NOTES/COMMENTS/REPLY MANUSCRIPT

Affective and non-affective psycholinguistic norms for 500 Chinese three-character idiomatic expressions

Yanlu Zhong[1] · Yimei Shao[2] · Wei Yi[3]

Abstract

Multiword expressions, especially idiomatic expressions, convey rich emotional information. To explore the relationship between affective and non-affective psycholinguistic dimensions in the processing of idiomatic expressions, we collected normative data for two affective (i.e., valence and arousal) and four non-affective variables (i.e., familiarity, transparency, ambiguity, age of acquisition) for 500 Chinese three-character idiomatic expressions, based on responses from 418 native Chinese speakers. Our results showed a quadratic U-shaped relationship between valence and arousal, with neutral idiomatic expressions rated as less arousing than negative or positive ones. Furthermore, more transparent idiomatic expressions were more strongly valenced (either negative or positive) and more arousing. Idiomatic expressions more often used figuratively evoked more negative or stronger emotional responses. In addition, idiomatic expressions with greater negativity were acquired earlier but occurred less frequently. Regarding the non-affective variables, more transparent idiomatic expressions were perceived as more familiar and acquired earlier. Idiomatic expressions more often used figuratively, acquired earlier, or occurred more frequently were rated as more familiar. Lastly, idiomatic expressions that were more often used literally were more transparent and frequent. This study provides a large-scale database for examining the impact of affective and non-affective psycholinguistic factors on Chinese idiomatic expression processing, and it sheds light on the interrelated cognitive mechanisms underlying figurative language processing and emotion.

Keywords Emotion · Idiomatic expression · Valence · Arousal · Non-affective variables

Introduction

According to usage-based approaches (e.g., Ellis et al., 2013; Yi et al., 2017), language acquisition emerges from exposure to linguistic input and the accumulation of language experience. Language consists of linguistic exemplars of various sizes, including individual words and multiword expressions (MWEs; Ellis et al., 2013). MWEs refer to combinations of words that co-occur more frequently than would be expected by chance. They include linguistic units with diverse semantic and syntactic properties, such as collocations (*black coffee*), binomials (*salt and pepper*), idiomatic expressions (*over the moon*), and lexical bundles (*is one of the*; Siyanova-Chanturia & Martinez, 2015).

Among various types of MWEs, idiomatic expressions are particularly challenging for second-language speakers due to their semantic figurativeness (Yi & Zhong, 2024). Idiomatic expressions usually have both figurative and literal interpretations. The figurative meaning of an idiomatic expression (e.g., *spill the beans*, meaning: to reveal a secret) cannot be directly derived from the literal meaning of its constituent words but must be retrieved from long-term memory (Siyanova-Chanturia & Martinez, 2015). Previous studies have found that idiomatic expressions are more commonly used than literal ones to express emotions (Drew & Holt, 1988). Emotions can be categorized into two bipolar dimensions, namely, valence and arousal. Valence describes the degree to which an emotion is pleasant or unpleasant, while arousal refers to the level of physiological activation caused by a stimulus, ranging from calm to excited (Lang et al., 1998). Many studies have collected norms for

✉ Wei Yi
   weiyisla@pku.edu.cn

1   University of California, Santa Barbara, CA, USA

2   The Hong Kong Polytechnic University, Kowloon, Hong Kong SAR

3   School of Chinese as a Second Language, Peking University, Haidian District, No. 5 Yiheyuan Road, Beijing 100871, China

affective and non-affective properties of single words (e.g., Yao et al., 2017), yet little has been done for MWEs, particularly idiomatic expressions. Most norming studies on idiomatic expressions (e.gBonin et al., 2013; Bulkes & Tanner, 2017; Li et al., 2016) have focused on non-affective psycholinguistic variables, with only a few providing norms for both affective and non-affective variables of idiomatic expressions (Citron et al., 2016a, 2016b, for German; Gavilán et al., 2021, for Spanish; Morid & Sabourin, 2024, for English). Chinese is the most widely spoken non-alphabetic language in the world, and three-character and four-character idiomatic expressions are the key components of Chinese. Only Li et al. (2016) provided norms for non-affective variables of Chinese four-character idiomatic expressions, and the affective variables were not examined. Chinese three-character idiomatic expressions, such as "铁饭碗" (tiě fàn wǎn, literal meaning: iron rice bowl, figurative meaning: job stability; positive valence), "结对子" (jié duì zi, literal meaning: forming a pair, figurative meaning: collaborating to achieve a goal, neutral valence), and "拖油瓶" (tuō yóu píng, literal meaning: dragging an oil bottle; figurative meaning: a burden, negative valence), play a vital role in communication by conveying complex semantic, syntactic, and emotional information. These three-character idiomatic expressions have distinctive cultural connotations and regional characteristics, and they are more colloquial compared to their four-character counterparts (Li et al., 2016; Zang et al., 2021). However, current knowledge of Chinese three-character idiomatic expressions remains more limited compared to four-character idiomatic expressions.

Previous norming research on idiomatic expressions, although primarily focused on non-affective features and conducted in languages other than Chinese, provides a foundation for examining both the affective and non-affective features of Chinese three-character idiomatic expressions and their interrelations. Relevant studies are reviewed below.

Emotion and idiomatic expressions

Language acquisition involves more than just encoding lexical-semantic information in the mental lexicon. Instead, it is also shaped by language users' emotional experiences in real life. Both single words and MWEs, such as idiomatic expressions, are effective tools for perceiving and expressing human emotions (see Citron et al., 2016a, 2016b, for a review).

Affective norming studies on single words have been conducted in a wide range of languages, such as English (Citron et al., 2014; Warriner et al., 2013), German (Briesemeister et al., 2011), Italian (Della et al., 2010), and Chinese (Chan & Tse, 2024; Ho et al., 2015; Lv et al., 2024; Xu et al., 2022; Yao et al., 2017; Yi et al., 2025). To the best of our knowledge, only a few studies (Citron et al., 2016a, 2016b; Gavilán et al., 2021; Morid & Sabourin, 2024) have collected both affective and non-affective psycholinguistic norms for idiomatic expressions. A large body of research has found a typical quadratic U-shaped relationship between valence and arousal, where words (Chan & Tse, 2024; Ho et al., 2015; Lv et al., 2024; Xu et al., 2022; Yao et al., 2017) or idiomatic expressions (e.gCitron et al., 2016a, 2016b; Gavilán et al., 2021; Morid & Sabourin, 2024) rated as either very positive or very negative generally have higher arousal ratings compared to neutral ones. This U-shaped distribution usually includes a strong negative linear correlation between valence and arousal for negative stimuli, which results in an overall negative linear correlation (e.g., Citron et al., 2016a, 2016b; Ho et al., 2015; Xu et al., 2022; Yao et al., 2017). However, some studies have not observed this overall negative relationship between valence and arousal in idiomatic expressions (e.g., Gavilán et al., 2021) or words (e.g., Warriner et al., 2013 for English; Eilola & Havelka, 2010 for Finnish).

Apart from norming studies, a substantial body of psycholinguistic research has shown that valence and arousal significantly influence language processing across various cognitive tasks. The processing advantage of positive stimuli over neutral stimuli has been well documented in online processing studies (e.g., Citron et al., 2014; Kousta et al., 2009). However, there remains debate about whether negative stimuli differ in processing compared to neutral and positive stimuli. The automatic vigilance model of emotion (Estes & Verges, 2008) posits that positive and negative stimuli engage distinct cognitive mechanisms. Negative stimuli are perceived as threats, and withdrawal from them is more critical for survival than approaching pleasant or neutral stimuli. As a result, negative stimuli lead to delayed attentional disengagement and slower processing compared to neutral and positive stimuli (Estes & Verges, 2008; Kousta et al., 2009). In contrast, the motivated attention and affective states model (Lang et al., 1998) argues that attention is captured by emotionally significant stimuli, regardless of polarity, and is sustained more than neutral stimuli to facilitate rapid behavioral responses. Thus, both negative and positive stimuli share similar cognitive mechanisms and are processed faster than neutral ones. The role of arousal in language processing is also debated. Estes and Adelman (2008) found that exciting words are recognized faster than calming ones, while Kousta et al. (2009) found no effect of arousal on lexical decision latencies when controlling for valence. These discrepancies may stem from the interaction between arousal and valence. The "approach–withdrawal" model of emotion processing (Robinson et al., 2004) indicates that low-arousal stimuli, similar to positive stimuli, are associated with a sense of safety and elicit a consistent "approach" tendency, whereas high-arousal and negative stimuli are

related to survival threats and elicit a consistent "withdrawal" tendency. Therefore, positive high-arousal and negative low-arousal stimuli evoke conflicting approach–withdrawal tendencies, resulting in slower processing compared to positive low-arousal and negative high-arousal stimuli. Citron et al., (2016a, 2016b) presented positive and negative words with varying arousal levels and asked participants to decide whether to approach or withdraw from the stimuli. Their results, consistent with the "approach–withdrawal" model, showed longer response times for positive high-arousal and negative low-arousal words compared to positive low-arousal and negative high-arousal words. Likewise, Larsen et al. (2008) found that in low-arousal conditions, participants responded more slowly to negative words than to positive ones in lexical decision tasks. Although previous research has shown that valence and arousal affect language processing, these affective effects might actually arise from interference from other non-affective features. For example, Kousta et al. (2009) found no effect of valence after controlling for non-affective variables. Similarly, Larsen et al. (2006) evaluated the lexical characteristics of 1033 words used in 32 published emotional Stroop studies and discovered that emotional words were less frequent, were shorter in length, and had fewer orthographic neighbors than neutral words. These results suggest that the observed affective effects in language processing may be partly due to lexical differences between emotional and neutral stimuli. Therefore, it is necessary to further explore the non-affective features of idiomatic expressions.

Non-affective characteristics of idiomatic expressions

Idiomatic expression processing is moderated by various non-affective properties, such as familiarity (Siyanova-Chanturia et al., 2011), transparency (Bulkes & Tanner, 2017), and ambiguity (Cieślicka & Heredia, 2011). This study, in addition to examining valence and arousal, focuses on non-affective factors of Chinese three-character idiomatic expressions, including familiarity, ambiguity, transparency, age of acquisition (AoA), and frequency.

Familiarity is a key variable in psycholinguistic research and has been extensively explored (e.g., Siyanova-Chanturia et al., 2011). Many studies assess familiarity of idiomatic expressions to either control for or examine its effects (e.g., Siyanova-Chanturia et al., 2011). Research has shown that highly familiar idiomatic expressions are processed more quickly and accurately than unfamiliar ones (Gyllstad & Wolter, 2016; Libben & Titone, 2008). Familiarity is usually defined in two ways in norming studies. First, it may refer to subjective intuition about the occurrence of a linguistic unit, which reflects how often individuals feel they hear or see a linguistic unit (Citron et al., 2016a, 2016b). Second, familiarity can be defined by a direct explanation of familiarity, which refers to the degree to which individuals feel familiar with a unit (Yao et al., 2017). While the second definition provides a more straightforward description of familiarity, it may result in varying interpretations of familiarity across participants (Yao et al., 2017). Regarding the relationship between familiarity and emotionality, Citron et al., (2016a, 2016b) found that more positive and arousing idiomatic expressions tended to exhibit greater familiarity.

Ambiguity refers to the extent to which an idiomatic expression can be interpreted literally (Cronk et al., 1993). Many idiomatic expressions have both figurative and plausible literal usage. For example, the expression "穿小鞋" (chuān xiǎo xié) is used figuratively to indicate "to deliberately make things difficult for someone" and literally to describe "wearing small shoes". Libben and Titone (2008) showed that idiomatic expressions with lower ambiguity are processed more quickly and accurately than those with higher ambiguity. In norming studies of idiomatic ambiguity, participants are often asked to rate how often idiomatic expressions are used literally (e.g., Bonin et al., 2013; Tabossi et al., 2011). Some studies adopt a binary classification of idiomatic expressions as either ambiguous or unambiguous, based on whether they have semantically plausible literal usage (e.g., Citron et al., 2016a, 2016b). Previous studies have found that more familiar idiomatic expressions are more likely to be interpreted literally (Bonin et al., 2013; Cronk et al., 1993; Li et al., 2016). Moreover, idiomatic expressions with more positive connotations in German (Citron et al., 2016a, 2016b) and Spanish (Gavilán et al., 2021) have been found to be more frequently used in their literal sense.

Transparency, also referred to as compositionality (e.g., Bonin et al., 2013; Li et al., 2016), describes how easily the figurative meaning of an idiomatic expression can be inferred from the literal meanings of its components. More transparent idiomatic expressions were found to be easier to process in semantic judgment tasks (Gyllstad & Wolter, 2016) and eye-tracking reading tasks (Cieślicka & Heredia, 2017). Transparency norming research has been conducted in German (Citron et al., 2016a, 2016b), French (Bonin et al., 2013), English (Bulkes & Tanner, 2017), Spanish (Gavilán et al., 2021), and Chinese four-character idioms (Li et al., 2016). These studies found that more transparent idiomatic expressions were rated as more familiar (Bonin et al., 2013; Bulkes & Tanner, 2017; Li et al., 2016; Tabossi et al., 2011), more arousing (Citron et al., 2016a, 2016b), or more positive (Citron et al., 2016a, 2016b; Gavilán et al., 2021). However, the relationship between transparency and ambiguity has yielded mixed results. While some studies have found no significant correlation between transparency and ambiguity (e.g., Tabossi et al., 2011), other research

has shown that more transparent idiomatic expressions are either associated with more plausible literal meanings and more frequent literal usage (Bulkes & Tanner, 2017; Gavilán et al., 2021; Li et al., 2016), or are used figuratively (Libben & Titone, 2008).

AoA refers to the age at which a language learner first acquires the figurative meaning of an idiomatic expression. It has been shown that words learned early in life are processed more quickly and accurately than those acquired later, as evidenced by studies in picture naming (e.g., de Zubicaray et al., 2012) and word recognition tasks (e.g., Wilson et al., 2012). For idiomatic expressions, the acquisition of figurative meanings does not always align with the acquisition of literal meanings. Children typically grasp transparent phrases by age 5 and develop the ability to understand idiomatic figurative meanings through context by ages 7–8 (Cain et al., 2009). Additionally, the AoA of idiomatic expressions largely depends on communication contexts (Cain et al., 2005). Previous norming studies have found that idiomatic expressions acquired earlier are more familiar to language learners (Bonin et al., 2013; Li et al., 2016; Libben & Titone, 2008; Tabossi et al., 2011), more often used literally (Li et al., 2016), and more transparent (Bonin et al., 2013; Li et al., 2016).

Frequency refers to the number of occurrences of a linguistic unit in a language. Frequency is typically obtained from corpora. Psycholinguistic studies have widely established frequency effects, showing that more frequent linguistic units are processed more quickly and accurately than less frequent ones (e.g., Yi, 2018). Frequency and familiarity are correlated but distinct concepts. Some research suggests that familiarity is a more effective measure for lexical processing than frequency (Gernsbacher, 1984). However, other studies (e.g., Brysbaert & Cortese, 2011) have shown that frequency better predicts language speakers' behavior in language processing than familiarity. To further understand how frequency relates to familiarity, Tanaka-Ishii and Terada (2011) investigated the relationship between frequency and language speakers' feelings about how familiar they were with English and Japanese words. Later, Siyanova-Chanturia and Spina (2015) and Yi et al. (2023) collected subjective intuitions about the occurrences of MWEs in English, which are often used as a measure of participants' familiarity with MWEs. All these studies found a significant correlation between the frequency and familiarity. In addition, Bonin et al. (2013) found that more frequent idiomatic expressions were acquired earlier.

Overall, descriptive norms for the aforementioned variables of idiomatic expressions have been provided in languages such as English (Bulkes & Tanner, 2017; Libben & Titone, 2008), French (Bonin et al., 2013), Bulgarian (Nordmann & Jambazova, 2017), and German (Citron et al., 2016a, 2016b). These studies suggest that the affective and non-affective psycholinguistic variables of idiomatic expressions are interrelated factors that may jointly influence idiomatic expression processing. In addition, existing research on idiomatic expressions supports usage-based theories (e.g., Ellis et al., 2013), showing that exposure to idiomatic expressions in language experience (e.g., frequency, AoA) enhances language learners' mental representation (e.g., familiarity, knowledge) of these expressions (Bonin et al., 2013). Previous studies have also examined whether word-level properties can be activated and influence phrasal-level processing of idiomatic expressions by investigating the effects of transparency and ambiguity, thereby deepening the understanding of the dual-route model of idiomatic expression processing (Carrol & Conklin, 2014).

However, prior databases mainly focused on non-affective properties of idiomatic expressions, and only a small number of studies (Citron et al., 2016a, 2016b; Gavilán et al., 2021; Morid & Sabourin, 2024) have examined the relationship between affective and non-affective factors. For Chinese, a widely spoken non-alphabetic language, only Li et al. (2016) provided non-affective norming for four-character idiomatic expressions, without examining affective variables. Furthermore, no attention has been given to norms for Chinese three-character idiomatic expressions. To address these gaps, the present study collected ratings for affective variables (valence, arousal) and non-affective variables (familiarity, imaginability, ambiguity, semantic transparency, and AoA) for a large set of Chinese three-character idiomatic expressions. We also obtained the frequency of these idiomatic expressions from a Chinese corpus. The relationships among these affective and non-affective psycholinguistic variables were examined. Our dataset will be a valuable resource for research on the processing of Chinese idiomatic expressions. This study will also highlight the irreplaceable role of figurative language in emotional expression in Chinese, and shed light on Chinese language teaching. In addition, it will provide insights into the cognitive mechanisms underlying the interaction between emotional experiences and linguistic knowledge in the use of figurative language.

Method

Participants

We recruited 418 participants (209 male) from various universities and colleges in mainland China. Their ages ranged from 17 to 41 years ($M = 22.68$, $SD = 3.45$). They were all native speakers of Mandarin Chinese. All participants reported themselves to be mentally and physically healthy.

Materials

The following steps were taken to select 500 idiomatic expressions: First, we compiled a dataset of 7013 idiomatic expressions from one of the most widely used Chinese idiomatic expression dictionaries (Huang, 2009). These expressions were then cross-checked by three native Chinese speakers. We removed idiomatic expressions that fall under any of the following conditions: (1) idiomatic expressions whose figurative meanings are difficult to activate or comprehend or are rarely used in daily communication among native speakers (e.g., "帮八股", bāng bā gǔ; figurative meaning: supporting rigid and outdated rules or methods); (2) three-character phrases whose overall meaning could be entirely inferred from their constituent words (e.g., "按手印", àn shǒu yìn; figurative meaning: to press a handprint); (3) idiomatic expressions that are part of larger linguistic constructions and could not stand alone—for instance, "壁上观" (bì shàng guān) is part of "作壁上观" (zuò bì shàng guān) (figurative meaning: being an onlooker); (4) idiomatic expressions with similar forms and meanings to other idiomatic expressions but less frequently used, based on frequency data from the Center for Chinese Linguistics PKU (CCL) corpus—for instance, "跑江湖" (pǎo jiāng hú) shares the same figurative meaning (i.e., making a living through various skills) as "走江湖"(zǒu jiāng hú), but"跑江湖" was used less frequently than "走江湖" and was thus excluded. Following these criteria, 500 of the most common Chinese idiomatic expressions were selected. These idiomatic expressions include the following structural categories: verb-noun structure ($N = 260$, 52.0%), adjective-noun structure ($N = 208$, 41.6%), adverb-verb structure ($N = 23$, 4.6%), noun-verb structure ($N = 6$, 1.2%), and verb-particle-adjective structure ($N = 2$, 0.6%).

Alongside the 500 idiomatic expressions, three calibrator idiomatic expressions were selected for each variable (i.e., valence, arousal, familiarity, ambiguity, transparency, AoA). Ten native Chinese speakers, who did not participate in the formal study, rated the calibrators on a nine-point scale.[1] These calibrators were used in the instructions at the beginning of the questionnaires to provide participants with a consistent sense of the rating scale range.

[1] The calibrator idiomatic expressions for the respective variable were as follows (in increasing order of ratings): valence: "丑八怪" (1.0), "斑马线"(4.9), "千里眼" (8.6); arousal: "定风波" (1.4), "斑马线" (5.2), "下地狱" (8.7); familiarity: "执牛耳" (1.1), "千里眼"(5.3), "丑八怪" (8.7); ambiguity: "执牛耳" (1.1), "下地狱" (5.1), "千里眼" (8.0); transparency: "执牛耳" (1.0), "下地狱" (5.3), "丑八怪" (8.6); AoA: "丑八怪" (5.0), "定风波" (11.9), "执牛耳" (20.9)

Procedure

Online questionnaires were generated using Qualtrics (www.qualtrics.com). The 500 idiomatic expressions were randomly divided into 10 lists, each containing 50 idiomatic expressions. For each list, six variables—valence, arousal, familiarity, ambiguity, transparency, and AoA—were distributed across two separate questionnaires, with each questionnaire containing three variables. Valence and arousal are interrelated emotional variables, and transparency and ambiguity are interrelated semantic variables. To prevent interference, valence and arousal were not included in the same questionnaire, and similarly, transparency and ambiguity were kept separate. Thus, the first questionnaire included valence, familiarity, and ambiguity, while the second included arousal, transparency, and AoA. The two questionnaires for each list were distributed on separate days to ensure the quality of participants' ratings. The order of the three dimensions within the same questionnaire was randomized, except that the familiarity ratings were to be completed first to prevent participants' familiarity with idiomatic expressions from increasing due to repeated exposure during the rating of other dimensions. The order of idiomatic expressions within the questionnaires was also randomized. Participants were randomly and evenly distributed across the 10 lists and asked to rate all six dimensions of the same 50 idiomatic expressions.

We employed a nine-point scale with an "I don't know" option to gather ratings of each variable for the idiomatic expressions. For valence, in line with previous norming studies for single words (Stadthagen-Gonzalez et al., 2017; Yao et al., 2017) and idiomatic expressions (Citron et al., 2016a, 2016b), participants were required to rate the extent to which the idiom's figurative meaning evoked negative (sad, angry, dissatisfied, melancholic, despairing = rating of 1) or positive (happy, joyful, satisfied, optimistic, hopeful = rating of 9) feelings. For arousal, following previous norming studies for single words (Yao et al., 2017) and idiomatic expressions (Citron et al., 2016a, 2016b), participants were required to rate the extent to which the idiom's figurative meaning evoked feelings of calmness (relaxed, sluggish, boring, drowsy, listless = rating of 1) or excitement (tense, excited, enthusiastic, alert, invigorated = rating of 9). For familiarity, to enhance the consistency and accuracy in participants' ratings, we drew on previous studies and asked participants to rate familiarity by simultaneously considering both how often they encountered each idiomatic expression in its figurative meaning (Citron et al., 2016a, 2016b) and the extent to which they were familiar with the figurative meaning of the idiomatic expression (Yao et al., 2017) (have never seen/used, completely unfamiliar = rating of 1; see/use very often, very familiar = rating of 9). For ambiguity, following previous norming studies (e.g., Bonin et al., 2013;

Table 1 Descriptive statistics for the 500 Chinese idiomatic expressions across all the variables

Variable	Mean	SD	Median	Min	Max	Q1	Q3	IRQ	Skewness	Kurtosis	Unknown no. (%)	Outliers no. (%)
Valence	4.111	1.405	3.801	1.675	8.024	3.024	5.054	2.030	0.669***	2.688***	527 (2.51%)	601 (3.0%)
Arousal	5.687	0.842	5.705	3.135	7.643	5.051	6.290	1.239	−0.075	2.492***	412 (2.04%)	460 (2.3%)
Familiarity	7.235	1.148	7.563	1.366	8.860	6.769	8.025	1.256	−1.577***	5.920***	0 (0.00%)	188 (0.9%)
Transparency	6.371	0.918	6.500	2.567	8.256	5.843	7.024	1.181	−0.816***	4.052***	308 (1.53%)	372 (1.9%)
Ambiguity	4.618	0.989	4.558	2.250	7.825	3.881	5.304	1.423	0.242*	2.849***	649 (3.09%)	659 (3.3%)
AoA	12.576	1.955	12.577	7.195	19.778	11.297	13.842	2.544	0.131	3.204***	993 (4.82%)	1452 (7.6%)
Frequency	8.847	2.165	8.863	0.000	14.654	7.488	10.280	2.792	−0.426***	4.043***	-	-

Tabossi et al., 2011), participants were required to estimate the probability that an idiomatic expression is used literally rather than figuratively (completely impossible to be used literally = rating of 1; almost always used literally = rating of 9). For transparency, following Citron et al., (2016a, 2016b), participants were asked to rate the extent to which the figurative meaning of an idiomatic expression could be inferred from the meaning of its constituent words (cannot infer the figurative meaning at all = rating of 1; can fully infer the figurative meaning = rating of 9). For AoA, participants were required to recall as accurately as possible the age at which they learned the figurative meaning of each idiomatic expression.

Participants were instructed to rate based on their own intuition and at their own pace, with the option to take breaks at any time. The entire norming study took approximately 60 minutes, and participants received 30 yuan for their participation. Detailed instructions for the norming task, including variable definitions, explanations of the nine-point Likert scale, and rating examples, were provided in the questionnaires (see Appendix A for the original Chinese instructions and their English translations).

Results and discussion

Data cleaning

A total of 123,879 data points were collected. For 19 participants (10 male), we only had data from the first questionnaire (valence, familiarity, and ambiguity) because they did not complete the second questionnaire (arousal, transparency, and AoA). The data from these 19 participants' first questionnaires were included in the analysis. For each dimension, "I don't know" responses were recorded as missing values (see Table 1 for the number and proportion of "I don't know" responses). All statistical analyses were conducted using R version 4.3.1. (R Core Team, 2023).

We adopted the following criteria[2] to exclude outliers (see Table 1 for the number and proportion of outliers for each variable): First, we excluded participants who selected the same rating for more than 80% of their total responses for each list (Yao et al., 2017). No participant was excluded based on this criterion. We then calculated the mean ratings for each dimension of each idiomatic expression across participants and removed data points deviating from the mean by more than $\pm$ 3 standard deviations, resulting in the removal of 3729 outliers (3.01%). Next, following the practice of Warriner et al. (2013), we computed the correlation coefficients between each participant's ratings for all dimensions for each set of 50 idiomatic expressions and the mean ratings for all participants in that set. If the correlation coefficient for a participant was less than 0.05, the data from that participant were deleted. Data from one participant were deleted based on this criterion. After these steps, 120,003 data points were retained for analysis.

---

[2] To ensure participants had sufficient knowledge of the idiomatic expressions they rated, we calculated the proportion of "I don't know" responses for each participant across different dimensions of the same idiomatic expression. If a participant selected "I don't know" for more than three dimensions out of six subjective dimensions for an idiomatic expression, we assumed they lacked comprehensive knowledge of the expression and excluded all their ratings for that expression. Based on this, we excluded 437 data points (0.36% of the total). After exclusion, the results remained unchanged from the original dataset. Additionally, to ensure that each idiomatic expression could be rated by most participants across all dimensions, within each dimension of an idiomatic expression, we calculated the proportion of participants who selected "I don't know". If more than 50% of participants rated an idiomatic expression as "I don't know" for a dimension, it would suggest general difficulty in assessing that dimension for that idiomatic expression. No idiomatic expression received more than 10% of "I don't know" responses, which is within a normal range given individual differences in language experience. Therefore, no data were excluded based on this criterion.

Reliability of the measures

To assess the reliability of participants' ratings across the six variables, we first randomly divided the ratings for each idiomatic expression within each dimension into two equally sized subgroups and calculated the mean ratings for each subgroup. Next, we computed the Spearman–Brown split-half coefficient (*r*) for the mean ratings of the 500 idiomatic expressions within the same variable across the two subgroups (Yao et al., 2017).

The Spearman–Brown split-half coefficient for the reliability analysis of ratings across all variables exceeded 0.690. This indicated that the ratings for all the measured variables were consistently reliable. Valence showed the highest reliability (*r* = .91), followed by familiarity (*r* = .83) and arousal (*r* = .81). Both AoA (*r* = .74) and transparency (*r* = .73) exhibited moderate reliability, while ambiguity (*r* = .69) had the lowest reliability. These results suggest that participants'

estimations of valence were the most consistent, indicating that native Chinese speakers uniformly perceive emotional information conveyed by idiomatic expressions. In contrast, there was greater individual variation in participants' ratings of ambiguity. This may be because the likelihood of using an idiomatic expression literally or figuratively is closely tied to each participant's unique linguistic experiences and knowledge.

Rating consistency

We examined rating consistency by computing the relationship between the mean and standard deviation of ratings for all idiomatic expressions across dimensions. As shown in Fig. 1, idiomatic expressions with moderate levels of valence, familiarity, ambiguity, and transparency received less consistent ratings, while those with extreme values were rated more consistently. In contrast, arousal
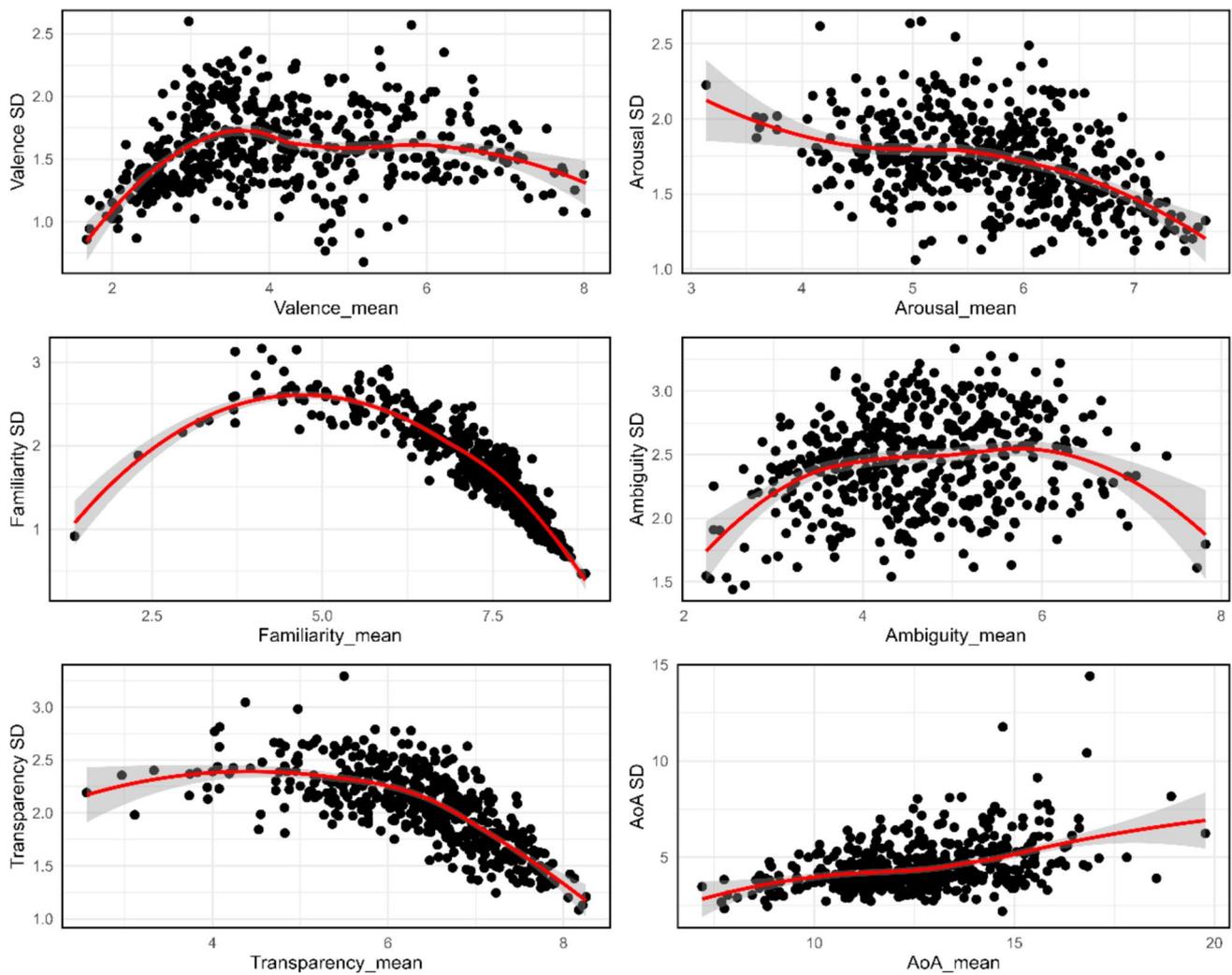

Fig. 1 Standard deviations of ratings for each variable plotted against respective mean ratings

demonstrated a negative linear correlation with the standard deviation of ratings, while AoA exhibited a positive linear correlation with the standard deviation of ratings among participants. This suggests that Chinese speakers exhibit smaller individual differences in ratings for more arousing or earlier-acquired idiomatic expressions.

Descriptive statistics

In addition to the six affective and non-affective psycholinguistic variables, we included the logged frequency of each idiomatic expression as a separate psycholinguistic variable, which was directly obtained from the CCL corpus. A complete list of 500 Chinese three-character idiomatic expressions, including their literal translations in English, figurative meanings in English, and the mean and standard deviation of ratings for each idiomatic expression across dimensions, can be found in Appendix B. The raw data for each participant's responses to each idiomatic expression in each dimension can be found in Appendix C. We also computed descriptive statistics for all 500 idiomatic expressions across valence, arousal, familiarity, transparency, ambiguity, AoA, and objective frequency (see Table 1).

Frequency was log-transformed. SD = standard deviation; Q1 = 25th percentile; Q3 = 75th percentile; IQR = interquartile range

Significance level: * $p < .05$, ** $p < .01$, *** $p < .001$. Unknown = "I don't know" responses

For the affective variables, Table 1 and Fig. 2 show the distribution of valence and arousal ratings for the idiomatic expressions. As discussed in the literature review, there is an ongoing debate on the similarities (e.g., Kousta et al., 2009) and heterogeneity in processing negative, neutral, and positive stimuli. Therefore, using the criteria from prior studies (e.g., Yao et al., 2017), we categorized Chinese idiomatic expressions into negative (valence < 4), neutral (4 ≤ valence ≤ 6), and positive (valence > 6) groups. The median and mean valence of the idiomatic expressions were both below 5, with the number of negative idiomatic expressions (valence < 4, $N = 273$) exceeding the number of neutral idiomatic expressions (4 ≤ valence ≤ 6, $N = 166$) and positive idiomatic expressions (valence > 6, $N = 61$). This finding replicates Citron et al., and and's (2016a, 2016b) study on German idiomatic expressions and Gavilán et al.'s (2021) study on Spanish idiomatic expressions, indicating a cross-linguistic consistency in terms of the dominance of negative idiomatic expressions. In contrast, previous research has found that positive Chinese two-character (Chan & Tse,
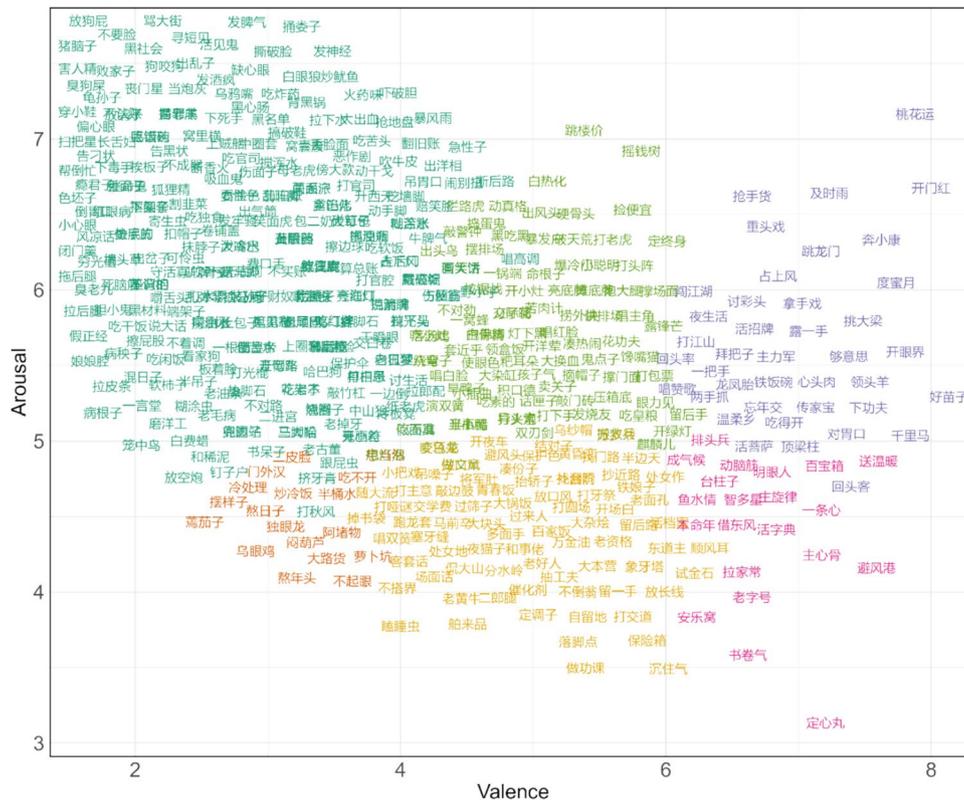

Fig. 2 The distribution of valence and arousal ratings for the idiomatic expressions

2024; Ho et al., 2015) or multi-character word combinations (ranging from 1 to 4 characters) are more common than negative ones (e.g., Lv et al., 2024; Xu et al., 2022). This suggests that Chinese idiomatic expressions and words play different roles in encoding and expressing emotional valence. Given the substantial inclusion of Chinese idiomatic expressions in our study, the observed dominance of negative emotions in Chinese idiomatic expressions is unlikely to be due to sample selection biases. Instead, our results may reflect that Chinese three-character idiomatic expressions are used to express negative emotions more commonly than neutral or positive emotions. This tendency could stem from language speakers' inclination to use figurative language to indirectly convey negative emotions to avoid conflicts, or because idiomatic expressions are more effective at expressing negative emotions than expressing neutral or positive ones (Cacciari & Tabossi, 1988; Drew & Holt, 1988).

The median and mean arousal values for idiomatic expressions were both above 5, indicating that a higher proportion of Chinese idiomatic expressions make people feel excited rather than calm. Our results are inconsistent with studies on Chinese words (Chan & Tse, 2024; Ho et al., 2015; Lv et al., 2024), which found a greater proportion of low-arousing words than that of high-arousing ones, indicating a difference in the encoding of emotional intensity between Chinese words and idiomatic expressions. This also contrasts with studies on German (Citron et al., 2016a, 2016b), Spanish (Gavilán et al., 2021), and English (Morid & Sabourin, 2024) idiomatic expressions, which identified a dominance of low-arousal idiomatic expressions. The finding suggests a cross-cultural difference, with Chinese idiomatic expressions conveying stronger emotional content compared to idiomatic expressions in other languages.

For non-affective variables, both the median and mean familiarity ratings exceeded 5, replicating previous studies on German (Citron et al., 2016a, 2016b) and Spanish idiomatic expressions (Gavilán et al., 2021). This indicates that most Chinese idiomatic expressions are well known to native speakers. Similarly, the median and mean transparency ratings were also above 5, suggesting that, similar to German idiomatic expressions (Citron et al., 2016a, 2016b) and Chinese four-character idiomatic expressions (Li et al., 2016), the figurative meanings of more than half of the Chinese three-word idiomatic expressions can be easily inferred from the semantic clues of the constituent words. The median and mean ambiguity were both below 5, suggesting that, unlike Spanish idiomatic expressions (Gavilán et al., 2021) and Chinese four-character idiomatic expressions (Li et al., 2016), Chinese three-character idiomatic expressions are more likely to be used figuratively rather than literally. In addition, the mean and median age for acquiring idiomatic expressions was around 12 years old, suggesting that the age of 12, which falls within the critical period of language

acquisition (Snow & Hoefnagel-Höhle, 1978), is a key period for the acquisition of idiomatic expressions, with at least half of the idiomatic expressions learned before this age. Our results also indicate that the acquisition of the figurative meanings of Chinese three-character idiomatic expressions is an important milestone and foundation for language development during this critical period (Wray, 2002).

We also examined the distributional characteristics of the ratings for the variables (see Table 1 for the skewness and kurtosis values and Fig. 3 for the distributions of ratings for each variable). Skewness tests revealed positive skewness for valence ($skewness = 0.669$, $p < .001$) and ambiguity ($skewness = 0.242$, $p < .05$), while familiarity ($skewness = -1.577$, $p < .001$) and transparency ($skewness = -0.816$, $p < .001$) demonstrated negative skewness. Arousal and AoA ratings showed a normal distribution in terms of skewness. Kurtosis tests indicated platykurtic distributions for valence ($kurtosis = 2.688$, $p < .001$), arousal ($kurtosis = 2.492$, $p < .001$), and ambiguity ($kurtosis = 2.849$, $p < .001$), while familiarity ($kurtosis = 5.920$, $p < .001$), transparency ($kurtosis = 4.052$, $p < .001$), and AoA ($kurtosis = 3.204$, $p < .001$) exhibited leptokurtic distributions. Since most variables did not meet the assumption of normality, we applied bootstrapping to all parametric statistical analyses (1000 resamples, 95% percentile confidence intervals) when examining the relationships among variables (Bradley & Lang, 1999; Citron et al., 2016a, 2016b).

Relationships between variables

We reported the linear partial correlation coefficients between variables, which control for the influence of other variables and reveal the residual correlation between two specific variables (Citron et al., 2016a, 2016b). Significant partial correlations up to ± .1 were categorized as "small correlations", those between ± .1 and ± .3 as "moderate correlations", and those between ± .3 and ± .5 as "large correlations" (Citron et al., 2016a, 2016b).

Valence is a bipolar dimension, whereas the ranges for all other variables span from absence to complete presence of a certain attribute (Citron et al., 2016a, 2016b; Yao et al., 2017). We then categorized Chinese idiomatic expressions into negative (valence < 4), neutral ($4 \leq$ valence ≤ 6), and positive (valence > 6) groups, and calculated the correlations among variables within each group. We also calculated the mean values of the variables, excluding valence, across the negative, neutral, and positive groups. Table 2 and Fig. 4 show the partial correlations among variable ratings for all idiomatic expressions, as well as negative, neutral, and positive idiomatic expressions. To explore the potential quadratic relationship between valence and other variables, we built quadratic regression models. In these models, the variables of interest were included as the dependent variables, while
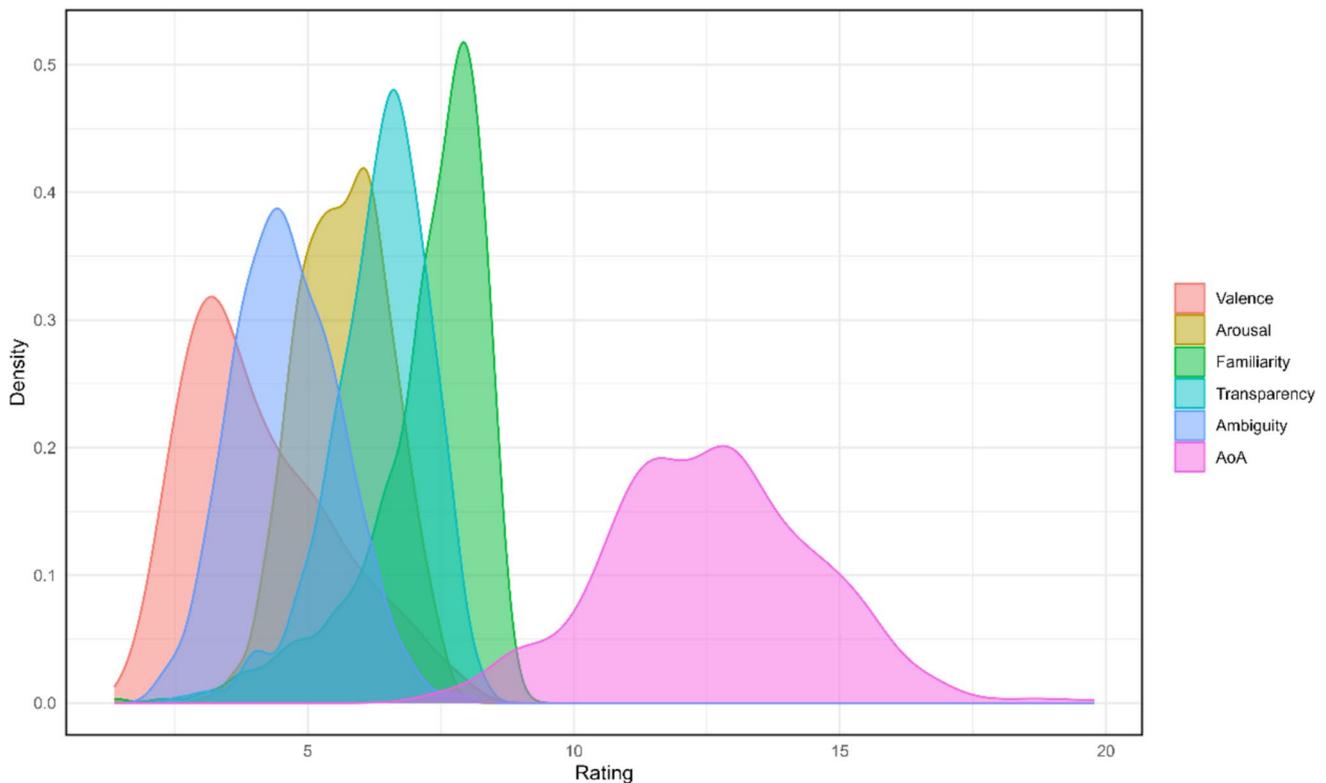
Fig. 3 The distributional characteristics of the ratings across variables

valence and valence squared were included as independent variables. Other affective and non-affective psycholinguistic variables that were not of interest were simultaneously included in the models as control variables. Table 3 presents the quadratic regression results for valence and other variables.

Relationships between affective variables

A strong negative correlation between arousal and valence was found across all idiomatic expressions ($r = -.519$, $p < .001$), as well as in negative ($r = -.496$, $p < .001$) and neutral idiomatic expressions ($r = -.227$, $p = .046$). However, no correlation was found within positive idiomatic expressions (see Table 2 and Fig. 5).

The negative correlation between valence and arousal observed in all idiomatic expressions in our study aligns with the findings of Citron et al., (2016a, 2016b) for German idiomatic expressions. However, Gavilán et al. (2021) reported a moderate positive correlation between valence and arousal in Spanish idiomatic expressions. This discrepancy indicates that the relationship between valence and arousal in mixed-valence idiomatic expressions varies across different languages.

After categorizing idiomatic expressions into negative, neutral, and positive, we found that neutral and negative idiomatic expressions with a stronger negative valence were more likely to evoke stronger emotional responses. This replicates most studies on isolated words (Citron et al., 2014), including those on Chinese two-character (Chan & Tse, 2024; Ho et al., 2015; Yao et al., 2017) and multi-character words (Lv et al., 2024; Xu et al., 2022), as well as on idiomatic expressions (Citron et al., 2016a, 2016b, for German; Gavilán et al., 2021, for Spanish). Additionally, we found that valence has a stronger effect on arousal in negative idiomatic expressions than in neutral or positive ones, with arousal being more sensitive to changes in negative information. This aligns with the "approach–withdrawal" model of emotion processing (Robinson et al., 2004), which posits that humans tend to withdraw from both negative and high-arousal stimuli, as they signal potential threats. Thus, negative and high-arousal information consistently trigger corresponding mental states and are highly correlated. Furthermore, as suggested by the automatic vigilance model of emotion (Estes & Verges, 2008), avoiding negative threats is more crucial for survival than approaching positive or neutral stimuli, which leads to delayed attentional disengagement and heightened emotional involvement in processing negative emotions. As a result, language speakers show greater emotional intensity changes in response to negative

Table 2 Linear partial correlations between the variables for all idiomatic expressions and negative, neutral, and positive idiomatic expressions

	Valence	Arousal	Familiarity	Transparency	Ambiguity	AoA	Frequency
	All idiomatic expressions						
Valence	1						
Arousal	−.519***	1					
Familiarity	−.019	.065	1				
Transparency	.129*	.241***	.520***	1			
Ambiguity	.051	−.164*	−.140*	.407***	1		
AoA	.119*	−.031	−.334***	−.174*	−. 117	1	
Frequency	.241***	−.061	.275***	−.085	.159*	.003	1
	Negative idiomatic expressions						
Valence	1						
Arousal	−.496***	1					
Familiarity	−.012	.107	1				
Transparency	−.172*	.101	.492***	1			
Ambiguity	.197*	−.005	−.025	.483***	1		
AoA	.005	−.052	−.324***	−.183*	−.062	1	
Frequency	.120	−.078	.311***	−.056	.136	.020	1
	Neutral idiomatic expressions						
Valence	1						
Arousal	−.227*	1					
Familiarity	−.045	.006	1				
Transparency	.234*	.274*	.546***	1			
Ambiguity	.043	.286*	−.062	.364***	1		
AoA	.219*	.042	−.378***	−.173	−.180	1	
Frequency	.137	−.020	.221*	−.176	.176	−.064	1
	Positive idiomatic expressions						
Valence	1						
Arousal	.250	1					
Familiarity	.124	.191	1				
Transparency	.201	.056	.451***	1			
Ambiguity	−.205	−.012	.158	.282	1		
AoA	−.153	.014	−.195	−.152	−.179	1	
Frequency	.179	−.207	.232	−.007	.175	.147	1

Frequency was log-transformed. Significance level: * $p < .05$, ** $p < .01$, *** $p < .001$

information than to positive or neutral information (Citron et al., 2014). However, Lv et al. (2024) found that, contrary to most previous word-based studies (e.g., Yao et al., 2017 for two-character Chinese words) and the current study on three-character idiomatic expressions, arousal was more sensitive to changes in positive information than negative information in Chinese multi-character words (ranging from 1 to 4 characters). This discrepancy may be due to structural differences, as Lv et al. (2024) did not control for the number of characters in the words, which may have led to greater variability.

A quadratic regression model was constructed with valence and valence squared as independent variables, arousal as the dependent variable, and other semantic variables as control variables (see Table 3). The model accounted for 51.1% of the variance ($R^2 = .511$), $F(7, 492) = 321.695$), with both valence and valence squared as significant predictors. Additionally, the arousal ratings of neutral idiomatic expressions ($M = 5.15$) were lower than those of positive ($M = 5.22$) [$t(87) = 8.73$, $p < .001$] and negative idiomatic expressions ($M = 6.11$) [$t(350) = 14.42$, $p < .001$].

The results revealed a typical asymmetric U-shaped curve between valence and arousal, with extremely valenced (either extremely positive or extremely negative) idiomatic expressions rated as the most arousing (see Fig. 6). This finding is consistent with previous research on single words (Chan & Tse, 2024; Ho et al., 2015; Lv et al., 2024; Warriner et al., 2013; Xu et al., 2022; Yao et al., 2017), as well as German (e.g., Citron et al., 2016a, 2016b), Spanish (e.g., Gavilán et al., 2021), and English (Morid & Sabourin, 2024)
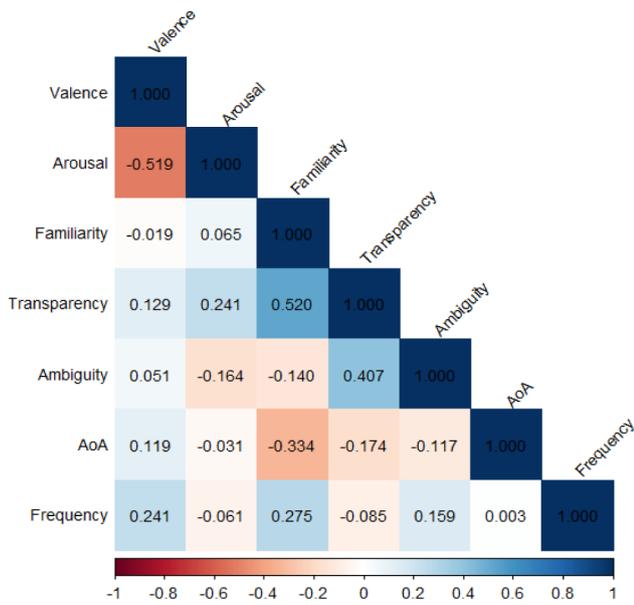
(a) All idiomatic expressions
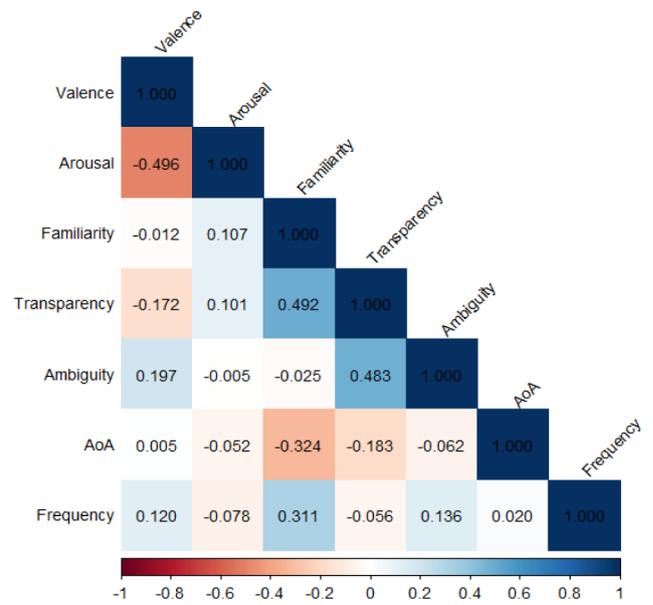
(b) Negative idiomatic expressions
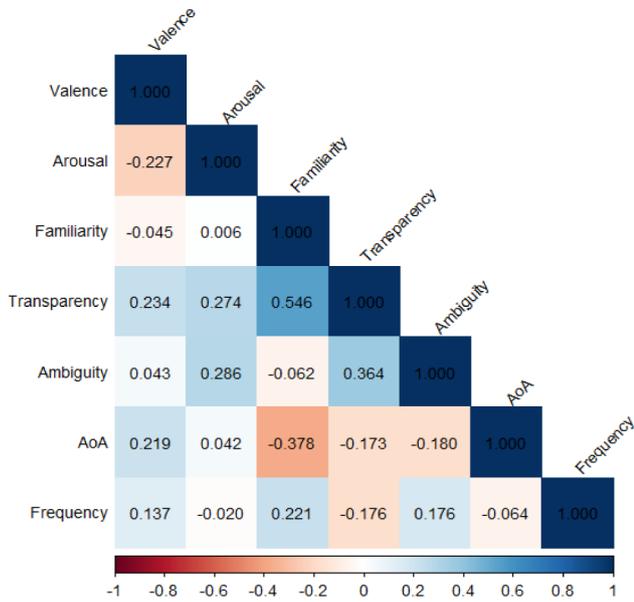
(c) Neutral idiomatic expressions
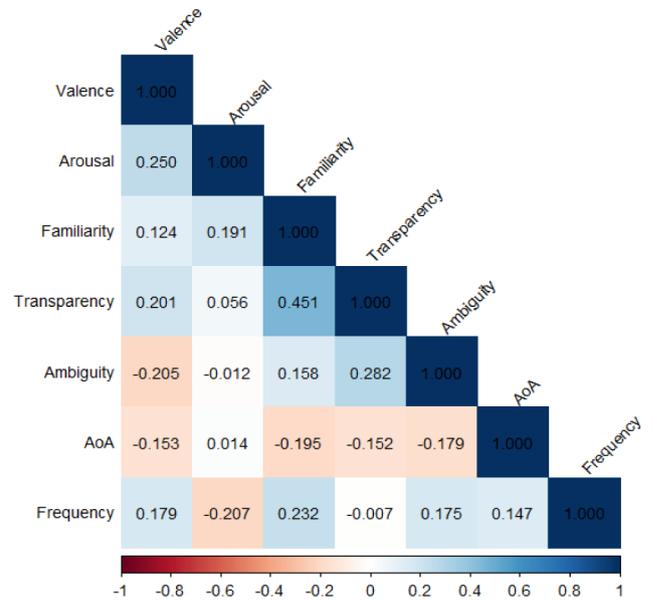
(d) Positive idiomatic expressions

Fig. 4 Partial correlations among all the variables for all idiomatic expressions, as well as for positive, neutral, and negative idiomatic expressions

idiomatic expressions. This suggests that, similar to other languages, the association between more valenced emotions and stronger arousal in Chinese extends beyond words (Chan & Tse, 2024; Ho et al., 2015; Lv et al., 2024; Yao et al., 2017) to figurative three-character idiomatic expressions.

Moreover, the U-shape pattern suggests that both negative and positive idiomatical expressions share similar emotional processing mechanisms, as proposed by the model of motivated attention and affective states (Lang et al., 1998). Our results on the valence–arousal relationship further support

Table 3 Quadratic regression results for valence and other variables in all idiomatic expressions

Predictor variables	Outcome variables					
	Arousal	Familiarity	Transparency	Ambiguity	AoA	Frequency
	Coefficient [CI] (*p*-value)	Coefficient [CI] (*p*-value)	Coefficient [CI] (*p*-value)	Coefficient [CI] (*p*-value)	Coefficient [CI] (*p*-value)	Coefficient [CI] (*p*-value)
Intercept	8.254 [7.241,9.329] (<.001)	3.661 [2.038, 5.255] (.001)	3.218 [1.969,4.427] (<.001)	1.820 [−0.041, 3.685] (.145)	19.762 [16.987, 22.332] (<.001)	4.778 [0.958, 8.739] (.080)
Valence	**−1.377**** [−1.656, −1.125] (.001)	0.174 [−0.139, 0.485] (.355)	**−0.452*** [−0.702, −0.208] (.012)	0.501 [0.124, 0.891] (.063)	0.611 [−0.082, 1.316] (.198)	−0.148 [−0.933,0.663] (.490)
Valence squared	**0.116***** [0.087, 0.148] (<.001)	−0.019 [−0.050, 0.011] (.337)	**0.054**** [0.028, 0.080] (.004)	−0.048 [−0.088, −0.011] (.081)	−0.047 [−0.116, 0.023] (.288)	0.059 [−0.024, 0.139] (.286)
Arousal	-	0.099 [−0.019,0.212] (.188)	0.144 [0.053, 0.237] (.021)	−0.153 [−0.280, −0.026] (.089)	−0.025 [−0.257, 0.226] (.464)	−0.252 [−0.520, 0.022] (.187)
Familiarity	0.067 [−0.002,0.139] (.181)	-	0.397 [0.340, 0.460] (.033)	−0.160 [−0.261, −0.064] (.022)	−0.688 [−0.877, −0.502] (<.001)	0.720 [0.502, 0.942] (<.001)
Transparency	0.161 [0.065, 0.254] (.020)	0.686 [0.559,0.806] (<.001)	-	0.611 [0.491,0.735] (<.001)	−0.425 [−0.711, −0.150] (.019)	−0.326 [−0.621, −0.048] (.115)
Ambiguity	−0.078 [−0.148, −0.012] (.095)	−0.127 [−0.204, −0.052] (<.001)	0.284 [0.234, 0.336] (<.001)	-	−0.222 [−0.420, −0.029] (.080)	0.377 [0.182, 0.568] (.009)
AoA	−0.003 [−0.042, 0.036] (.475)	−0.164 [−0.210, −0.119] (<.001)	−0.060 [−0.098, −0.021] (.021)	−0.067 [−0.120, −0.010] (.075)	-	0.010 [−0.106, 0.116] (.506)
Frequency	−0.025 [−0.051, 0.004] (.188)	0.107 [0.074, 0.142] (<.001)	−0.028 [−0.050, −0.003] (.116)	0.071 [0.036,0.106] (.007)	0.006 [−0.060,0.082] (.509)	-
R^2	0.511	0.599	0.635	0.327	0.426	0.246
F value	$F(7, 492) = 321.695$ [220.802, 442.047] **$p < .001$**	$F(7, 492) = 4.985$ [0.007, 19.615] $p = .254$	$F(7, 492) = 2.449$ [0.002 13.051] $p = .453$	$F(7, 492) = 27.228$ [7.146, 58.893] **$p = .001$**	$F(7, 492) = 17.979$ [2.676, 42.354] **$p = .006$**	$F(7, 492) = 71.41$ [0.186, 0.309] **$p < .001$**

Frequency was log-transformed. Significance level: * $p < .05$, ** $p < .01$, *** $p < .001$

the view that arousal and valence are not independent emotional dimensions but may interactively influence language processing.

Relations between affective and non-affective variables

Valence and transparency

We observed moderate positive correlations between transparency and valence for all idiomatic expressions ($r = .129$, $p = .034$) and neutral idiomatic expressions ($r = .234$, $p = .034$), and a moderate negative correlation for negative idiomatic expressions ($r = −.172$, $p = .041$; see Fig. 7). Transparency of neutral idiomatic expressions ($M = 6.13$) was lower than that of positive ($M = 6.74$) [$t(123) = −3.07$,

$p = .033$] and negative idiomatic expressions ($M = 6.43$) [$t(348) = 3.27$, $p = .022$]. Negative idiomatic expressions ($M = 6.43$) were significantly less transparent than positive idiomatic expressions ($M = 6.74$) [$t(151) = −5.49$, $p < .001$].

Gavilán et al. (2021) similarly reported that figurative meanings of Spanish idiomatic expressions with richer positive emotions were easier to infer from their constituent words. However, beyond the overall positive correlation, our study revealed a more complex U-shaped pattern in Chinese three-character idiomatic expressions: more valenced idiomatic expressions were more semantically transparent than neutral ones. Specifically, the facilitative effect of positive information on the derivation of figurative meaning from constituent words was limited to neutral idiomatic expressions, whereas in negative idiomatic expressions, richer negative information contributed to such derivation. The
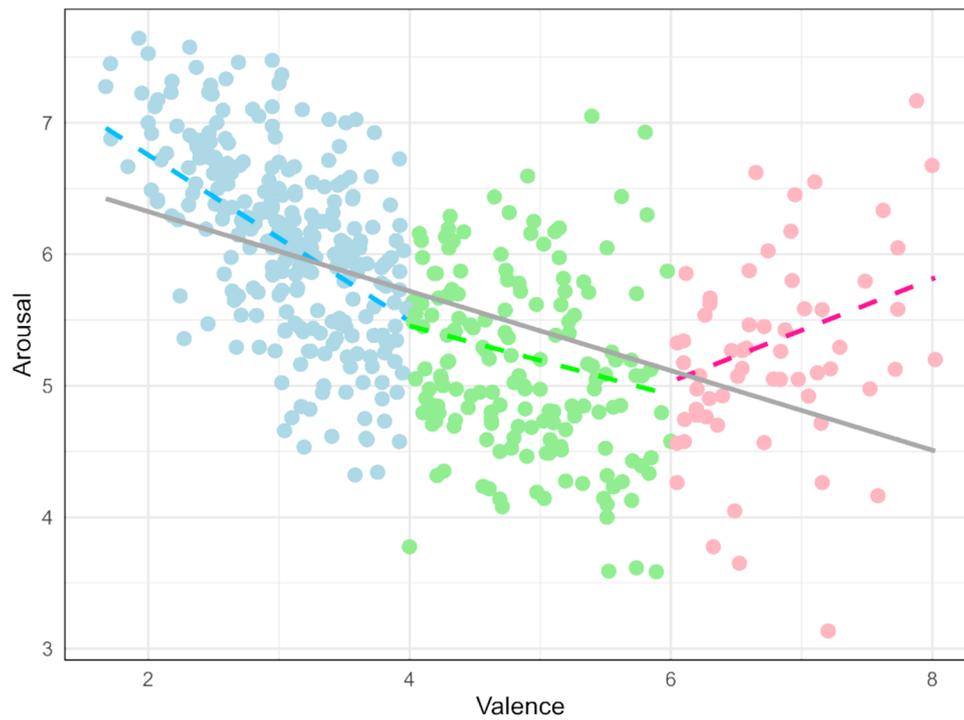
Fig. 5 Linear correlations between arousal and valence ratings of idiomatic expressions. Scatter points and/or linear fits (with other variables controlled) are shown for all idiomatic expressions (gray), as well as negative (blue), neutral (green), and positive idiomatic expressions (pink)
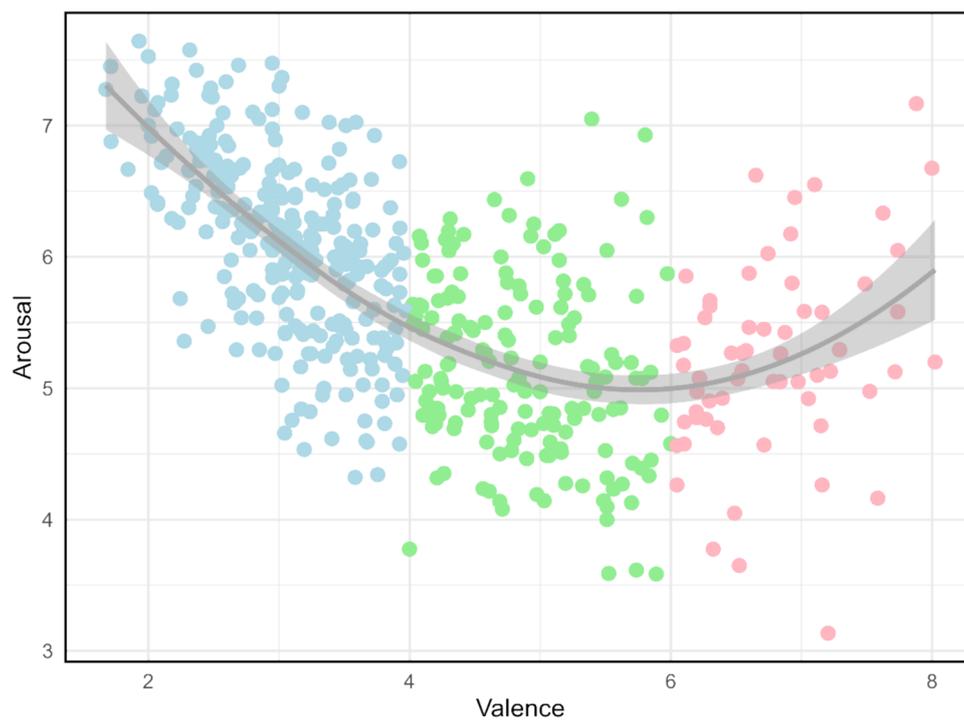

Fig. 6 U-shaped curve between valence and arousal ratings of idiomatic expressions. Scatter points and/or nonlinear fits are shown for all idiomatic expressions (gray), as well as negative (blue), neutral (green), and positive idiomatic expressions (pink)
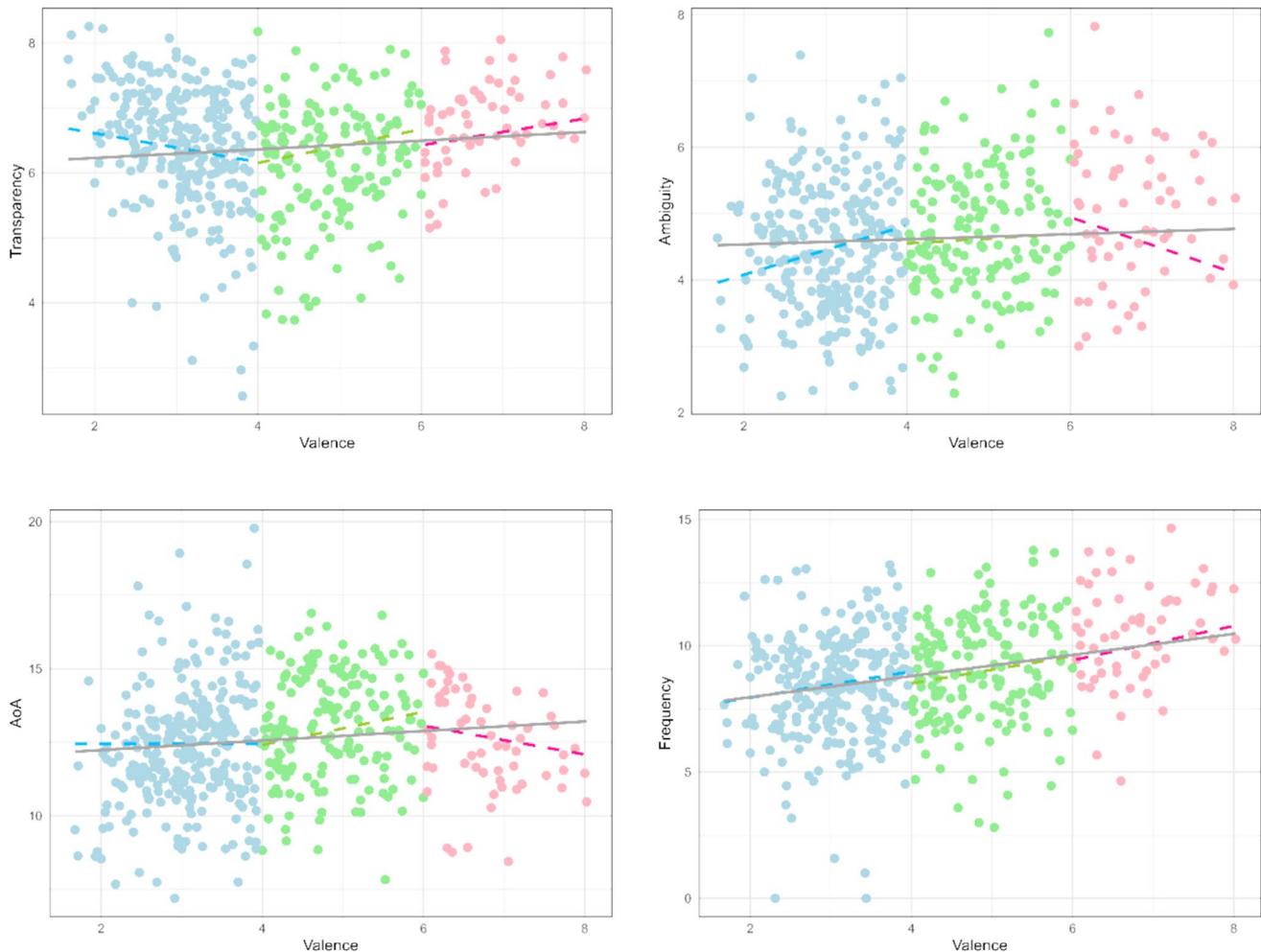
Fig. 7 Linear correlations between valence and non-affective variables. Scatter points and/or linear fits (with other variables controlled) are shown for all idiomatic expressions (gray), as well as negative (blue), neutral (green), and positive idiomatic expressions (pink)

U-shaped pattern also suggests that idiomatic expressions whose figurative meanings are easier to infer from their constituent words tend to convey richer positive or negative emotions. This further indicates that the emotional content of constituent words enhances the emotionality of the figurative meaning of idiomatic expressions, which supports the dual-route model of idiomatic expression processing (Carrol & Conklin, 2014).

Valence and ambiguity

A positive linear relationship between ambiguity and valence was observed only in negative idiomatic expressions ($r = .197$, $p = .024$; see Fig. 7). Additionally, the ambiguity of neutral idiomatic expressions ($M = 4.67$) was significantly lower than that of positive idiomatic expressions ($M = 5.00$) [$t(87) = -3.54$, $p = .013$].

Our results suggest that more negatively valenced or less positively valenced idiomatic expressions are more likely to be used figuratively, which aligns with the findings of Gavilán et al. (2021) on Spanish idiomatic expressions, and Citron et al., (2016a, 2016b) on German negative idiomatic expressions. These results indicate that the figurative usage of Chinese three-character idiomatic expressions plays a significant role in activating and conveying negative emotionality. This may be because the figurative usage of three-character idiomatic expressions allows for a more indirect expression of negative emotions than transparent expressions, helping to avoid conflicts in communication.

From a cognitive perspective, the figurative use of idiomatic expressions conveys abstract meanings and reflects internal emotional states (Citron et al., 2016a, 2016b), while the literal interpretations of constituent words in idiomatic expressions are associated with concrete objects and actions, reflecting embodied language experiences in

external contexts (Yao et al., 2017; Vigliocco et al., 2014). Our findings therefore suggest that humans are less prone to associate concrete embodied experiences with negative emotions (Vigliocco et al., 2014), likely due to their tendency to avoid negative stimuli in real life (Robinson et al., 2004).

Valence and AoA

We found a significant positive linear relationship between AoA and valence for all idiomatic expressions ($r = .119$, $p = .049$) and within neutral idiomatic expressions ($r = .219$, $p = .041$; see Fig. 7). The AoA of negative idiomatic expressions ($M = 12.31$) was significantly lower than that of neutral idiomatic expressions ($M = 13.05$) [$t(380) = -3.93$, $p < .001$].

This finding suggests that more negatively valenced idiomatic expressions are acquired earlier, consistent with word-level findings (Citron et al., 2014, for English words; Della Rosa et al., 2010, for Italian words). As mentioned earlier, Chinese three-character idiomatic expressions play a crucial role in indirectly conveying negative emotionality, likely as a strategy to avoid conflict in communication. Due to such communicative demand, language users are more likely to encounter and acquire negative idiomatic expressions than neutral or positive ones at an earlier age. In addition, negative stimuli are associated with delayed attentional disengagement and require more cognitive resources for processing compared to neutral and positive stimuli (Estes & Verges, 2008). Therefore, negative idioms could be more salient in the mental lexicon and prioritized for acquisition.

Our results further highlight an asymmetry between positive, neutral, and negative idiomatic expressions in child language development. During the early stages of language acquisition, language users are more inclined to use idiomatic expressions to express negative information, while neutral and positive expressions may be often expressed through non-idiomatic literal phrases. Consequently, encoding neutral and positive information within a figurative language may necessitate richer, usage-based linguistic experiences, emerging later in development (Ellis et al., 2013).

Valence and frequency

We found significant positive correlations between frequency and valence across all idiomatic expressions ($r = .241$, $p < .001$; see Fig. 7). Additionally, the frequency of negative idiomatic expressions ($M = 8.37$) was significantly lower than that of neutral idiomatic expressions ($M = 9.04$) [$t(343) = -3.26$, $p = .025$] and positive idiomatic expressions ($M = 10.46$) [$t(119) = -4.85$, $p = .001$]. Furthermore, the frequency of neutral idiomatic expressions ($M = 9.04$) was significantly lower than that of positive idiomatic expressions ($M = 10.46$) [$t(96) = -7.62$, $p < .001$].

The results showed that positive idiomatic expressions were used most frequently, followed by neutral idiomatic expressions, and finally negative idiomatic expressions. This pattern may be explained by the greater variety of negative Chinese three-character idiomatic expressions compared to neutral or positive ones. As a result, the average frequency of individual negative idiomatic expressions may be lower, as usage is distributed across a greater variety of idiomatic expressions in negative categories than in positive and neutral ones. Another possible explanation is that participants might avoid assigning negative ratings to frequently used idiomatic expressions (Citron et al., 2014). Such avoidance of negative experiences aligns with the "approach–withdrawal" model of emotion processing (Robinson et al., 2004).

Arousal and transparency

We found a moderate positive correlation between arousal and transparency across all idiomatic expressions ($r = .241$, $p < .01$) and within neutral idiomatic expressions ($r = .274$, $p = .021$; see Fig. 8). These results indicate that if the figurative meaning of an idiomatic expression is more easily inferred from the literal meaning of its constituent words, it is more likely to evoke a stronger emotional response. This finding is consistent with Citron et al., and and's (2016a, 2016b) research on German idiomatic expressions and further suggests that the literal meanings of constituent words in idiomatic expressions significantly contribute to the expression and perception of intense emotions.

Arousal and ambiguity

Our results showed moderate negative correlations between ambiguity and arousal across all idiomatic expressions ($r = -0.164$, $p = .013$) and within neutral idiomatic expressions ($r = -0.286$, $p = .015$; see Fig. 8). The finding indicates that idiomatic expressions with a higher likelihood of being used figuratively tend to convey greater emotional intensity, which echoes the findings of Citron et al., (2016a, 2016b) on German idiomatic expressions. In a neuroimaging study, Citron and Goldberg (2014) found that the figurative usage of language elicited strong activation in the amygdala and the anterior portion of the hippocampus compared to its literal usage. We can therefore infer that the figurative usage of idiomatic expressions was more emotionally engaging and encoded more intense emotions in real-life contexts compared to their literal usage.
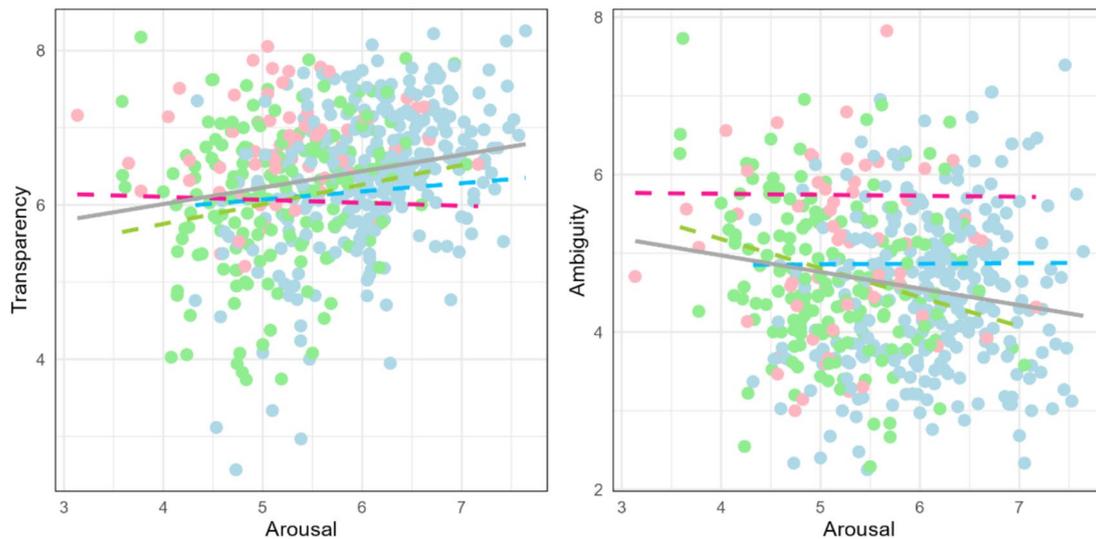
Fig. 8 Linear correlations between arousal and non-affective variables. Scatter points and/or linear fits (with other variables controlled) are shown for all idiomatic expressions (gray), as well as negative (blue), neutral (green), and positive idiomatic expressions (pink)

Relations between non-affective variables

Familiarity and transparency

The results revealed significant positive correlations between familiarity and transparency across all idiomatic expressions ($r = .520$, $p < .001$), as well as for positive ($r = .492$, $p < .001$), neutral ($r = .546$, $p < .001$), and negative idiomatic expressions ($r = .451$, $p < .001$; see Fig. 9). Our findings align with previous studies on English idiomatic expressions (Libben & Titone, 2008) and Chinese four-character idiomatic expressions (Li et al., 2016). This suggests that when the figurative meanings of idiomatic expressions can be easily inferred from their literal meanings, language users can draw on their existing knowledge of the constituent words' literal meanings to enhance their familiarity with the idiomatic expressions. This also supports the dual-route model of idiomatic expression processing (Carrol & Conklin, 2014), which posits that the acquisition of idiomatic meanings occurs simultaneously with the activation of constituent word meanings.

However, our findings differ from those of Citron et al., (2016a, 2016b) on German idiomatic expressions, where familiarity and transparency were not significantly related. This discrepancy may arise from cross-linguistic differences in the acquisition of idiomatic expression. Specifically, speakers of languages like German may rely more on holistic storage and retrieval of idiomatic expressions during language acquisition, rather than analyzing the literal meanings of constituent words. Nonetheless, this is an exploratory explanation, and future research could further investigate this possibility.

Familiarity and ambiguity

Familiarity and ambiguity showed a significant negative correlation across all idiomatic expressions ($r = -.140$, $p = .025$; see Fig. 9). This suggests that the literal use of Chinese three-character idiomatic expressions may consume cognitive resources required for extracting and memorizing their figurative meanings, thereby reducing language users' familiarity with the idiomatic expressions (Cacciari & Tabossi, 1988). This finding contrasts with previous studies on English idiomatic expressions (Bonin et al., 2013; Cronk et al., 1993) and Chinese four-character idiomatic expressions (Li et al., 2016), where familiarity and ambiguity were positively correlated, indicating the unique characteristics of Chinese three-character idiomatic expressions. A possible explanation for the difference between Chinese four-character and three-character idiomatic expressions is that four-character idiomatic expressions, with their strong written style, are typically learned in formal educational settings. In such contexts, learners systematically study the origins and semantic evolution of idiomatic expressions, simultaneously learning their original literal usage and how they evolve into the derived figurative meanings. Thus, the literal usage facilitates learners' deeper understanding of and familiarity with their figurative meanings. In contrast, three-character idiomatic expressions are more colloquial and are acquired implicitly through real-life usage (Li et al., 2016; Zang et al., 2021), where learners grasp only the figurative meanings without focusing on the connection between the figurative and literal usage. As a result, increased literal usage of Chinese three-character idioms does not enhance language learners' familiarity with their figurative meanings
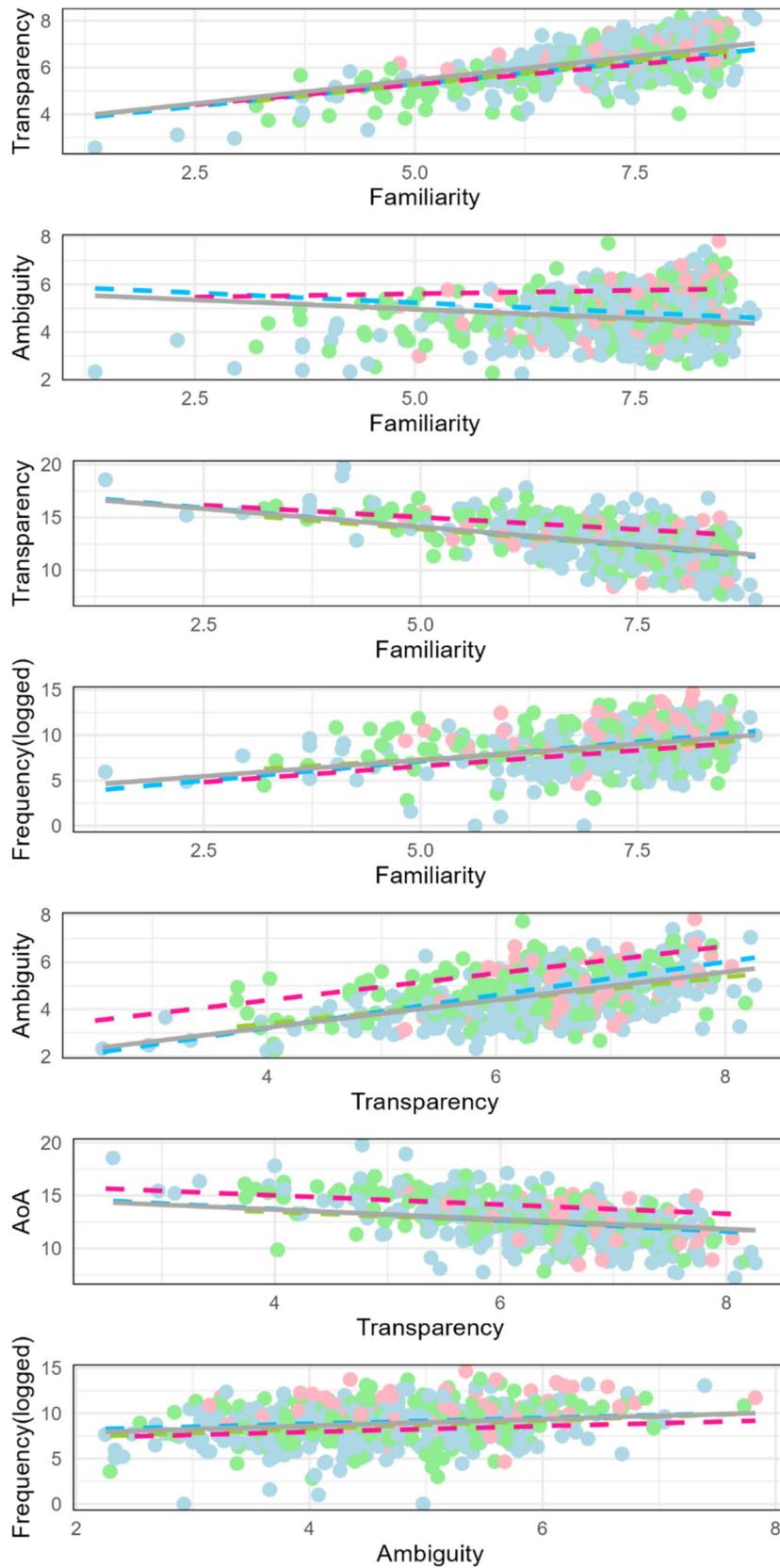
Fig. 9 Linear correlations between non-affective variables. Scatter points and/or linear fits (with other variables controlled) are shown for all idiomatic expressions (gray), as well as negative (blue), neutral (green), and positive idiomatic expressions (pink)

and may instead reduce opportunities to acquire these figurative meanings.

Familiarity and AoA

Familiarity and AoA showed significant negative correlations across all idiomatic expressions ($r = -.334$, $p < .001$), as well as in negative ($r = -.324$, $p < .001$) and neutral idiomatic expressions ($r = -.378$, $p < .001$; see Fig. 9). This indicates that idiomatic expressions learned at a younger age are more familiar to language speakers, especially for negative and neutral ones. Our findings align with those on French idiomatic expressions (Bonin et al., 2013) and Chinese four-character idiomatic expressions (Li et al., 2016). Such results can be explained from several perspectives. First, from a usage-based perspective, language acquisition involves exemplar-based learning (Ellis et al., 2013). Idiomatic expressions acquired earlier are encountered and used more frequently over time, leading to stronger entrenchment in memory. Second, the existence of a critical period for language acquisition (Snow & Hoefnagel-Höhle, 1978) suggests that idiomatic expressions encountered earlier, particularly before the critical period, tend to be mastered more effectively.

Familiarity and frequency

Our study revealed moderate to large positive correlations between familiarity and frequency across all idiomatic expressions ($r = .276$, $p < .001$), as well as negative ($r = .331$, $p < .001$), and neutral idiomatic expressions ($r = .221$, $p = .049$; see Fig. 9), which is consistent with previous studies on individual words (e.g., Tanaka-Ishii & Terada, 2011). Linguistic units that occur more frequently are easier to process and recall, thus becoming more familiar to language users (Schmidt & Seger, 2009). Our findings further support that language acquisition is usage-based and that language users are sensitive to the statistical properties of language units (Ellis et al., 2013), even for figurative Chinese three-character idiomatic expressions.

In familiarity ratings, participants considered both how often they encountered a linguistic unit (Citron et al., 2016a, 2016b) and how familiar they felt with it (Yao et al., 2017). The strong correlation between familiarity and frequency observed in our study suggests that subjective intuitions about the occurrence of linguistic units (how often), the direct explanation of familiarity (how familiar), and frequency are closely related constructs rather than independent ones, consistent with previous research (Siyanova-Chanturia & Spina, 2015; Tanaka-Ishii & Terada, 2011; Yi et al., 2023). However, this does not mean that subjective intuition, frequency, and direct explanation of familiarity are equally effective in predicting language processing behaviors

(Gernsbacher, 1984; Brysbaert & Cortese, 2011; Yao et al., 2017). Future research should simultaneously integrate these measures to improve predictive accuracy.

Transparency and ambiguity

Transparency showed strong positive correlations with ambiguity across all idiomatic expressions ($r = .407$, $p < .001$), as well as in negative ($r = .483$, $p < .001$) and neutral idiomatic expressions ($r = .364$, $p < .001$; see Fig. 9). This is consistent with findings on English (Bulkes & Tanner, 2017), Spanish (Gavilán et al., 2021), and Chinese four-character idiomatic expressions (Li et al., 2016). These results suggest that a stronger contribution of the literal meanings of constituent words to the figurative meaning of a Chinese three-character idiom is associated with greater plausibility of its literal interpretation (Gavilán et al., 2021; Li et al., 2016) and more frequent use of the literal interpretation.

Transparency and AoA

Transparency and AoA showed moderate negative correlations across all idiomatic expressions ($r = -.174$, $p = .012$) and within positive idiomatic expressions ($r = -.183$, $p = .049$; see Fig. 9). This indicates that positive idiomatic expressions whose figurative meanings are more easily inferred from their literal meanings are acquired earlier. Our findings align with previous norming studies on French idiomatic expressions (Bonin et al., 2013) and Chinese four-character idiomatic expressions (Li et al., 2016). The results support the developmental trajectory of MWE acquisition in children as reported in previous longitudinal studies (Wray, 2002). Specifically, children tend to acquire idiomatic expressions with decomposable and highly transparent meanings first, while idiomatic expressions with lower transparency are acquired later as their language experience and cognitive resources expand.

Ambiguity and frequency

Ambiguity and frequency showed moderate positive correlations in all idiomatic expressions ($r = .159$, $p = .010$; see Fig. 9). The results suggest that idiomatic expressions that are more likely to be used literally occur more often. This may be because the literal interpretation of an idiomatic expression has a broader range of use, with fewer contextual constraints compared to its figurative meaning. Additionally, the literal meanings of idiomatic expressions can be understood by language learners who do not grasp their figurative meanings, allowing idiomatic expressions with more literal usage to be used by a broader group of language speakers, thereby increasing their frequency.

General discussion

Our study is the first to provide subjective ratings of affective and non-affective psycholinguistic variables for a large set of Chinese three-character idiomatic expressions and to explore the relationships among affective variables (valence and arousal) and non-affective variables (i.e., familiarity, transparency, ambiguity, AoA, and frequency). First, we found that a large portion of Chinese three-character idiomatic expressions express negative and intense emotions. Additionally, over half of these idiomatic expressions have figurative meanings that can be easily inferred from the constituent words, are more likely to be used figuratively, are familiar to native Chinese speakers, and are typically acquired before the age of 12. This suggests that Chinese three-character idiomatic expressions play a crucial role in expressing negative and intense emotions, and the acquisition of their figurative usage is essential to language development.

We then examined the relationships between affective and non-affective variables across all idiomatic expressions as well as in negative, neutral, and positive idiomatic expressions. Valence and arousal exhibited a U-shaped curve, replicating findings from previous studies on single Chinese words (e.g., Chan & Tse, 2024; Ho et al., 2015; Lv et al., 2024; Xu et al., 2022; Yao et al., 2017) and idiomatic expressions in other languages (Citron et al., 2016a, 2016b; Gavilán et al., 2021; Morid & Sabourin, 2024). Such results suggest that more valenced idiomatic expressions were more arousing. In addition, arousal was more sensitive to the changes in negative information than to positive or neutral information, which supports the "approach–withdrawal" model of emotion processing (Robinson et al., 2004).

Further analysis of the relationship between affective and non-affective variables revealed that more transparent Chinese three-character idiomatic expressions were more valenced and more arousing. This suggests that the emotionality of constituent words can be activated and facilitate the expression of emotional content in figurative meanings (Citron et al., 2016a, 2016b; Gavilán et al., 2021), as proposed by the dual-route model of idiomatic expression processing (Carrol & Conklin, 2014). Moreover, idiomatic expressions more commonly used figuratively elicited stronger negative emotional responses (particularly for negative words) and more intense emotional reactions (especially for positive words) than those used more literally. This suggests that figurative usage of idiomatic expressions is more closely associated with encoding and conveying more negative and more intense emotional experiences (Yao et al., 2017). It was also found that idiomatic expressions with stronger negative emotionality were acquired earlier (i.e., earlier AoA), possibly because idiomatic expressions are crucial for expressing negative emotions (Kousta et al., 2009). Lastly, more frequent idiomatic expressions were rated more positively, which may be related to the tendency for participants to assign more positive rather than negative ratings to frequently used phrases (Robinson et al., 2004).

Regarding relations between non-affective variables, when the figurative meaning of Chinese three-character idiomatic expressions could be more easily inferred from the constituent words (i.e., higher transparency), language speakers were more familiar with the idiomatic expressions and acquired their figurative meanings earlier—especially for negative idiomatic expressions. This again fits with the dual-route model of idiomatic expression processing (Carrol & Conklin, 2014). Additionally, consistent with usage-based approaches (Ellis et al., 2013), we found that idiomatic expressions that were used more often in figurative senses (i.e., with low ambiguity), acquired earlier (i.e., with earlier AoA), and occurred more frequently—especially for negative and neutral ones—tended to be more efficiently stored and retrieved from memory (i.e., higher familiarity). Finally, idiomatic expressions that were more likely to be used literally (i.e., higher ambiguity) were more transparent and occurred more frequently.

Our results indicate that the affective information and the multilevel linguistic properties interact and share cognitive mechanisms in the usage of idiomatic expressions. We observed similar relationships between the affective and non-affective variables in negative and positive idiomatic expressions, which appears to support the model of motivated attention and affective states (Lang et al., 1998). However, compared to positive idiomatic expressions, stronger correlations between affective and non-affective variables were also found for negative and neutral idiomatic expressions. These results further suggest that positive, neutral, and negative idiomatic expressions engage distinct—yet related—physiological and cognitive mechanisms, supporting the automatic vigilance model of emotion (Estes & Verges, 2008). We also identified distinct features of Chinese three-character idiomatic expressions compared to individual words in expressing emotion. Chinese three-character idiomatic expressions are more likely to convey negative and stronger emotions, contrasting with prior research showing that individual Chinese words exhibit a dominance of positive and low-arousal emotionality (Chan & Tse, 2024; Ho et al., 2015; Lv et al., 2024). Moreover, unique affective and non-affective features of Chinese three-character idiomatic expressions in comparison to Chinese four-character idiomatic expressions, as well as idiomatic expressions from languages such as German, Spanish, and English, were also observed, which supports the irreplaceable role of three-character Chinese idiomatic expressions in language usage.

Conclusion

This study sought to provide descriptive norms for the affective and non-affective psycholinguistic properties of 500 Chinese three-character idiomatic expressions and to explore the interrelationships among the properties. In practical terms, our database provided valuable measures of various affective and non-affective psycholinguistic properties of Chinese three-character idiomatic expressions for future empirical studies. Furthermore, our study suggests that language instructors should emphasize teaching Chinese learners to use idiomatic expressions for conveying negative or emotionally charged statements, thereby mitigating potential conflicts in social interactions. Additionally, different affective and non-affective variables measure distinct but important aspects of idiomatic expressions (Li et al., 2016). For example, frequency is distinct from familiarity, and ambiguity is not the same as decomposability. Future research and teaching should, therefore, consider multiple features of idiomatic expressions simultaneously. Theoretically, our findings indicate that various types of information within idiomatic expressions, at both the word and phrase levels, interact during idiomatic expression processing and are shaped by accumulated language experience. This supports both the dual-route model (Carrol & Conklin, 2014) and usage-based approaches (Ellis et al., 2013). We also observed both similarities and differences in the relationships among affective and non-affective psycholinguistic variables across positive, neutral, and negative idiomatic expressions, providing evidence for both the model of motivated attention and affective states and the automatic vigilance model of emotion (Estes & Verges, 2008). Future studies could explore the potential coexistence of the two mechanisms. Furthermore, our study highlights the unique psycholinguistic features of Chinese three-character idiomatic expressions, distinguishing them from Chinese words and four-character idiomatic expressions. Finally, the cross-linguistic comparisons in our study highlight the need for future research to explore how cultural differences shape language processing.

Declarations

References

Bonin, P., Méot, A., & Bugaiska, A. (2013). Norms and comprehension times for 305 French idiomatic expressions. *Behavior Research Methods, 45*, 1259–1271. https://doi.org/10.3758/s13428-013-0331-4

Bradley, M. M., & Lang, P. J. (1999). *Affective norms for English words (ANEW): Instruction manual and affective ratings* (Vol. 30, No. 1, pp. 25–36). Technical report C-1, the Center for Research in Psychophysiology, University of Florida.

Briesemeister, B. B., Kuchinke, L., & Jacobs, A. M. (2011). Discrete emotion norms for nouns: Berlin affective word list (DENN–BAWL). *Behavior Research Methods, 43*, 441–448. https://doi.org/10.3758/s13428-011-0059-y

Brysbaert, M., & Cortese, M. J. (2011). Do the effects of subjective frequency and age of acquisition survive better word frequency norms? *Quarterly Journal of Experimental Psychology, 64*(3), 545–559. https://doi.org/10.1080/17470218.2010.503374

Bulkes, N. Z., & Tanner, D. (2017). "Going to town": Large-scale norming and statistical analysis of 870 American English idiomatic expressions. *Behavior Research Methods, 49*, 772–783. https://doi.org/10.3758/s13428-016-0747-8

Cacciari, C., & Tabossi, P. (1988). The comprehension of idiomatic expressions. *Journal of Memory and Language, 27*(6), 668–683. https://doi.org/10.1016/0749-596X(88)90014-9

Cain, K., Oakhill, J., & Lemmon, K. (2005). The relation between children's reading comprehension level and their comprehension of idiomatic expressions. *Journal of Experimental Child Psychology, 90*(1), 65–87. https://doi.org/10.1016/j.jecp.2004.09.003

Cain, K., Towse, A. S., & Knight, R. S. (2009). The development of idiom comprehension: An investigation of semantic and contextual processing skills. *Journal of Experimental Child Psychology, 102*(3), 280–298. https://doi.org/10.1016/j.jecp.2008.08.001

Carrol, G., & Conklin, K. (2014). Getting your wires crossed: Evidence for fast processing of L1 idiomatic expressions in an L2. *Bilingualism: Language and Cognition, 17*(4), 784–797.

Chan, Y. L., & Tse, C. S. (2024). Decoding the essence of two-character Chinese words: Unveiling valence, arousal, concreteness, familiarity, and imageability through word norming. *Behavior Research Methods*, 1–28. https://doi.org/10.3758/s13428-024-02437-w

Cieślicka, A. B., & Heredia, R. R. (2011). Hemispheric asymmetries in processing L1 and L2 idiomatic expressions: Effects of salience

and context. *Brain and Language, 116*(3), 136–150. https://doi.org/10.1016/j.bandl.2010.09.007

Cieślicka, A., & Heredia, R. R. (2017). How to save your skin when processing L2 idiomatic expressions: An eye movement analysis of idiom transparency and cross-language similarity among bilinguals. *Iranian Journal of Language Teaching Research*, *5*(3 (Special Issue)), 81–107. https://doi.org/10.30466/IJLTR.2017.20306

Citron, F. M., & Goldberg, A. E. (2014). Metaphorical sentences are more emotionally engaging than their literal counterparts. *Journal of Cognitive Neuroscience, 26*(11), 2585–2595. https://doi.org/10.1162/jocn_a_00654

Citron, F. M., Cacciari, C., Kucharski, M., Beck, L., Conrad, M., & Jacobs, A. M. (2016a). When emotions are expressed figuratively: Psycholinguistic and Affective Norms of 619 Idiomatic expressions for German (PANIG). *Behavior Research Methods, 48*, 91–111. https://doi.org/10.3758/s13428-015-0581-4

Citron, F. M., Güsten, J., Michaelis, N., & Goldberg, A. E. (2016b). Conventional metaphors in longer passages evoke affective brain response. *NeuroImage, 139*, 218–230. https://doi.org/10.1016/j.neuroimage.2016.06.020

Citron, F. M., Weekes, B. S., & Ferstl, E. C. (2014). How are affective word ratings related to lexicosemantic properties? Evidence from the Sussex Affective Word List. *Applied Psycholinguistics, 35*(2), 313–331. https://doi.org/10.1017/S0142716412000409

Cronk, B. C., Lima, S. D., & Schweigert, W. A. (1993). Idiomatic expressions in sentences: Effects of frequency, literalness, and familiarity. *Journal of Psycholinguistic Research, 22*, 59–82. https://doi.org/10.1007/BF01068157

de Zubicaray, G. I., Miozzo, M., Johnson, K., Schiller, N. O., & McMahon, K. L. (2012). Independent distractor frequency and age-of-acquisition effects in picture-word interference: FMRI evidence for post-lexical and lexical accounts according to distractor type. *Journal of Cognitive Neuroscience, 24*(2), 482–495. https://doi.org/10.1162/jocn_a_00141

Della Rosa, P. A., Catricalà, E., Vigliocco, G., & Cappa, S. F. (2010). Beyond the abstract—concrete dichotomy: Mode of acquisition, concreteness, imageability, familiarity, age of acquisition, context availability, and abstractness norms for a set of 417 Italian words. *Behavior Research Methods, 42*, 1042–1048. https://doi.org/10.3758/BRM.42.4.1042

Drew, P., & Holt, E. (1988). Complainable matters: The use of idiomatic expressions in making complaints. *Social Problems, 35*(4), 398–417. https://doi.org/10.2307/800594

Eilola, T. M., & Havelka, J. (2010). Affective norms for 210 British English and Finnish nouns. *Behavior Research Methods, 42*(1), 134–140. https://doi.org/10.3758/BRM.42.1.134

Ellis, N. C., O'Donnell, M. B., & Römer, U. (2013). Usage-based language: Investigating the latent structures that underpin acquisition. *Language Learning, 63*, 25–51. https://doi.org/10.1111/j.1467-9922.2012.00736.x

Estes, Z., & Adelman, J. S. (2008). Automatic vigilance for negative words in lexical decision and naming. *Emotion, 8*(4), 441–444. https://doi.org/10.1037/1528-3542.8.4.441

Estes, Z., & Verges, M. (2008). Freeze or flee? Negative stimuli elicit selective responding. *Cognition, 108*(2), 557–565. https://doi.org/10.1016/j.cognition.2008.03.003

Gavilán, J. M., Haro, J., Hinojosa, J. A., Fraga, I., & Ferré, P. (2021). Psycholinguistic and affective norms for 1,252 Spanish idiomatic expressions. *PLoS ONE, 16*(7), e0254484. https://doi.org/10.1371/journal.pone.0254484

Gernsbacher, M. A. (1984). Resolving 20 years of inconsistent interactions between lexical familiarity and orthography, concreteness, and polysemy. *Journal of Experimental Psychology: General, 113*(2), 256. https://doi.org/10.1037/0096-3445.113.2.256

Gyllstad, H., & Wolter, B. (2016). Collocational processing in light of the phraseological continuum model: Does semantic transparency

matter? *Language Learning, 66*(2), 296–323. https://doi.org/10.1111/lang.12143

Ho, S. M., Mak, C. W., Yeung, D., Duan, W., Tang, S., Yeung, J. C., & Ching, R. (2015). Emotional valence, arousal, and threat ratings of 160 Chinese words among adolescents. *PLoS ONE, 10*(7), e0132294. https://doi.org/10.1371/journal.pone.0132294

Huang, B. (2009). A dictionary of Chinese three-character idiomatic expressions (汉语惯用语词典). Commercial Press.

Kousta, S. T., Vinson, D. P., & Vigliocco, G. (2009). Emotion words, regardless of polarity, have a processing advantage over neutral words. *Cognition, 112*(3), 473–481. https://doi.org/10.1016/j.cognition.2009.06.007

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1998). Emotion and motivation: Measuring affective perception. *Journal of Clinical Neurophysiology, 15*(5), 397–408.

Larsen, R. J., Mercer, K. A., & Balota, D. A. (2006). Lexical characteristics of words used in emotional Stroop experiments. *Emotion, 6*(1), 62. https://doi.org/10.1037/1528-3542.6.1.62

Larsen, R. J., Mercer, K. A., Balota, D. A., & Strube, M. J. (2008). Not all negative words slow down lexical decision and naming speed: Importance of word arousal. *Emotion, 8*(4), 445–452. https://doi.org/10.1037/1528-3542.8.4.445

Li, D., Zhang, Y., & Wang, X. (2016). Descriptive norms for 350 Chinese idioms with seven syntactic structures. *Behavior Research Methods, 48*(4), 1678–1693. https://doi.org/10.3758/s13428-015-0692-y

Libben, M. R., & Titone, D. A. (2008). The multidetermined nature of idiom processing. *Memory & Cognition, 36*, 1103–1121. https://doi.org/10.1016/j.jneuroling.2013.10.001

Lv, Y., Ye, R., Ni, C., Wang, Y., Liu, Q., Zhou, Y., & Gao, F. (2024). ANCW: Affective norms for 4030 Chinese words. *Behavior Research Methods, 56*(5), 4893–4908. https://doi.org/10.3758/s13428-023-02226-x

Morid, M., & Sabourin, L. (2024). Affective and sensory-motor norms for idioms by L1 and L2 English speakers. *Applied Psycholinguistics, 45*(1), 138–155. https://doi.org/10.1017/S0142716423000504

Nordmann, E., & Jambazova, A. A. (2017). Normative data for idiomatic expressions. *Behavior Research Methods, 49*, 198–215. https://doi.org/10.3758/s13428-016-0705-5

R Core Team (2023). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Robinson, M. D., Storbeck, J., Meier, B. P., & Kirkeby, B. S. (2004). Watch out! That could be dangerous: Valence-arousal interactions in evaluative processing. *Personality and Social Psychology Bulletin, 30*(11), 1472–1484. https://doi.org/10.1177/0146167204266647

Schmidt, G. L., & Seger, C. A. (2009). Neural correlates of metaphor processing: The roles of figurativeness, familiarity and difficulty. *Brain and Cognition, 71*(3), 375–386. https://doi.org/10.1016/j.bandc.2009.06.001

Siyanova-Chanturia, A., & Martinez, R. (2015). The idiom principle revisited. *Applied Linguistics, 36*(5), 549–569. https://doi.org/10.1093/applin/amt054

Siyanova-Chanturia, A., & Spina, S. (2015). Investigation of native speaker and second language learner intuition of collocation frequency. *Language Learning, 65*(3), 533–562. https://doi.org/10.1111/lang.12125

Siyanova-Chanturia, A., Conklin, K., & Schmitt, N. (2011). Adding more fuel to the fire: An eye-tracking study of idiom processing by native and non-native speakers. *Second Language Research, 27*(2), 251–272. https://doi.org/10.1177/0267658310382068

Snow, C. E., & Hoefnagel-Höhle, M. (1978). The critical period for language acquisition: Evidence from second language learning. *Child Development*, 1114–1128. https://www.jstor.org/stable/1128751

Stadthagen-Gonzalez, H., Imbault, C., Pérez Sánchez, M. A., & Brysbaert, M. (2017). Norms of valence and arousal for 14,031 Spanish words. *Behavior Research Methods, 49*, 111–123.

Tabossi, P., Arduino, L., & Fanari, R. (2011). Descriptive norms for 245 Italian idiomatic expressions. *Behavior Research Methods, 43*, 110–123. https://doi.org/10.3758/s13428-010-0018-z

Tanaka-Ishii, K., & Terada, H. (2011). Word familiarity and frequency. *Studia Linguistica, 65*(1), 96–116. https://doi.org/10.1111/j.1467-9582.2010.01176.x

Vigliocco, G., Kousta, S. T., Della Rosa, P. A., Vinson, D. P., Tettamanti, M., Devlin, J. T., & Cappa, S. F. (2014). The neural representation of abstract words: The role of emotion. *Cerebral Cortex, 24*(7), 1767–1777. https://doi.org/10.1093/cercor/bht025

Warriner, A. B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods, 45*, 1191–1207. https://doi.org/10.3758/s13428-012-0314-x

Wilson, M. A., Ellis, A. W., & Burani, C. (2012). Age-of-acquisition affects word naming in Italian only when stress is irregular. *Acta Psychologica, 139*(3), 417–424. https://doi.org/10.1016/j.actpsy.2011.12.012

Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge University Press.

Xu, X., Li, J., & Chen, H. (2022). Valence and arousal ratings for 11,310 simplified Chinese words. *Behavior Research Methods, 54*(1), 26–41. https://doi.org/10.3758/s13428-021-01607-4

Yao, Z., Wu, J., Zhang, Y., & Wang, Z. (2017). Norms of valence, arousal, concreteness, familiarity, imageability, and context availability for 1,100 Chinese words. *Behavior Research Methods, 49*, 1374–1385. https://doi.org/10.3758/s13428-016-0793-2

Yi, W. (2018). Statistical sensitivity, cognitive aptitudes, and processing of collocations. *Studies in Second Language Acquisition, 40*(4), 831–856. https://doi.org/10.1017/S0272263118000141

Yi, W., & Zhong, Y. (2024). The processing advantage of multiword sequences: A meta-analysis. *Studies in Second Language Acquisition, 46*(2), 427–452. https://doi.org/10.1017/S0272263123000542

Yi, W., Lu, S., & Ma, G. (2017). Frequency, contingency and online processing of multiword sequences: An eye-tracking study. *Second Language Research, 33*(4), 519–549. https://doi.org/10.1177/0267658317708009

Yi, W., Man, K., & Maie, R. (2023). Investigating first and second language speaker intuitions of phrasal frequency and association strength of multiword sequences. *Language Learning, 73*(1), 266–300. https://doi.org/10.1111/lang.12521

Yi, W., Xu, H., & Man, K. (2025). Perception of emotion across cultures: Norms of valence, arousal, and sensory experience for 4923 Chinese words translated from English in Warriner et al. (2013). *Behavior Research Methods, 57*(1), 43. https://doi.org/10.3758/s13428-024-02580-4

Zang, C., Fu, Y., Bai, X., Yan, G., & Liversedge, S. P. (2021). Foveal and parafoveal processing of Chinese three-character idioms in reading. *Journal of Memory and Language, 119*, 104243. https://doi.org/10.1016/j.jml.2021.104243