The Annotated Lexicon of Chinese Emotion Words

Bee Chin NG, Can CUI and Francesco CAVALLARO

Abstract:

This study explores language specificity in the organisation and distribution of emotion words in Mandarin Chinese. Anecdotal references to the prevalence of use of verbs in some languages (e.g. Russian and Mandarin) for expressing emotion words compared to English has not been supported by empirical evidence from a comprehensive study of the emotion words vocabulary, and, despite a proliferation of cross-linguistic studies of emotion words, a database of a corpus of emotion words across languages is absent. Using the framework proposed by Pavlenko (2008), emotion words in Mandarin Chinese are extracted, sorted into three semantic categories; emotion words, emotion laden words and emotion related words, and tagged for frequency of use, valency, intensity and parts of speech. Corpus data analysis method was then employed to study the patterns of the data. Consistent with other reports on Mandarin Chinese in other domains (e.g. acquisition), it was found that verbs occupied the biggest percentage in both emotion words and emotion-related words categories. An analysis of the valence and intensity of emotion words shows cross-linguistic divergence from other studies reported. The study also represents a significant attempt at providing a working template for the identification of emotion words in emotion research.

Keywords: emotions; emotion-laden words; emotion-related words; Mandarin Chinese; valence

1. Introduction

Emotion research or 'Affective Science', is a burgeoning field attracting different lines of enquiry from multiple disciplines like psychology, anthropology and computer science. Just inputting the term 'emotion' in the Google Scholar search engine throws up 2.9 million published research from a variety of disciplines. This interdisciplinary interest in emotion is not surprising given its indisputable and vital role in human interaction. While emotion has traditionally been perceived as a mental state and independent of language, many scientists are revisiting the Sapir-Whorfian concept that our language plays a role in influencing our interpretations of events and cognition (Pavlenko 2014). This presents the scientific community with a challenging field of inquiry into the cultural aspects of expressing affect in diverse communicative settings. However, despite this surge in interest and studies on language and emotion, there is still a lack of agreement among researchers on the representations and classification of emotions across languages and cultures to allow for a coherent comparison. A cursory view of the field reveals disparate approaches with groups of researchers working within their own circumscribed understanding of the issue (cf. Scherer & Moors 2018;

Kringelbach & Berridge 2017; Just, Pan, Cherkassky et al. 2017 for a sample of a recent varied approaches to the study of emotion).

Apart from English, Chinese is a language that is spoken by great numbers of people in many parts of the world (Simons & Fennig 2018; Siegel 2018). However, while English has been intensively studied, there is still a lack of corroborated data with regards to the classification of the emotion terms in Chinese. While various researchers have worked on list of emotion words or specific emotion in Chinese (e.g. Bai 2015; Chang et al. 2000; Lee 2010; Li, Wang & Fischer 2004), there is no systematic list of Mandarin Chinese emotion words. Xu & Tao (2003) did compile a list of 374 Chinese emotion terms by extracting them from a dictionary, but the extracted emotion words contain only verbs and adjectives. They left out nouns and adverbs without really explaining the process by which these words were culled. As pointed out by Lin and Yao (2016), the selection of these words was subjectively made by the researchers themselves without a clear indication of what criteria were used.

The lack of Chinese emotion lists or corpus provides the impetus for this study which aims to find a suitable framework to categorise emotion terms, so as to select and compile a list of emotion terms in the Chinese language. The development of a definitive list of emotion words will enable the study of the patterns of how Chinese speakers express their emotions in daily conversations and also will provide a proper coding of emotion terms for future studies, especially in the field of sentiment analysis, where the knowledge of our emotion lexicon can play a key role in how we analyse emotion quality in text (Vinodhini & Charasekaran 2012). It also has critical implications for the choice and use of emotion words stimuli in psycholinguistics research where we commonly draw conclusions about how we process different emotions. For example, Altarriba and Basnight-Brown (2011) found that different emotion words in English and Spanish provided different priming effects.

Up till now, what qualifies as an emotion word is still a site for debate. What we attempt to do in this study is to adapt the model proposed by Pavlenko (2008) to Mandarin Chinese. In doing so, we hope that the data will provide a starting point for an overall comparison of cross-linguistic emotion terms and concepts.

2. Background

Traditional studies on emotion originally adopted a naturalist approach to classifying emotion terms (Ogarkova 2007), generally by determining universal states of emotion without linguistic or cultural considerations. The common problem encountered by this approach however, is determining the number of basic emotional states that exist. This is one major stumbling block, which researchers have failed to agree upon (Pavlenko 2014). In his studies of facial expressions recognition, Ekman (2003) proposed that there are seven discrete basic emotions that can be identified universally across cultures – *anger*, *disgust*, *contempt*, *surprise*, *fear*, *happiness* and *sadness*. Although a more recent study (based also on facial expressions) by Jack, Garrod and Schyns (2014) proposed instead a classification of four basic emotions– *glad*, *sad*, *mad* and *scared*. Attesting the universality of emotion has been an issue taken up mainly by researchers working from a 'universalist' perspective. It is however, difficult, if not impossible to dissociate the study of emotion from language/culture and many of the early

research have been criticised for its eurocentric approach to the field (Pavlenko 2008). Indeed, as pointed out by Wierzbicka (1999: 24), the concept of 'emotion' itself is in fact very unique to English in its "reference to 'feeling', a reference to 'thinking', and a reference to a person's body." She aptly pointed out while one can talk about 'a feeling of hunger' or 'a feeling of heartburn', an expression like 'an emotion of hunger' or 'an emotion of heartburn' makes no sense. Even closely related languages like German do not even have a translational equivalent for a concept that is similar to the English concept of 'emotion'. Wierzbicka instead, favours the use of 'feelings' which, she argues, is less culturally bound than 'emotion'. Chinese also does not have an superordinate term that is equivalent to 'emotion' in English but the term 情 *qing* or 情感 *qinggan* 'literally emotion + feeling' has been loosely used as a translation for 'emotion', though there is acknowledgement that they do not refer to the same thing as 'emotion' in English (Eifring 2004).

Apart from criticism of an ethnocentric standpoint, another concern is that the language used in academic writing, particularly English, is the technical medium commonly used to describe research findings on emotion studies and English is after all, not a culturally neutral language (Enfield & Wierzbicka 2002; Eifring 2004; Pavlenko 2008). In their discussion on the treatment of emotion in the literature, it is clear that using English as a metalanguage to discuss emotions, especially in other languages risks dismissing lexical diversity and the "lexical categories of any distant culture in favour of one's own" (Wierzbicka 1999: 25). In addition, it is evident that not all languages encode emotional meaning in the same manner, as different studies have found variances in emotion words across languages. While it is not unusual to find that the emotion categories identified by Eckman to be universal, the experience and interpretation of these categories have been found to deviate across cultures and even gender (e.g. Markus & Kitayama 1991; Russell & Yik 1996; Kitayama, Mesquita & Karasawa 2006; Fischer, Rodrigruez Mosquera, van Vianen & Manstead 2004).

The subjective evaluation of speakers across cultures to each emotion is also known to be different. Pavlenko (2008) illustrated this point with an extensive list of crosslinguistic examples. For example, the source or the cause of jealousy is salient in Russian. *Revnost* 'jeolousy' is caused by one's boyfriend and *zavist* 'envy' is caused by someone's good fortune. This distinction in the cause is not the case for 'envy' in Chinese. In fact, *xianmu* 羡慕 in Chinese is often used in a positive context to show admiration. Furthermore, 'envy' in English has a far more intensive negative semantics than its translational equivalent in Chinese. Other examples include differences in appraisals, captured in appraisal theories. Appraisal theories approach the study of emotions from the perspective that emotions are made up of a collection of components. Frijda and Mesquita (1998) propose that emotions involve seven components namely, *Antecedent Events, Appraisal, Significance, Regulation, Arousal, Experience and Action Readiness*[1]. The sense of each emotion may change according to the changes in one or more component. A situation or event may, therefore, be interpreted in many ways which lead to the experience of different emotion (Scherer, Schorr & Johnstone 2001). Over the course of a

[1] *Antecedent Events* are situations or events that evoke emotions.
Appraisal refers to the evaluation that links perception of an event to emotional meaning.
Significance is the evaluation of the experience in relation to oneself.
Regulation refers to the inhibition or enhancement of emotions.
Arousal refers to the physiological responses to emotions.
Experience refers to the awareness of some or all components.
Action readiness is how inidviduals act to modify or entertain some form of relationship with the environment.

decade, Scherer and his colleagues (Scherer 2005; Scherer, Shuman, Fontaine & Soriano 2013) extended their version of Appraisal Theory for emotion which is now widely used. Hence, an emotion is experienced based on personal evaluation of the event in relation to the values and goals held by the speaker. The approach of viewing an emotion as a sequence of components in an event allows any emotion to be dissected and discussed in components and allows for cultural differences in interpretations.

Hence, the cause of 'jeolousy' is important in Russian, the subjective evaluation of 'jealousy' or 'envy' is not as intensively negative in Chinese as it is in English. The Appraisal Theory approach also incorporates the Dimensional Approach (Zammuner 1998) to emotion as each emotion can also be assessed in terms of valence (positive-negative) or intensity (high intensity – low intensity). Cross-culturally, we also differ in the way we choose to regulate or control our emotion. As pointed out by Pavlenko (2008), English speakers typically see regulation and control of 'anger' as a sign of maturity, for Samoan and Ilongot speakers, expressing 'anger' is a sign of having attained manhood. Similarly, the concept of *en* 恩 in Chinese describes 'the strong sense of obligation that needs to be fulfilled', not adequately captured in the English equivalent 'gratitude'. This is expressed in idioms like *en zhong ru shan* 恩重如山 '*en* as heavy as a mountain' and 'to forget *en* and loyalty' (*wan gen fu yi* 忘恩负义) is one of the worst descriptions you could use to describe a Chinese person.

While the idea of 'love' may be universal, across cultures we talk about love in very different ways. Soueif (1999: 386-87) describes Arabic love lexicon in intriguing detail:

- *Hubb* is love
- *ishq* is love that entwines two people together,
- *shaghaf* is love that nests in the chambers of the heart,
- *hayam* is love that wanders the earth,
- *teeh* is love in which you lose yourself,
- *walah* is love that carries sorrow within it,
- *sababah* is love that exudes from your pores,
- *hawa* is love that shares its name with 'air' and with 'falling',
- *gharm* is love that is willing to pay the price

It seems languages are idiosyncratic about what emotion they elaborate. For example, while Arabic has many words for 'love', Chinese encodes the concept of 'shame' in an extensive way. In fact, Li, Wang and Fischer (2004) identified 113 terms denoting 'shame' in Chinese, argueing that 'shame' is an important organisational principle governing social relationship amongst the Chinese.

Researchers working on emotion are also aware of many emotions which are unique to particular languages and often has no translational equivalence, such as the example of Japanese *amae* reported above. Lomas (2016) listed several such complex emotions that have no translation in English:

- *Gula* – Spanish for the desire to eat simply for the taste
- *Sobremesa* – Spanish for when the food has finished but the conversation is still flowing
- *Mbukimvuki* – Bantu for "to shuck off one's clothes in order to dance"
- *Schnapsidee* – German for coming up with an ingenious plan when drunk
- *Volta* – Greek for leisurely strolling the streets
- *Gokotta* – Swedish for waking up early to listen to bird songs

- *Suaimhneas croi* – Gaelic for the happiness that comes from finishing a task
- *Iktsuarpok* – Inuit for the anticipation felt when waiting for someone
- *Vacilando* – Greek for the idea of wandering, where the act of travelling is more important than the destination
- *Philotimo* – Greek virtue – to love to honour

Many other terms, e.g. *schadenfreude* (from German – describing 'pleasure derived from someone's misfortune') or *mono no aware* (from Japanese, 'describing a gentle beautiful feeling of sadness') have been borrowed into the English lexicon because of the lexical gap. While *schadenfreude* has entered common parlance, *mono no aware* is still currently confined to cultural studies of aesthetics.

Underpinning study of emotion expressions is the assumption that emotion words exist and they are identifiable. To date, there have been a few separate attempts at compiling emotion lexicon in different languages. To this end, there also seems to be a wide range in the number of emotion words across different languages. Dutch is believed to have 1,501 words for describing emotions, while Taiwanese Chinese seems to have much fewer at 750 words and Malay could have as few as 230 words (Russell 1991). Chewong, an Aslian language is reported to have only seven emotion words (Howell 1984). One would also suspect that categorisation and enumeration of emotion or emotion words largely depend on the set of criteria used and the number, range and categories of emotion words may vary considerably across research studies. For example, based on a dictionary-derived search for nouns and adjectives, Wallace and Carson (1973) generated an initial master list of over 2000 English words that encode emotional meaning. While this appears to work for English, this may not be as applicable to other languages that tend to lexicalise emotions in verbs such as, Polish and Russian (Pavlenko 2002). For example, in the Russian example, *pis'mo ee rasstraivaet* 'the letter angers her', *rasstraivaet* is a verb meaning to upset, and this is a common structure used to express the feeling of anger. Such examples are common in Chinese. *Wo huo*le 我火了 (literally I+fired) 'I am furious' or *wo bengk*uile 我崩溃了(literally I+ crumbled from extreme stress) 'I collapsed' are emotions expressed structurally by verbs. The differences in the emotion lexicon across languages thus make it essential to carry out further in-depth studies to consider how different languages uniquely encode for emotional meaning and even more fundamentally, how do researchers identify emotion words across languages.

Scherer (2013) advocated taking a lexical approach to the study of emotion and suggested that, given the primacy of emotion in social interaction, adequate and meaningful representations of emotion states must be present in the substance of language (see lexical sedimentation hypothesis, John, Angleitner & Ostendorf 1988; Saucier & Goldberg 1996). Examining the semantic organisation and distribution of emotion words may be able to shed light on the structure of emotion concepts, revealing universal as well as culture- and language-specific features. Unfortunately, this line of inquiry has not been extensively explored – linguistic observations, such as the prevalence of verbs that express emotion in Russian and Mandarin Chinese for example, remain anecdotal or rooted in discourse-based work. Yet, the preponderance of nouns versus verbs is psychologically a critical distinction as it has implications for the status of emotion as 'state focused', in the case of languages with a noun bias and 'process focused' in the case of languages with a verb bias. Noun or Verb bias has

been found to have a significant impact on acquisition processes and language processing in general (Peter, Chang, Pine, Blything & Rowland 2015; Thothathiri, Evans & Poudel 2017). Empirical approaches, as a complement to linguistic investigation, could aid in verifying such observations and understanding their implications in this regard.

Up to now, despite a proliferation of cross-linguistic studies of emotion words, a database of a corpus of emotion words across languages is absent. Existing emotion word lists, such as the Affective Norms for English Words (ANEW; Bradley & Lang 1999), were designed to serve as stimuli databases and do not necessarily reflect the full working emotion vocabularies of a given language. Additionally, these word lists are difficult to compare cross-linguistically as they vary widely in their manner of creation and characterisation, stemming from the different emotion models upon which they were based. The ANEW (Bradley & Lang 1999), for example, is based on a dimensional model of emotion and its words were selected on a "common-sense" basis and characterised on three dimensions of valence, arousal and dominance. Bradley and Lang's (1999) study has also been carried out in other languages, such as Italian (Montefinese, Ambrosini, Fairfield & Mammarella 2014), German (Schmidtke, Schröder, Jacobs & Conrad 2014), Portuguese (Soares et al. 2012), and Spanish (Redondo, Fraga, Padrón et al. 2007; Ferré, Guasch, Moldovan et al. 2012; Ferré, Guasch, Martinez-Gracia et al. 2017). These translations of the ANEW list to other languages often report cross-linguistic and cross-cultural differences in ratings (see also Eilola & Havelka 2010), suggesting that the same ANEW word in different languages may not bear the same emotional connotation or express equivalent concepts. The very fact that the list of words is derived from translations is in itself, a limitation.

Some studies have used semantic dimensions to specifically distinguish emotion from non-emotion words and have also utilised different combinations (e.g. valence, intensity and duration in Zammuner (1998); valence, arousal, and imagery in the Dictionary of Affect in Language or DAL, Whissell (2008)). Other word lists are instead compiled by discrete emotion models. For example, in Hamid et al.'s (1999) study, emotion words were elicited from participants imagining and describing a basic emotion state. While Stevenson, Mikels, and James (2007) and Briesemeister, Kuchinke and Jacobs (2011) provided word norms, which limit the categorisation of emotion words to five discrete emotion categories – happiness, sadness, anger, fear and disgust. The Berlin Affective Word List (BAWL) works along the same lines as the DAL and ANEW, in that it comprises the emotional valence and imageability ratings of over 2,200 German words (Võ, Jacobs & Conrad 2006). In its upgraded version, renamed the 'Berlin affective Word List Reloaded' (BAWL-R), it now comprises the valence, imagineability, and arousal ratings of over 2,900 German words (Võ, Conrad, Urton, Hofman & Jacobs 2009). Riegel, Wierzba, Wypych et al. (2015) adapted the BAWL-R to Polish (see the Nencki Affective Word List (NAWL)).

Such corpora of emotion words are and have been useful for the semantic mapping of emotion words across languages. For example, after adapting the original ANEW data set to European Portuguese, Soares et al. (2012) then compared their results with those from the American (see Bradley & Lang 1999) and Spanish (see Redondo, Fraga, Padrón et al. 2007) studies. They found that their European Portuguese sample had lower emotional reactivity to

the ANEW words in comparison to the American participants. The Spanish had the highest scores in terms of arousal and the Americans the highest scores in terms of valence.

Currently, what we know is that existing word lists can inform cross-linguistic inquiry. It is clear that levels of arousal/intensity and degrees of valence do vary across languages. However, the lack of consistency in how these wordlists are compiled impinges on the reliability of these comparisons. Comprehensive emotion corpora then become essential to exploring these cross-linguistic emotion concepts in an empirically supported manner.

The creation of emotion corpora is unfortunately hampered by a lack of comparable ways of identifying emotion words. One possibility involves propositional analysis – for example, Wallace and Carson (1973) searched dictionaries for adjectives and nouns that fit the syntactic contexts "He has a feeling of X" and "He feels X" (where "X" refers to the emotion term). However, it is debatable what would be the best context to use (for other examples see Ortony, Clore & Foss 1987; Johnson-Laird & Oatley 1989). Additionally, this approach is language-specific, as the linguistic structures commonly used to express emotion in one language may not be applicable to others. In English and French, emotions are usually expressed using adjectives, while emotions in Russian or Polish languages are more frequently described using verbs referring to processes and relationships (Pavlenko 2002; Wierzbicka 1994, 2004). The cognitive implications of these differences in encoding emotion are still largely unexplored. For example, does a focus on verbs and as a result a focus on processes and relationships influence the way emotions are experienced and conceptualised?

Attempting to create a language- and theory-general method of building emotion corpora, Pavlenko (2008) proposed a framework for the identification and extraction of emotion words by describing three categories of emotion terms. The first are emotion words which directly refer to a particular emotion state (e.g. happy, angry) or process (e.g. to worry, to rage), and generally fit the syntactic context "I am/I feel X". The second are emotion-laden words which are words that elicit emotions from interlocutors such as swear words, endearments, reprimands, and interjections. The third are emotion related words which describe the behaviours of particular emotions without naming the actual emotions, including facial expressions (e.g. smile, frown), bodily symptoms (e.g. cry, shiver), and action tendencies (e.g. scream, escape). This is a productive way of approaching the issue as it neatly classifies emotion terms into those that refers explicitly to emotions and those that are related to emotions or have emotive overtones. This distinctive categorisation is useful as it allows an incisive way of separating lexical items that are strictly describing an emotion to those that are only associated with emotion – a procedure which is not attended to in previous compilations.

In the present paper, we describe our application and adaptation of Pavlenko's framework in the development of a Chinese language emotion corpus. While substantial computationally driven attempts at a Chinese emotion lexicon do exist (e.g. Dong & Dong 2006; Xu, Meng & Wang 2010), these remain insufficient for practical application and also represent different approaches to the present problem. The limitations of the various methods described thus far indicate that a proper framework for the identification and categorisation of Chinese emotion terms is required to facilitate empirically supported study and cross-linguistic comparison.

3. Method

3.1 *Procedures and classification guidelines*

The main objective was for a group of coders to extract emotion words from a representative Chinese dictionary. A system of counter coding was used where two coders coded the same list and any differences was resolved by consensus. To ensure consistency, we adhered to the counter coding practice commonly used in inter-judge agreement protocols employed in qualitative studies (see Syed & Nelson 2015 and Kovacs & Hill 2015 for detailed descriptions of the process). Words from the *Xiandai Hanyu Cidian* [Modern Chinese Dictionary] (Lu & Ding 2010) were entered into a framework for identifying emotion words, emotion-laden words and emotion-related words, as proposed by Pavlenko (2008). This framework is a revision of Pavlenko's classification of the emotion terms and is shown in Figure 1. Pavlenko's demarcations of emotion, emotion-related and emotion-laden words were used as the 'coding manual' to train the coders. The aim was for all coders to approach the data set with consistency.

The selected dictionary was divided into nine sections which were then each assigned to a pair of researchers for coding. A group of 18 native Mandarin Chinese speakers, comprising undergraduates and postgraduates at Nanyang Technological University, were recruited as Research Assistants (RAs) and trained to identify and code emotion words within their allocated word list. Each of the nine lists of words were coded independently by two RAs. Once the two RAs compiled and coded the list of emotion terms, non-agreements in the identification and coding processes were discussed and resolved together with the principal researcher. When the nine word lists of identified and coded emotion terms were completed, a final check was conducted on each list to assess their reliability and accuracy. As the purpose of the task was to correctly identify the categories of emotion words as stipulated by Pavenko (2008), it was critical that the rate of agreement between annotators had to be 100% accurate. Several meetings were held to discuss contentious examples which were resolved through consensus. The point was to achieve a judgement that is comparable across annotators. Hence, all examples which were annotated differently were jointly discussed with the principal researcher and resolved.

Additionally, all identified emotion words, emotion-laden words and emotion-related words were coded for valence (positive, negative, neutral) and part-of-speech (noun, verb, adjective, adverb, conjunction, mimetic, interjection). Frequency ranks were taken from the *Xiandai Hanyu Changyong Cibiao* (Common Lexicon in Modern Chinese, 2008), which includes data from over 50,000 Chinese terms. All words not found in the frequency ranking data were labelled accordingly.

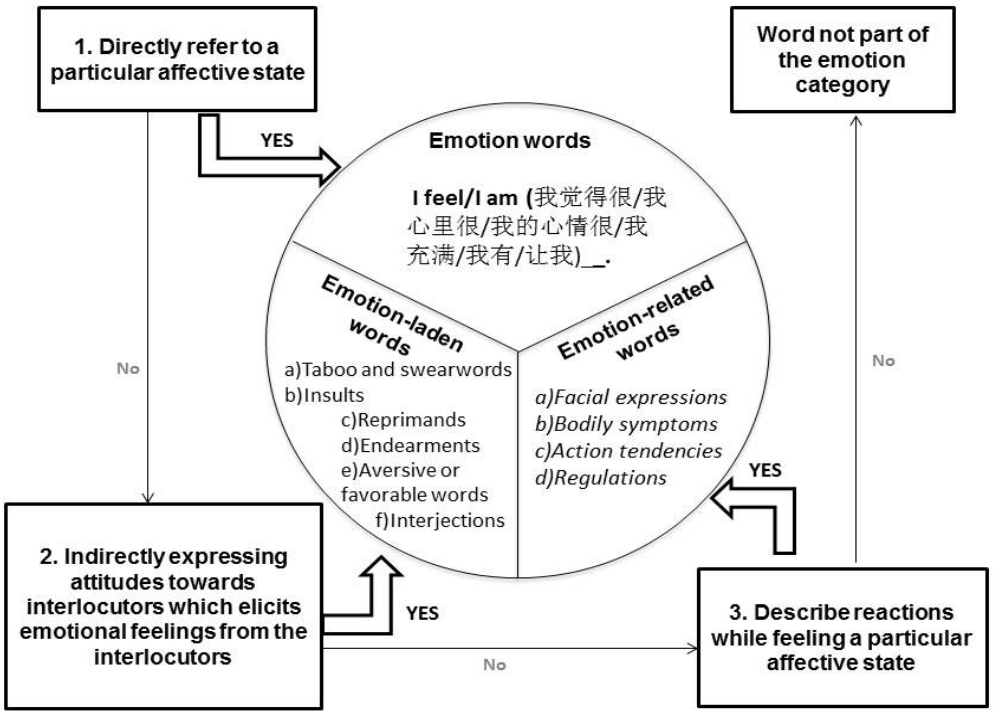

Figure 1. Framework for identifying emotion terms.

3.2. *Emotion words*

In the first phase, words were examined to check if indeed they were emotion words. In the revised framework, emotion words directly refer to a particular affective state (e.g. angry, scared) and fit in the syntactic context "I feel/I am X" (where "X" refers to the emotion word). However, as the linguistic structures in the Chinese language differ from English, several different expressions for "I feel/I am X" in Chinese were included to the framework. They are:

- "*wo jue de hen* X" (I feel X)
- "*wo xin li hen* X" (my heart feels X)
- "*wo de xin qing hen* X" (my mood is X)
- "*wo chong man* X"/ "*wo you* X" (I am full of X)
- "*rang wo* X" (it lets me feel X)

Emotion words were also identified with one or more of the basic emotion types described in the Geneva Emotion Wheel (GEW) shown in Figure 2 (Scherer 2005). The emotion categories in the GEW served as a guiding template for the consideration of emotion words membership. However, the RAs were aware that many emotions terms were not translatable and they only used the GEW as a reference.

Identified emotion words were then coded in several ways, including valence and part-of-speech. Additionally, specific term types were labelled accordingly:

- Complex emotions: emotion words which describe more than one emotional state, e.g. "*beichou*" 悲愁 (sad and worried)

- Emotion idioms: in Chinese, idioms usually consist of four characters, e.g. "*bu han er li*" 不寒而栗 (shuddering in fear)
- Emotion metaphors: emotion words that are expressed metaphorically, e.g. "*duanchang*" 断肠 (literally, broken intestines; heartbroken)
- Emotions in transition: words that describe emotions which are in a transitional state, e.g. "*jiechou*" 解愁 (to release worry/sadness)

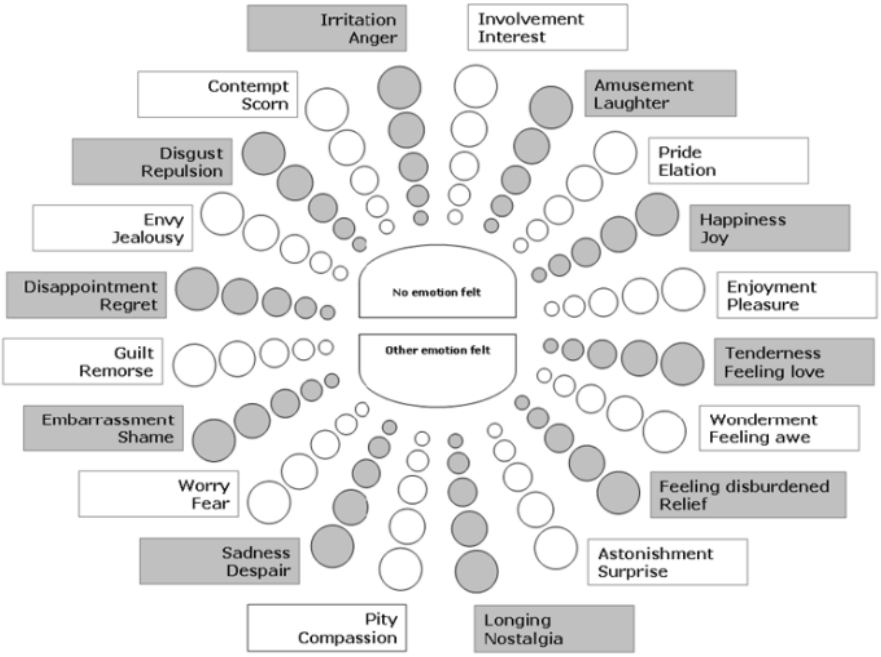

Figure 2. Geneva Emotion Wheel (Scherer 2005)

3.3 *Emotion-laden words*

If the target word was not an emotion word, it was then entered into the second phase of the framework where it was determined if it was an emotion-laden word. Emotion-laden words are used when speakers want to project their attitudes towards interlocutors, which might elicit emotional feelings from the interlocutors. The attitude and feelings elicited may be positive and negative. These words, which were labelled accordingly, include:

- expletives (e.g. *biaozi* 婊子'bitch')
- insults (e.g. *wangbadan* 王八蛋 'bastard')
- reprimands (e.g. *jiandian* 检点 'behave')
- endearments (e.g. *qin'aide* 亲爱的 "darling")
- valence-laden words (aversive, favourable; e.g. *qiangjian* 强奸 '*rape'*, *aizheng* 癌症 'cancer', *jiere* 节日 'festival'")
- interjections (e.g. *aiyo* 哎哟 'ouch')

3.4.1 Emotion-related words

If the word was neither an emotion word nor an emotion-laden word, it was entered into the third phase where it was determined whether it was an emotion-related word. Emotion-related

words describe reactions while feeling a particular affective state, with reference to one or more basic emotions depicted in the GEW or the previously identified Chinese emotion words. Emotion-related words were included in this category if they involved either facial and/or motor expressions (e.g. *cangbai* 苍白 'become pale/turn white'), physiological symptoms (e.g. *chandou* 颤抖 'shiver'), action tendencies (e.g. *taobi* 逃避 'escape and hide') or regulatory actions (e.g. *yinman* 隐瞒 "conceal"). Additionally, the following parameters help to identify emotion-related words when in doubt.

- Novelty: words which are related to evaluating a situation when it is sudden or unexpected (e.g. 'shock', 'surprise' or *turan* 突然 'sudden')
- Standard: words which describe adherence to or violation of rules and standards (e.g. *fanfa* 犯法 'against the law)
- Consequence: words which describe the positivity or negativity of outcomes (e.g. *yanzhong* 严重 "serious")
- Attribution: words which are related to judgments of the actor of an event or behaviour (e.g. *zize* 自责 "self-blame")

3.5. *Abstract Emotion terms*

Finally, abstract emotion terms were extracted from the remaining terms in the list. In Chinese, the majority of characters have no specific meaning and are rarely or never used singly. Chinese words are usually composed of two or more characters. For example, "*ai*" 哀 is associated with the meaning of sadness. However, "*ai*" 哀 is not commonly used singly to describe sadness; it is usually used in conjunction with other characters to form emotion terms such as "*aichou*" 哀愁 (sad, sorrowful) and "*aishang*" 哀伤 (grieving sorrow). Abstract emotion terms were treated as a separate category of not naturally used words.

Words that remained after the third phase and were not abstract emotion terms were classified as unrelated. Examples of excluded terms include words describing personality ("honest", "cunning") and courtesy expressions, such as *bugandang* 不敢当 "you flatter me".

4. Results

A total of 3766 words derived from Chinese characters in the dictionary were entered into the framework. These words were further sorted out into 953 emotion words, 978 emotion-related words, 1416 emotion-laden words, 52 abstract emotion terms, and 367 unrelated words (see Table 1). In this section, we describe only the valence and part-of-speech distributions for each of the three emotion term categories, Emotion words, Emotion-laden words, and Emotion-related words.

Finally, words without a frequency ranking (i.e., words that are so infrequent that they do not merit a frequency ranking) were omitted. When words without a frequency ranking were excluded from the three emotion term categories, 824 emotion words (28.47%), 855 emotion-related words (29.22%), and 1236 emotion-laden words remained (42.31%) (Figure 3).

Table 1. Number and percentage of emotion terms in the various categories of the corpus data, including and excluding infrequent words.

Category	Number of words	Percentage	Excluding infrequent words	Percentage
Emotion words	953	25.30%	824	25.30%
Emotion-laden	1416	37.60%	1236	37.90%
Emotion-related	978	26.00%	855	26.30%
Emotion abstract	52	1.40%	27	0.80%
Unrelated	367	9.70%	315	9.70%
Total	3766	100%	3257	100%

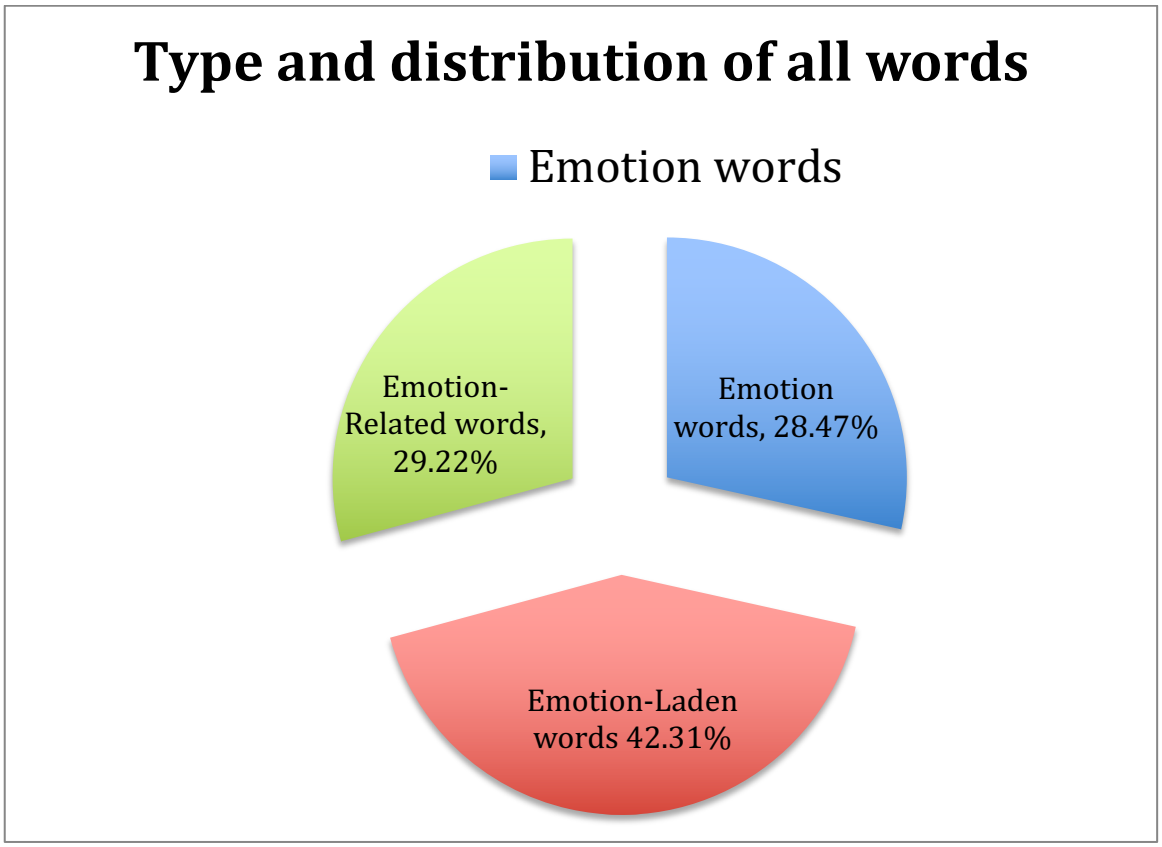

Figure 3 Type and distribution of emotion words

The following results also describe all the emotion words identified, regardless of availability of the frequency data obtained from the Common Lexicon in Modern Chinese (2008). Additionally, some words were tagged for more than one classification in the same category, where applicable. For example, the word "*anwei*" 安慰 in Chinese can act as both an adjective (comforting) and a verb (to comfort).

For the category of emotion words (see Table 2), 191 of the terms were classified into one of the four defined label types. The majority of the labelled words consisted of emotion idioms, (n=112) representing 11.75% of all emotion words. This was followed by complex emotion words which accounted for 4.41% of the data. As for valence, 62.01% of the emotion

words were tagged as negative emotion words, 33.89% as positive emotion words, 2.10% as neutral emotion words and 1.99% as emotion words that are emotions in transition and thus valence is not applicable. For parts-of-speech, the majority of emotion words were identified as verbs and adjectives with a number of 415 (43.55%) and 472 (49.53%) words respectively. On the other hand, nouns and adverbs made up only a smaller portion of the category, with a number of 92 (9.65%) and 9 (0.94%) words respectively. The valence distribution is shown in Figure 4.

Table 2. Number and percentage of emotion words classified by type, valence and part-of-speech (including and excluding infrequent words).

Category	Classification	Number of words	Percentage	Excluding infrequent words	Percentage
Type of emotion words[#]	Complex emotion	42	4.41%	31	3.76%
	Emotion idiom	112	11.75%	100	12.14%
	Emotion metaphor	21	2.20%	20	2.43%
	Emotion in transition	19	1.99%	19	2.31%
	N.A.	762	79.96%	656	79.61%
Valence	Positive	323	33.89%	280	33.98%
	Negative	591	62.01%	509	61.77%
	Neutral	20	2.10%	16	1.94%
	N.A.	19	1.99%	19	2.31%
Part-of-speech[#]	Verb	415	43.55%	372	44.42%
	Adjective	472	49.53%	402	48.79%
	Noun	92	9.65%	82	9.95%
	Adverb	9	.94%	8	.97%

[#]The total sum of words for the category exceeds the actual sum (n=953) as words may be tagged for more than one classification.

For the classification of emotion-laden words (Table 3), valence coding of emotion terms identified 75.14% with negative valence, 23.23% with positive valence and 1.62% having neutral valence.

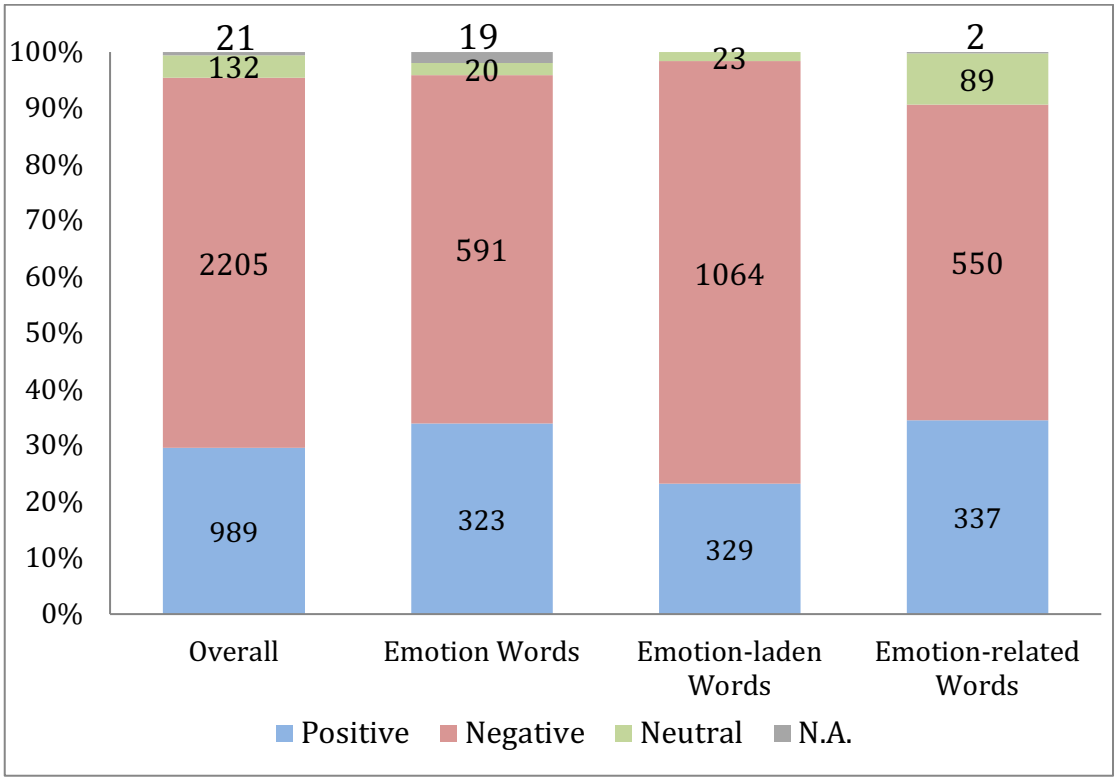

Figure 4 Valence distribution

Part-of speech coding of emotion-laden words identified a large number of words under three different classifications – 33.47% are verbs, 32.49% are adjectives, 35.59% are nouns. The remaining words consisted of interjections (e.g. "*ah*"), and adverbs. The distribution by part of speech is shown in Figure 5.

The decision of which part of speech each word belongs to is by no means, a simple issue. The method used in this study to resolve this issue was for the RAs in this study to make a judgement by doing a concordance of each word in the Lancaster Corpus of Mandarin Chinese (LCMC) (McEnery & Xiao 2004). We acknowledge the limitation of this approach and that at the same time future compilations could consider actual use in both written and spoken corpus. This is a highly involved task and is currently in progress in the second stage of the study.

Table 3. Number and percentage of emotion-laden words classified by type, valence and part-of-speech (including and excluding infrequent words).

Category	Classification	Number of words	Percentage	Excluding infrequent words	Percentage
Type of emotion-laden words	Expletive	1	.07%	1	.08%
	Insult	171	12.08%	151	12.22%
	Reprimand	4	.28%	4	.32%
	Endearment	15	1.06%	11	.89%
	Aversive/Favourable word	1203	84.96%	1053	85.19%
	Interjection	22	1.55%	16	1.29%
Valence	Positive	329	23.23%	288	23.30%
	Negative	1064	75.14%	927	75.00%
	Neutral	23	1.62%	21	1.70%
Part-of-speech[#]	Verb	474	33.47%	428	34.14%
	Adjective	460	32.49%	398	32.20%
	Noun	504	35.59%	442	35.76%
	Adverb	7	.49%	6	.49%
	Interjection	19	1.34%	13	1.05%

[#]The total sum of words for the category exceeds the actual sum (n=1416) as words may be tagged for more than one classification.

Valence of each word was evaluated by the team of 18 RAs. The words were not presented in context and the RAs were instructed to rate each word as positive, neutral or negative. A larger proportion of the emotion-related words (Table 4) was tagged as having negative valence (56.24%) as compared to positive valence (34.46%). For the identification of part-of-speech, 67.18% of the total number of emotion-related words were coded as verbs, 25.26% as adjectives, 6.65% as nouns, 2.66% as adverbs, 0.72% as mimetics (e.g. "he he"), and 0.1% as conjunctions. In comparison to the other categories, there are more verbs in emotion-related words and the number of verbs is close to the number of adjectives in emotion words.

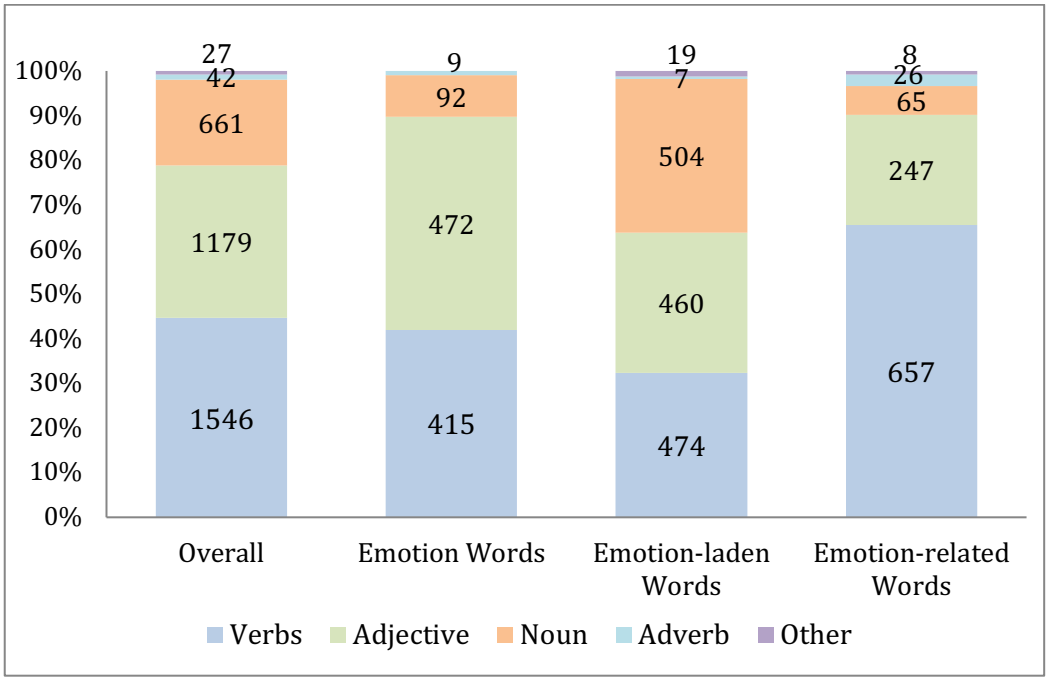

Figure 5 Part of speech Distribution

Table 4. Number and percentage of emotion-related words classified by type, valence and part-of-speech (including and excluding infrequent words).

Category	Classification	Number of words	Percentage	Excluding infrequent words	Percentage
Type of emotion-related words [#]	Novelty	4	.41%	3	.35%
	Standard	8	.82%	6	.70%
	Consequence	7	.72%	4	.47%
	Attribution	26	2.66%	22	2.57%
	N.A.	936	95.71%	822	96.14%
Valence	Positive	337	34.46%	300	35.09%
	Negative	550	56.24%	477	55.79%
	Neutral	89	9.10%	78	9.12%
	N.A.	2	.20%	-	-
Part-of-speech[#]	Verb	657	67.18%	580	67.84%
	Adjective	247	25.26%	214	25.03%
	Noun	65	6.65%	57	6.67%
	Adverb	26	2.66%	22	2.57%
	Conjunction	1	.10%	-	-
	Mimetic	7	.72%	4	.47%

[#]The total sum of words for the category exceeds the actual sum (n=978) as words may be tagged for more than one classification.

5. Discussion

The objective of this study was to compile and analyse a list of Chinese emotion terms from a Chinese lexicon in order to study the pattern of emotion expressions that could be found in the Mandarin Chinese language. This purpose was achieved by utilising a revised framework based on Pavlenko's proposal on the definition of the three emotion terms: emotion words, emotion-laden words and emotion-related words. In addition, the corpus data was also coded for specific emotion indicators, as well as valence and part-of-speech. Valence ratings revealed that the three emotion term categories vary significantly with regards to the distribution of different valence types, $\chi^2(6, N = 3347) = 201.1, p < .001$). This supports Pavlenko's caution that the three categories have different semantic profiles and should be treated separately, unlike dimensional approaches which do not distinguish between word type and typically only focus on intensity and valence.

Based on the results, there are several observations worth noting from the compiled list of emotion terms. Firstly, the present data appears to support the idea of a prevalence of verbs in the Chinese language, although comparable data for other languages would be required to fully support this claim. Generally, verbs are used extensively in Mandarin Chinese to encode emotion. Verbs make up the largest proportion overall but when you examine individual emotion categories, we see verbs predominantly in the emotion-related category only. In fact, studying the parts of speech distribution across emotion words categories, we see that these are distributed differently in each category. There were more adjectives in the emotion words category and an equal distribution of nouns, verbs and adjectives in the emotion-laden category. It has been suggested that different ways of talking about emotions give rise to different cultural schema, and also reflect different properties of their respective societies (Pavlenko 2002; Wierzbicka 1994, 2004). For instance, prevalence of adjectives when describing emotion states in English may be indicative of a more individualistic societal outlook which focuses more on internal states, while expression of emotion via verbs as in Russian or Polish may reflect a more collectivistic outlook which encodes emotions as personal or interpersonal processes or relations (Pavlenko 2008). This interpretation seems to agree with our present findings with Chinese, as it has been typically described as a more collectivistic culture. This needs to be corroborated by studies conducted in other collectivistic cultures.

The results also show that, unlike emotion words and emotion-related words, emotion-laden words contain the greatest number of nouns. This is also not unexpected as by the definition of emotion-laden words, they are words that can elicit emotions indirectly from the interlocutors and many nouns have aversive meanings that elicit feelings from others.

As for valence, there were more negative emotion terms coded than positive emotion terms across all the three emotion categories. Additionally, the number of neutral valence words was highest for emotion-related words. This could be due to the fact that this category captured the most terms whose valence is context specific. That is, the emotion these particular emotion-related words express can vary and thus the valence is dependent on the context. For example, "crying" is an emotion-related action that may be associated with the emotion "sadness", a negative valence word. However, the same action may also occur and be attributed to the emotion "joy", which has a positive valence. However, the participants in this study were

only given the restrictive options of 'positive', 'neutral' or 'negative' to rate each item. A more nuanced study allowing the participants to rate on a Likert scale will likely yield different results.

In summary, the present framework seems to provide a working and viable template for the identification of emotion words and terms as a method of understanding emotional expressions in a language. The value of an overall understanding or identity of valence (whether positive or negative) in an emotion corpus cannot be underestimated as demonstrated in the field of sentiment analysis. So far, a comprehensive list of Mandarin Chinese emotion words does not exist and this study represents a first step at unravelling the emotion language landscape of Mandarin Chinese speakers, the biggest speech community in the world. The list is potentially useful for researchers who need to develop stimuli for emotion experiments. Moreover, the classification of emotion expressions into the three emotion categories may also be a useful resource, as there is emerging evidence that they trigger different responses in experimental settings. Altarriba and Basnight-Brown (2011) found that emotion word type (emotion or emotion laden) moderates the Affective Simon Task. Also, adopting the same approach and same method of lexical selection will allow a more systematic way of comparing emotion lexicon across languages. In the past, this was not possible as each team of researchers used very different identification methods to compile their list. This current study represents the first effort in explicitly describing the process and methodology adopted. Currently, the team is also building the lexicon for Malay, Indonesian and English. When completed, it will facilitate a comparison of three typologically different languages (Mandarin Chinese, Malay/Indonesian and English) and two typologically similar languages (Malay and Indonesian). It is our hope that this crosslinguistic emotion lexicon will be useful for extending our understanding of emotion research across languages as well as help improve the accuracy of sentiment analysis.

Extending this framework to other languages bears a lot of potential for cross-linguistic and cross-cultural work. At this point, further research could be directed at identifying the categories of specific emotional states. Furthermore, valence and intensity ratings by a bigger group of participants would serve to enhance the validity of the data set.

Dr Bee Chin NG
Linguistics and Multilingual Studies
School of Humanities
Nanyang Technological University
Singapore
Email: mbcng@ntu.edu.sg

Ms Can CUI
Linguistics and Multilingual Studies
School of Humanities
Nanyang Technological University
Singapore
Email: cuic0003@e.ntu.edu.sg

Dr Francesco CAVALLARO (Corresponding Author)
Nanyang Technological University
Linguistics and Multilingual Studies
School of Humanities
48 Nanyang Avenue,
#03-83
Singapore 639818
Email: cfcavallaro@ntu.edu.sg

Acknowledgments

We wish to acknowledge the funding support for this project from Nanyang Technological University under the Undergraduate Research Experience on CAmpus (URECA) programme.

We are also grateful to the New Silk Road Grant (NSR M4080398.100) for supporting the Research Assistants used in this project.

We thank Liu Hong Yong (Post doctoral Fellow), Teh Hui Chian (Project officer) and April Ching (Project Officer) who all assisted with coordinating the study.

References

Altarriba, J. & D. M. Basnight-Brown. 2011. The representation of emotion vs. emotion-laden words in English and Spanish in the Affective Simon Task. *International Journal of Bilingualism,* 15(3). 310-328.

Bai, Lin 2015. *Emotion categorization and early bilinguals: A study of Chinese-English bilinguals' understanding of the shame category*. Ph.D. dissertation. Singapore: Nanyang Technological University.

Bradley, M. M. & P.J. Lang. 1999. Affective norms for English words (ANEW): Instruction manual and affective ratings. *Technical Report C-1*, The Center for Research in Psychophysiology, University of Florida.

Briesemeister, B. B., L. Kuchinke & A. M. Jacobs. 2011. Discrete emotion norms for nouns: Berlin affective word list (DENN–BAWL). *Behavior Research Methods*, 43(2). 441-448.

Chang, L., K. Chen & C. R. Huang. 2000. Alternation across semantic field: A study of Mandarin verbs of emotion. In Yung-O Biq (ed.), *Chinese Verbal Semantics, Special Issue, Computational Linguistics and Chinese Language Processing* 5(1). 61–80.

Dong, Z. & Q. Dong. 2006. *HowNet and the Computation of Meaning*. Singapore: World Scientific.

Eifring, H. 2004. Introduction: Emotions and the Conceptual History of Qing. In H. Eifting, (ed.), *Love and Emotions in Traditional Chinese Literature* (pp.1-36). Leiden ad London: Brill.

Eilola, T. M. & J. Havelka. 2010. Affective norms for 210 British English and Finnish nouns. *Behavior Research Methods*, 42(1). 134-140.

Ekman, P. 2003. Darwin, Deception, and Facial Expression. In P. Ekman, J. J. Campos, R. J. Davidson & F. B. M. de Waal, (eds.), *Emotions inside out: 130 years after Darwin's: The expression of the emotions in man and animals* (pp. 205-221). New York, NY: New York Academy of Sciences.

Enfield, N. J. & A. Wierzbicka. 2002. Introduction: The body in description of emotion. *Pragmatics and Cognition,* 10(1). 1-24.

Ferré, P., M. Guasch, N. Martinez-Gracia & I. Fraga. 2017. Moved by words: Affective ratings for a set of 2,266 Spanish words in five discrete emotion categories. *Behavior Research Methods*, 49(3). 1082-1094.

Ferré, P., M. Guasch, C. Moldovan & R. Sánchez-Casas. 2012. Affective norms for 380 Spanish words belonging to three different semantic categories. *Behavior Research Methods*, 44(2). 395-403.

Fischer, A. H., P. M. Rodriguez Mosquera, A. E. M van Vianen & A. S. R. Manstead. 2004. Gender and Culture Differences in Emotion. *Emotion*, 4(1). 87-94.

Frijda, N. H. & B. Mesquita. 1998. The analysis of emotions. What develops in emotional development. In M. F. Mascolo, & S. Griffin (eds.), *What develops in emotional development?* (pp. 273-295). New York: Plenum Publishing Corporation.

Hamid, T., N. Doost, A. R. Moradi, M. R. Taghavi, W. Yule & T. Dalgleish. 1999. The development of a corpus of emotional words produced by children and adolescents. *Personality and Individual Differences*, 27(3). 433-451.

Howell, S. 1984. *Society and Cosmos: Chewong of Peninsular Malaysia*, Oxford: Oxford University Press.

Jack, R. E., O. G. B. Garrod & P. G. Schyns. 2014. Dynamic facial expressions of emotion transmit an evolving hierarchy of signals over time. *Current Biology*, 24(2). 187-192.

John, O. P., A. Angleitner & F. Ostendorf. 1988. The lexical approach to personality: A historical review of trait taxonomic research. *European journal of Personality*, 2(3). 171-203.

Johnson-Laird, P. N. & K. Oatley. 1989. The language of emotions: An analysis of a semantic field. *Cognition and emotion*, 3(2). 81-123.

Just, M. A., L. Pan, V. L. Cherkassky, D. L. McMakin, C. Cha, M. K. Nock & D. Brent. 2017. Machine learning of neural representations of suicide and emotion concepts identifies suicidal youth. *Nature Human Behaviour*, 1. 911–919.

Kitayama, S., B. Mesquita & M. Karasawa. 2006. Cultural affordances and emotional experience: Socially engaging and disengaging emotions in Japan and the United States. *Journal of Personality and Social Psychology*, 91(5). 890-903.

Kovacs, T. & K. Hill. 2015. A tutorial on reliability testing in AAC language sample transcription and analysis. *Augmentative and Alternative Communication*, *31*(2). 148-158.

Kringelbach, M. L. & K. C. Berridge. 2017. The affective core of emotion: Linking pleasure, subjective well-being, and optimal metastability in the brain. *Emotion Review,* 9(3). 191–199.

Lee, S. Y. M. 2010. *A linguistic approach to emotion detection and classification*. Ph.D. dissertation. Hong Kong: The Hong Kong Polytechnic University.

Li, J., L. Wang & K. Fischer. 2004. The organization of Chinese shame concepts. *Cognition and Emotion*, 18(6). 767–797.

Lin, J. & Y. Yao. 2016. Encoding emotion in Chinese: A database of Chinese emotion words with information of emotion type, intensity, and valence. *Lingua Sinica*, 2(1). Article 6 (No page numbers).

Lomas, T. 2016. Towards a positive cross-cultural lexicography: Enriching our emotional landscape through 216 'untranslatable' words pertaining to well-being. *Journal of Positive Psychology***,** 11/5. 546-558.

Lu, S. & Ding, S. (eds.), 2010. *Xiandai Hanyu Cidian (Modern Chinese Dictionary*), revised edition. Beijing: The Commercial Press.

Markus, H. R. & S. Kitayama. 1991. Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review,* 98(2). 224-253.

McEnery, T. & R. Xiao. 2004. *The Lancaster Corpus of Mandarin Chinese*. Lancaster University.

Montefinese, M., E. Ambrosini, B. Fairfield & N. Mammarella. 2014. The adaptation of the affective norms for English words (ANEW) for Italian. *Behavior Research Methods,* 46(3). 887–903.

Ogarkova, A. 2007. Green-eyed monsters: a corpus-based study of the concepts of ENVY and JEALOUSY in modern English. *metaphorik.de* 13. 87-147

Ortony, A., G. L. Clore & M. A. Foss. 1987. The referential structure of the affective lexicon. *Cognitive Science*, 11(3). 341-364.

Pavlenko, A. 2002. Emotions and the body in Russian and English. *Pragmatics & Cognition*, 10(1). 207-241.

Pavlenko, A. 2008. Emotion and emotion-laden words in the bilingual lexicon. *Bilingualism: Language and Cognition*, 11(02). 147-164.

Pavlenko, A. 2014. *The bilingual mind and what it tells us about language and thought*. Cambridge: Cambridge University Press

Peter, M., F. Chang, J. Pine, R. Blything & C. Rowland. 2015. When and how do children develop knowledge of verb argument structure? Evidence from verb bias effects in a structural priming task. *Journal of Memory and language*, 81. 1-15.

Redondo, J., I. Fraga, I. Padrón & M. Comesaña. 2007. The Spanish adaptation of ANEW (affective norms for English words). *Behavior Research Methods*, 39(3). 600-605.

Riegel M., M. Wierzba, M. Wypych, Ł Żurawski, K. Jednoróg & A. Grabowska. 2015. Nencki affective word list (NAWL): The cultural adaptation of the Berlin affective word list-reloaded (BAWL-R) for Polish. *Behavior Research Methods*, 47(4). 1222–1236.

Russell, J. A. 1991. Culture and the categorization of emotion. *Psychological Bulletin*, 110. 426-450.

Russell, J. A. & M. S. M. Yik. 1996. Emotion among the Chinese. In M. H. Bond (ed.), *The handbook of Chinese psychology* (pp. 166-188). New York: Oxford University Press.

Saucier, G. & L. R. Goldberg. 1996. The language of personality: Lexical perspectives on the five-factor model. *The five-factor model of personality: Theoretical perspectives*, 21-50.

Scherer, K. R. 2005. What are emotions? And how can they be measured? *Social Science Information*, 44(4), 693-727.

Scherer, K. R. 2013. Measuring the meaning of emotion words: A domain-specific componential approach. In Johnny J. R. Fontaine, Klaus R. Scherer & Cristina Soriano (eds.), *Components of emotional meanings. A sourcebook* (pp. 7-30). Oxford : Oxford University Press.

Scherer, K. R. & A. Moors. (in press) 2019. The emotion process: Event appraisal and component differentiation. *Annual Review of Psychology.* 70(1). Review in Advance first posted online on August 15, 2018. https://doi.org/10.1146/annurev-psych-122216-011854

Scherer, K. R., A. Schorr & T. Johnstone. 2001. *Appraisal Processes in Emotion: Theory, Methods, Research.* New York: Oxford University Press.

Scherer, K.R., V. Shuman, J. R. J. Fontaine & C. Soriano. 2013. The GRID meets the Wheel: Assessing emotional feeling via self-report. In J. R. J. Fontaine, K. R. Scherer & C. Soriano (eds.), *Components of emotional meaning: A sourcebook* (pp. 281-298). Oxford: Oxford University Press.

Schmidtke, D. S., T. Schröder, A. M. Jacobs & M. Conrad. 2014. ANGST: Affective norms for German sentiment terms, derived from the affective norms for English words. *Behavior Research Methods*, 46(4). 1108-1118.

Siegel J. S. 2018. *Demographic and Socioeconomic Basis of Ethnolinguistics*. Cham, Switzerland: Springer.

Simons, G. F. & C. D. Fennig, (eds.). 2018. *Ethnologue: Languages of the World, Twenty-first edition.* Dallas, Texas: SIL International. Online http://www.ethnologue.com.

Soares, A. P., M. Comesaña, A. P. Pinheiro, A. Simões & C. S. Frade. 2012. The adaptation of the affective norms for English words (ANEW) for European Portuguese. *Behavior Research Method*s, (44)1. 256-269.

Soueif, Ahdaf. 1999. *The Map of Love*. New York: Bloomsbury Publishing.

Stevenson, R. A., J. A. Mikels & T. W. James. 2007. Characterization of the Affective Norms for English Words by discrete emotional categories. *Behavior Research Methods*, 39(4). 1020-1024.

Syed, M. & S. C. Nelson. 2015. Guidelines for establishing reliability when coding narrative data. *Emerging Adulthood*, *3*(6). 375-387.

Thothathiri, M., D. G. Evans & S. Poudel. 2017. Verb bias and verb-specific competition effects on sentence production. *PLoS One* 12(7). 1-18.

Vinodhini G. & R. M. Chandrasekaran. 2012. Sentiment analysis and opinion mining: A survey. *International Journal of Advanced Research in Computer Science and Software Engineering,* 2(6). 282-292.

Võ, M. L.-H., M. Conrad, K. Urton, M. J. Hofman & A. M. Jacobs. 2009. The Berlin affective word list. *Behavior Research Methods*, 41(2). 534-538.

Võ, M. L.-H., A. M. Jacobs & M. Conrad. 2006. Cross-validating the Berlin affective word list. *Behavior Research Methods*, 38(4). 606–609.

Wallace, A. & M. Carson. 1973. Sharing and diversity in emotion terminology. *Ethos*, 1(1). 1-29.

Whissell, C. 2008. A comparison of two lists providing emotional norms for English words (ANEW and the DAL). *Psychological Reports*, 102(2). 597-600.

Wierzbicka A. 1999. Emotional universals. *Language Design,* 2. 23-69.

Wierzbicka, A. 1994. Emotion, language, and cultural scripts. In S. Kitayama & H. Markus (eds.), *Emotion and culture: Empirical studies of mutual influence* (pp. 133–196). Washington, DC: American Psychological Association.

Wierzbicka, A. 2004. Preface: Bilingual lives, bilingual experience. *Special issue of the Journal of Multilingual and Multicultural Development*, 25 (2/3). 94–104.

Xu, G., X. Meng & H. Wang. 2010. Build Chinese emotion lexicons using a graph-based algorithm and multiple resources. In Aravind K. Joshi, Chu-Ren Huang & Dan Jurafsky (eds.), *Proceedings of the 23rd International Conference on Computational Linguistics* (pp. 1209-1217). Stroudsburg, PA, USA: Association for Computational Linguistics.

Xu, X. & J. Tao. 2003. The study of affective word categorization in Chinese 汉语情感系统中情感划分的研究. In *Proceedings of the 1st Chinese Conference on Affective Computing and Intelligent Interaction*. Accessed on 6 October 2018 at http://speakit.cn/Group/file/EmotionWord_ACII03.pdf

Zammuner, V. L. 1998. Concepts of emotion:" Emotionness," and dimensional ratings of Italian emotion words. *Cognition and Emotion*. 12(2). 243-272.