories could be used for further analysis. Eigenvalues > 1 were detected from the EFA in the first group of word categories, while the second and third groups had eigenvalues < 1. All word categories could explain 99.9% of the variance in total.

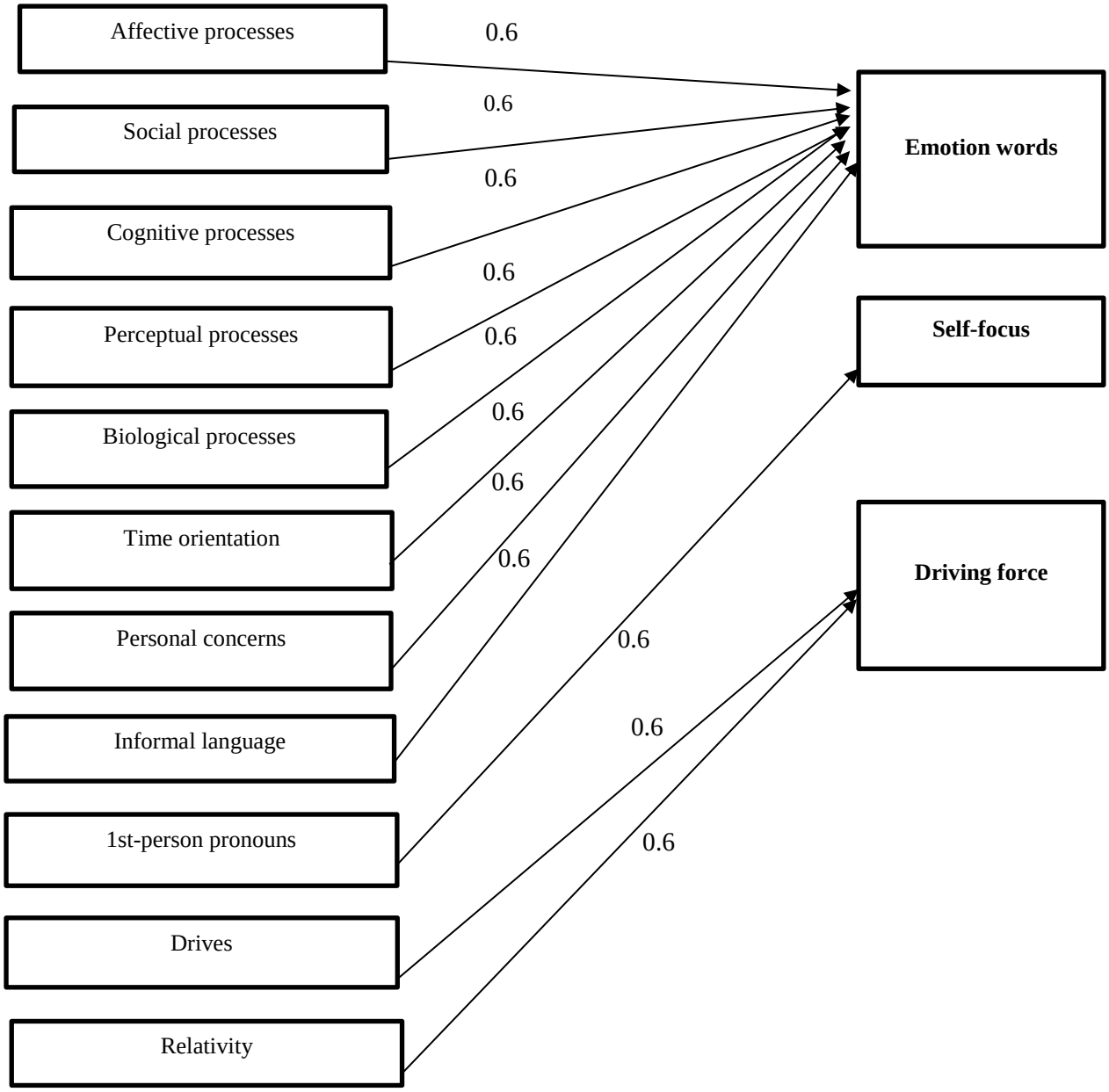
The first word category included eight subcategories, namely, affective processes, social processes, cognitive processes, perceptual processes, biological processes, time orientations, personal concerns, and informal language, with factor loadings of 0.562 – 0.627. This category was titled “Emotions words”.

The second word category included first-person pronouns with factor loadings of 0.544 – 0.625. This category was titled “Self-focus.”

The third word category included drives and relativity, with factor loadings values of 0.523 – 0.592. This category was titled “Driving force.”

The factors and their contributions led to three word categories, as shown in Figure 1.

Figure 1
Word use model for depressed Thai teenagers



5. Discussion

Depressed individuals reflect their thoughts through their verbalized language. This study aimed to analyze the word categories used by depressed Thai teenagers based on LIWC and to quantify a model of word use for depressed Thai teenagers' utterances.

Words signalling social processes, such as friend, family, and brother, are the most used by depressed Thai teenagers. Words signalling cognitive processes account for 17.13% of

words, first-person pronouns for 15.57%, and the drive category for 2.01%. These findings are consistent with the studies of Angkapanichkit et al. (2019) and Sathientharadol (2020).

Similar to previous research in other languages, first-person pronouns are highly used by depressed Thai teenagers. This finding is in line with that of Zimmermann et al. (2017) and a quantitative study of Edwards and Holtzman (2017). The findings on the speech of the depressed Thai teenagers' in this study confirm Pyszczynski and Greenberg's (1987) model of self-focus. It is widely accepted that depressed individuals have lower self-regulation and compensate for this with self-focus. This assumption is also consistent with the research of Ramirez-Esparza et al. (2009) and Stirman and Pennebake (2001). Compared to Angkapanichkit et al. (2019), which found "we" to be an inclusive pronoun used by depressed teenagers in the Thai language, this present study is more likely to be congruent with Beck's theory of self-focus. This might be due to the different methods of data collection, as Angkapanichkit et al. (2019) mainly used the diary writings of depressed teenagers, while other studies used social media content or speech. The style of language use seems to affect the words used by depressed teenagers. There is another point suggesting that the style of language use among depressed individuals affects the research result. Yoo and Ha (2019) suggested that informal words, such as slang and abbreviations, were used as much as storytelling; however, the present study found a contrasting result, showing that the informal word category was less used by the depressed Thai teenagers. This difference is due to the fact that Yoo and Ha (2019) collected their data from social media, while the present study used interview data.

The drives category was found to be the least used by depressed Thai teenagers; this can be explained by the suggestion of Angkapanichkit et al. (2019) that depressed individuals feel tired and less motivated to live their lives. While Angkapanichkit et al. (2019) found that negative emotion was expressed by depressed individuals, Sathientharadol (2020) semantically linked hopeful words with social processes. Sathientharadol reached the conclusion that understanding from family and friends can give hope to depressed individuals.

Quantitative analysis using EFA in this study significantly indicated that there were approximately three word categories used by depressed Thai teenagers.

The first word category includes eight subcategories, namely, affective processes, social processes, cognitive processes, perceptual processes, biological processes, time orientations, personal concerns, and informal language with factor loadings of 0.562 – 0.627. This category is titled "Emotion words."

The second word category includes first-person pronouns, with factor loadings of 0.544 – 0.625. This category is titled "Self-focus."

The third word category includes drives and relativity, with factor loadings of 0.523 – 0.592. This category is titled "Driving force."

It is interesting that the first word category used by depressed Thai teenagers in this study includes eight LIWC word categories, including words signalling emotion among depressed individuals. This word category, renamed "Emotions words," appears to serve as the foundation for semantic categories as mentioned in Rauh (2016).

First-person pronouns were statistically classified as an independent category called "Self-focus." However, as the eigenvalue is less than 1; this category is unstable and requires more data to confirm its relationship with depression in Thai teenagers. Similarly, the third word category found in this present study called "Driving force" comprises drives and relativity and also has an eigenvalue less than 1. This category is undoubtedly in need of more data to confirm its significance. Regarding limitations, as it was found after employing EFA that the eigenvalues in the second and the third categories were less than 1; thus, further studies with larger sample sizes must be conducted.

Rude et al. (2004) and Schwartz (2018) suggest that the aforementioned results show depressed Thai teenagers are deeply dominated with heightened emotion and preoccupied by self-awareness thoughts that were then verbalized into language. Word use in depressed Thai teenagers signals vulnerability to depression.

However, the drives word category includes positive words, such as success and hope. This is in line with the results of Sathientharadol (2020), who found that the hope word category is used by depressed Thai teenagers. Sathientharadol (2020) explained that hope words signal help from people in society, which affects the feeling of depressed speakers. Moreover, it can be inferred that the co-occurrence between word categories might be a trend among depressed Thai individuals; for example, 'I' + 'success' was a word combination structure found to signal depression in the Thai language by Angkapanichkit et al. (2019). It seems that such words and their combinations are crucial for identifying depression. The present study has also yielded congruent results that depressed Thai teenagers tend to use the storytelling in the "Driving force" category, such as 'My mother helps me out,' which supports the findings of Sathientharadol (2020) that depressed Thai teenagers seek help from the others. The words in the social processes category (words related to friends and family) were the most used among the depressed Thai teenagers in this study, which also supports this finding.

Regarding the number of word categories used by depressed Thai teenagers, this study has proposed three categories of words linked to depression or word-use model for depressed Thai teenagers: "Emotion words," "Self-focus," and "Driving force." Meanwhile, previous research on the Thai language has found higher numbers of categories. Angkapanichkit et al. (2019) reported approximately six word categories, and Sathientharadol (2020) found four word categories used by depressed Thai teenagers. Applying the EFA method to establish a linguistic word-use model associated with Thai depression postulates different word-use categories. This present quantified results lead to ample evidence of a language word-use model to entail depressive signs than by a self-evaluated form as mentioned in Rude et al. (2001). The first category titled "Emotion words" had an eigenvalue higher than 1, which seems to confirm the cognitive model that depressed individuals highly express their feelings. This category of words can be used as an indicator for the diagnosis of depression (O'Dea et al., 2021; Zapata-Vega et al., 2010).

However, pragmatics of the Thai language used by Thai teenagers may have significant influence on the utterances; qualitative analysis of words is suggested for further research.

6. Conclusion

Based on our statistical analyses, three categories of words used in the speech of Thai depressed individuals are proposed. Utterances and word categories can be used for the early

detection of depression. Most of the identified word categories signal emotive meaning. This present result indicates the importance of emotive and self-centred thought to enact depressive vulnerability. In addition, driving force word-use category also indicates help-seeking of the depressed Thai teenagers.

Acknowledgements

I would like to extend my deepest gratitude to Program Management Unit on Area Based Development (PMU A) and Unit of Excellence of the School of Liberal Arts for research grants with human ethics number UP-HEC 3/018/62 and UP-HEC 2.2/053/64. I would also like to express my gratitude to Professor Dr. Unchalee Wongwattana for her valuable comments on the first draft of the paper.

Author

Phanintra Teeranon is an Associate Professor at the School of Liberal Arts, University of Phayao, teaching courses primarily in linguistics and area-based research. She received her PhD at Chulalongkorn University and has been researching and writing articles on linguistics and education. Conducting research on social issues to develop alternative solutions is currently her major concern. She has been awarded the "Excellent Research Award" from the University of Phayao and National Research Council of Thailand.

References

Angkapanichkit, J., Rojanahasadin, A., & Inthasian, S. (2019). *Language communication and depression: A survey of communication development in detection depression for sustainable life*. A Research Report, Thailand Research Fund. (in Thai)

Beck, A. T., & Clark, D. A. (1997). An information processing model of anxiety: Automatic and strategic processes. *Behaviour Research and Therapy*, *35*(1), 49-58.

Beck, A. T., & Haigh, E. A. P. (2014). Advances in cognitive theory and therapy: The generic cognitive model. *Annual Review of Clinical Psychology*, *10*, 1-24.

Breznitz, Z. (1992). Verbal indicators of depression. *The Journal of General Psychology*, *119(4)*, 351-363.

Chaveepojnkamjorn, W., Pichainarong, N., Adthasangsri, V., Sativipawee, P., & Prasertsong, C. (2017). Depression and its associated factors among senior high school students in Nonthaburi Province, Thailand: A cross-sectional study. *JPHDC*, *2(3)*, 224-234.

Chen, Y.-H., & Jin, C. (2017). *A corpus-based study of depressive language in online teen health communications*. An Unpublished Working Paper of the University of Nottingham, University of Nottingham.

Chua, Y. P. (2014). *Statistik penyelidikan lanjutan: Ujian regresi, analisis faktor dan analisis sem*. McGraw Hill (Malaysia) Sdn Bhd.

Connell, J., Carlton, J., Grundy, A., Taylor, B. E., Keetharuth, A. D., Ricketts, T., Barkham, M., Robotham D., Rose D., & Brazier J. (2018). The importance of content and face validity in instrument development: lessons learnt from service users when developing the Recovering Quality of Life measure (ReQoL). *Quality of Life Research*, *27(7)*, 1893-1902.

Crawford, M. J., Robotham, D., Thana, L., Patterson, S., Weaver, T., Barber, R., Wykes, T., & Rose, D. (2011). Selecting outcome measures in mental health: The views of service users. *Journal of Mental Health, 20(4),* 336-346.

Cummins, N., Scherer, S., Krajewski, J., Schnieder, S., Epps, J., & Quatieri, T. (2015). A review of depression and suicide risk assessment using speech analysis. *Speech Communication, 71,* 10-49.

De Choudhury, M., Gamon, M., Counts, S., & Horvitz, E. (2013). Predicting depression via social media. *Proceedings of the International AAAI Conference on Web and Social Media*, *7*(1), 128-137.

Edwards, T., & Holtzman, N. S. (2017). A meta-analysis of correlations between depression and first person singular pronoun use. *Journal of Research in Personality, 68,* 63-68.

Evans, V. (2012). Cognitive linguistics. *Wiley Interdisciplinary Reviews: Cognitive Science, 3*(2), 129-141.

Fávero, L. P., & Belfiore, P. (2019). *Data science for business and decision making*. Academic Press.

Francis, M. E., & Pennebaker, J. W. (1993). LIWC: Linguistic inquiry and word count. *Technical report*. Southern Methodist University.

Girden, E. R. (2001). *Evaluating research articles from start to finish*. Sage Publications.

Greden, J. F., & Carroll, B. J. (1980). *Decrease in speech pause times with treatment of endogenous depression*. *Biological Psychiatry*, *15*(4), 575-587.

Greden, J. F., Albala, A. A., Smokler, I. A., Gardner, R., & Carroll, B. J. (1981). Speech pause time: a marker of psychomotor retardation among endogenous depressives. *Biological Psychiatry*, *16*(9), 851-859.

Hair, J., Black, W., Babin, B., Anderson, R., & Tatham, R. (2006) *Multivariate data analysis*. Pearson Prentice Hall.

Jarrold, W., Javitz, H. S., Krasnow, R., Peintner, B., Yeh, E., Swan, G. E., & Mehl, M. (2011). Depression and self-focused language in structured interview with older men. *Psychological Reports*, *109*(2), 686-700.

Jiang, H., Hu, B., Liu, Z., Wang, G., Zhang, L., Li, X., & Kang, H. (2018). Detecting depression using an ensemble logistic regression model based on multiple speech features. *Computational and Mathematical Methods in Medicine, 2018,* 1-9.

Lieberman, M. A., & Goldstein ,B. A. (2006). Not all negative emotions are equal: The role of emotional expression in online support groups for women with breast cancer. *Psychooncology, 15*(2), 160-168.

Litvinova, T., Zagorovskaya, O., Litvinova, O., & Seredin, P. (2016). Profiling a set of personality traits of a text's author: A corpus-based approach. *International Conference on Speech Compututer, 9811,* 555-562.

Marcoulides, G. A. (1998). *Modern methods for business research*. Psychology Press.

Mairesse, F., Walker, M., Mehl, M., & Moore, R. (2007). Using linguistic cues for the automatic recognition of personality in conversation and text. *Journal of Artificial Intelligence Research (JAIR), 30,* 457-500.

Nilsome, A. (1998). Acoustic analysis of speech variables during depression and after improvement under antidepressants. *European Neuropsychopharmacology, 76*(3), 235-245.

O' Dea, B., Boonstra, T., Larsen, M., Nguyen, T., Venkatesh, S., & Christensen, H. (2021). The relationship between linguistic expression in blog content and symptoms of depression, anxiety, and suicidal thoughts: A longitudinal study. *PLOS ONE, 16*(5), 1-29.

Pinker, S. (1998). *Words and rules: The ingredients of language*. Weidenfeld and Nicolson.

Pyszczynski, T., & Greenberg, J. (1987). Self-regulatory preservation and the depressive self-focusing style: A self-awareness theory of reactive depression. *Psychological Bulletin*, *102*, 122-138.

Ramírez-Esparza, N., Chung, C. K., Sierra-Otero, G., & Pennebaker, J. W. (2009). *El lenguaje de la depresión: Categorías lingüísticas y temas usados en foros de discusión en el internet en Inglés y en Español.* [The language of depression: Linguistic categories and themes used in discussion forums on the internet in English and Spanish]. Revista de la Asociación de Psicoterapia de la República Argentina, Julios.

Rauh, G. (2016) Linguistic categories and the syntax-semantics interface: Evaluating competing approaches. In J. Fleischhauer, A. Latrouite & R. Osswald (Eds.), *Explorations of the syntax-semantics interface: Studies in language and cognition*, *3*, pp. 15-55. Düsseldorf University Press.

Rude, S. S., Covich, J., Jarrold, W., Hedlund, S., & Zentner, M. (2001). Detecting depressive schemata in vulnerable individuals: Questionnaires versus laboratory tasks. *Cognitive Therapy and Research*, *25*(1), 103-116.

Rude, S. S., Gortner, E. M., & Pennebaker, J. W. (2004). Language use of depressed and depression vulnerable college students. *Cognition and Emotion*, *18*(8), 1121-1133.

Sathientharadol, P. (2020). A semantic network analysis of Thai utterances used by depression patients. *Academic Journal of Humanities and Social Sciences Burapha University*, *28*(3), 116-142. (in Thai)

Schwartz, A. (2018). *Linguistic analysis of written language used by young adults with and without invisible disabilities*. [Unpublished honors thesis, The University of Tennessee at Chattanooga].

Stirman, S. W., & Pennebaker, J. W. (2001). Word use in the poetry of suicidal and nonsuicidal poets. *Psychosomatic Medicine*, *63*(4), 517-522.

Teeranon, P. (2020). An acoustic study of fundamental frequency and pauses in depressive teenagers in Thailand: A pilot study. *Journal of Universal Language*, *21*(1), 89-111.

Vicsi, K., Sztahó, D., & Kiss, G. (2012). Examination of the sensitivity of acoustic-phonetic parameters of speech to depression. *Proceeding of the 3rd IEEE International Conference on Cognitive Infocommunications*. December 2-5, 2012, Kosice, Slovakia.

Vogel, A. P., Fletcher, J., Snyder, P. J., Fredrickso, A., & Maruff, P. (2011). Reliability, stability, and sensitivity to change and impairment in acoustic measures of timing and frequency. *Journal of Voice*, *25*(2), 137-149.

Williamson, J. R., Godoy, E., Cha, M., Schwarzentruber, A., Khorrami, P., Gwon, Y., Kung, H-T., Dagli, C., & Quatieri, T. F. (2016). Detecting depression using vocal, facial and semantic communication cues. *Proceedings of the 6th International Workshop on Audio/Visual Emotion Challenge*. ACM, 11-18.

Wolohan, J. T., Hiraga, M., Mukherjee, A. Sayyed, Z. A., & Millard, M. (2018). Detecting linguistic traces of depression in topic-restricted text: Attending to self-stigmatized depression with NLP. *Proceedings of the First International Workshop on Language Cognition and Computational Models*. Retrieved from https://www.aclweb.org/anthology/W18-4102.pdf

World Health Organization. (2018). *Depression: Let's talk.* Retrieved August 17, 2019 from https://www.who.int/mental_health/management/depression/en/

Yang, Y., Fairbairn, C., & Cohn, J. F. (2013). *Detecting depression severity from vocal prosody*. *IEEE Transactions on Affective Computing*, *4*(2), 142-150.

Yoo, M. L., & Ha, T. (2019). Semantic network analysis for understanding user experiences of bipolar and depressive disorders on Reddit. *Information Processing & Management*, *56*(4), 1565-1575.

Zapata-Vega, M. I., De la Grecca, R., Altamirano, A., Gonzaga, V., & Vega-Dienstmaier, J. M. (2010). Diagnosing depression: Symptoms and vocabulary used by a sample of general practitioners. *Revista de Neuro-Psiquiatría*, *73*(3), 77-83.

Zimmermann, J., Brockmeyer, T., Hunn, M., Schauenburg, H., & Wolf, M. (2017). First-person pronoun use in spoken language as a predictor of future depressive symptoms: Preliminary evidence from a clinical sample of depressed patients. *Clinical Psychology and Psychotherapy*, *24*(2), 384-391.