Sensory and Affective Dimensions in Mandarin Monosyllabic Adjectives

Yimei Shao, Yu-Yin Hsu, Chu-Ren Huang

Sensory and Affective Dimensions in Mandarin Monosyllabic Adjectives

Yimei Shao, Yu-Yin Hsu, Chu-Ren Huang
Department of Language Science and Technology
The Hong Kong Polytechnic University, Hong Kong, SAR
{yimei.shao, yyhsu, churen.huang}@polyu.edu.hk

Abstract

This study examined how sensory modalities and affective dimensions jointly shape the semantics of Mandarin monosyllabic adjectives—a lexical class that is both perceptually rich and emotionally salient yet understudied in prior research. Drawing on embodied cognition theory, we integrated newly collected perceptual strength ratings (visual, auditory, haptic, gustatory, olfactory) for 165 adjectives with existing sensory norms (Chen et al., 2019) and affective ratings (valence, arousal; Peng et al., 2024), yielding a dataset of 298 items. Analyses revealed an asymmetrical sensory organization, with strong gustatory–olfactory coupling and relative independence of auditory imagery. Visual strength predicted more positive valence, whereas auditory and haptic strength predicted higher arousal, and modality exclusivity differentiated pleasantness from arousal. These results indicate that sensory experience systematically contributes to affective semantics, supporting embodied accounts of conceptual meaning. The study provides cross-linguistic insights into the perceptual grounding of emotion and offers a resource for future psycholinguistic and computational modeling.

1 Introduction

"We walked down the path to the well-house, attracted by the fragrance of the honeysuckle with which it was covered. Someone was drawing water and my teacher placed my hand under the spout. As the cool stream gushed over one hand she spelled into the other the word water, first slowly, then rapidly. I stood still, my whole attention fixed upon the motions of her fingers. Suddenly I felt a misty consciousness as of something forgotten—a thrill of returning thought; and somehow the mystery of language was revealed to me. I knew then that 'w-a-t-e-r' meant the wonderful cool something that was flowing over my hand. That living word awakened my soul, gave it light, hope, joy, set it free!"

— Helen Keller, The Story of My Life (1903)

This vivid account captures a profound insight: word meaning can arise directly from embodied sensory experience. It illustrates how perception and language intertwine—the tactile sensation of flowing water, the auditory rhythm of spelling, and the affective awakening that follows.

A single word can engage multiple senses and emotions at once. In Chinese, 清 (clear) may evoke visual imagery, the sound of flowing water, and a sense of calm; 香 (fragrant) blends olfactory impressions with pleasant affect; 刺耳 (piercing) conveys auditory harshness and negative emotion. Such examples show that perceptual and affective experiences are systematically intertwined in lexical representation—yet this interplay remains insufficiently examined in Chinese.

The idea that language both reflects and shapes human thought has long guided cognitive and linguistic theory (Hardin, 1993; Rosch, 1974; Vygotsky, 2000; Whorf, 1956). Early work on linguistic relativity argued that linguistic structure provides the scaffolding through which perceptual and conceptual distinctions are formed. This tradition highlights that meaning is not neutral but shaped by how language encodes experience.

Building upon this tradition, embodied cognition approaches extend the discussion by emphasizing that meaning is not only linguistically mediated but also grounded in direct sensory and emotional experience (Barsalou, 1999, 2008). Within the embodied cognition framework, meaning is grounded in the reactivation of sensory, motor, and emotional traces associated with a word's referent (Barsalou, 1999, 2008; Zwaan et al., 2002). Research on sensory dimensions has typically focused on five modalities—vision, audition, touch, taste, and smell—measured using modality strength and exclusivity norms (Lynott and Connell, 2009). In parallel, studies on affective dimensions have identified valence (pleasantness) and arousal (emotional

intensity) as key factors influencing lexical processing and conceptual organization (Bradley and Lang, 1999; Warriner et al., 2013). While each dimension has been widely studied for English, relatively few studies have addressed how sensory profiles systematically relate to emotional evaluations of words.

In Chinese, research on sensory norms has begun to emerge (e.g. Chen et al., 2019; Zhong and Ahrens, 2023), but it remains fragmented in lexical coverage, with limited attention to adjectives and minimal integration with affective ratings. Particularly, adjectives merit focused investigation: they are semantically compact, frequently encode perceptual features, and often carry strong emotional connotations. Moreover, cross-modal patterns such as gustatory–olfactory coupling or the independence of auditory modality, reported in English and other languages, have not been systematically tested in Mandarin adjectives.

Beyond modality strengths, embodied cognition posits that sensory experience forms a fundamental component of emotional representation—bodily sensations not only accompany emotions but constitute part of their experiential core (Damasio, 1994; Prinz, 2005). Yet it remains unclear to what extent different types of sensory experience contribute to emotional activation, and whether specific sensory channels (e.g., visual, auditory, haptic) are more strongly linked to affective dimensions such as pleasantness and arousal. Addressing this question helps clarify how emotion is grounded in distinct perceptual systems within the mental lexicon.

In addition to modality-specific effects, the overall structure of sensory engagement may also shape emotional meaning. Modality exclusivity, which captures the extent to which a word's meaning is concentrated in a single sensory channel (Lynott and Connell, 2009), provides a useful index of the breadth versus focus of embodied experience. Words that engage multiple senses may support richer, integrative simulations, whereas those confined to one dominant modality may evoke more vivid but narrower experiences. Examining how both sensory diversity and perceptual focus relate to emotional valence and arousal thus offers a deeper understanding of how embodied experience contributes to affective meaning in language.

The present study addresses these gaps by integrating sensory modality ratings and affective ratings into a unified dataset of 298 Chinese monosyllabic adjectives. Building on embodied cognition

theory, we pursue three main questions:

1. Interrelationships among sensory modalities – Do certain modalities tend to co-occur in Chinese adjectives, and how do these patterns compare to those reported in other languages?

2. Prediction of affective dimensions – To what extent can valence and arousal be systematically predicted from sensory modality strengths?

3. Role of modality exclusivity – Does perceptual focus on a single dominant modality enhance or diminish affective evaluations?

By focusing on monosyllabic adjectives and examining both perceptual and emotional dimensions simultaneously, this study contributes (a) novel empirical evidence on the embodied structure of Mandarin word meaning, (b) potential cross-linguistic comparisons with English and other languages, and (c) a lexical resource for psycholinguistic, computational, and applied linguistic research.

2 Literature Review

Embodied cognition theory offers a framework in which conceptual knowledge is grounded in sensory, motor, and emotional experiences, challenging the traditional amodal view of semantic representation (Barsalou, 1999, 2008). This perspective has been supported by evidence from cognitive neuroscience, such as the discovery of mirror neurons (Rizzolatti et al., 1996) and mechanisms described by Hebbian learning, which demonstrate how sensory, motor, and linguistic systems can become functionally integrated. In language processing, comprehension and production involve the simulation of sensory and motor experiences associated with linguistic input (Zwaan, 2003; Kogan et al., 2020). For example, understanding the word chair may engage visual imagery of its form, the motor schema of sitting, and even tactile sensations of its surface. Within Barsalou's Perceptual Symbol Systems account (Barsalou, 1999), these perceptual traces—encoded during prior experiences—are reactivated during linguistic tasks, forming an embodied route to meaning. This theoretical framework provides the foundation for investigating how sensory experiences are represented and re-engaged in lexical semantics.

Language often encodes modality-specific features such as vision (bright), audition (loud), touch

(rough), taste (sweet), and smell (fragrant). The concept of modality exclusivity quantifies the degree to which a word is associated with a single sensory modality versus multiple modalities (Lynott and Connell, 2009, 2013). In English, large-scale modality norms have revealed systematic relationships among modalities, such as the strong gustatory–olfactory coupling and the moderate co-occurrence of visual and haptic features (Lynott and Connell, 2009; Lynott et al., 2020; Speed and Brybaert, 2022). These norms have informed research in psycholinguistics, cognitive semantics, and computational modeling by providing quantitative measures of perceptual grounding. They have also been applied to examine processing effects, such as the modality-switch effect in sentence–picture verification tasks, and to explore how cross-modal integration supports conceptual organization.

Compared to English and other Indo-European languages, modality norm research in Mandarin is relatively recent. Chen et al. (2019) produced the first large-scale Mandarin modality exclusivity norms for monosyllabic and disyllabic adjectives, reporting perceptual strength ratings across five basic modalities and analyzing the influence of orthographic semantic radicals on modality judgments. While their dataset provides a crucial baseline, it did not include affective ratings, making it impossible to examine direct links between perceptual profiles and emotional dimensions. Zhong and Ahrens (2023) extended the scope by examining modality–emotion relationships in disyllabic nouns, finding that olfactory and interoceptive modalities were more emotionally charged, especially in arousal and absolute valence. However, their dataset had notable imbalances, with very few tactile and olfactory items (n = 8 for each), and their focus on nouns leaves open questions about whether similar patterns hold for adjectives—an important lexical class for perceptual and affective meaning.

Valence (pleasantness) and arousal (emotional intensity) are two fundamental affective dimensions shaping lexical processing (Bradley and Lang, 1999; Kousta et al., 2011; Warriner et al., 2013). These dimensions influence a range of cognitive processes, including lexical decision, memory, and semantic categorization. In Mandarin, several affective norms have been established (e.g., Yao et al., 2017), but very few studies have integrated them with sensory modality data. Yi et al. (2025) represents a rare attempt, examining sensory–emotion links in a large-scale dataset of disyllabic nouns translated from English norms (Warriner et al., 2013). Their results suggest that certain modalities, particularly olfactory and interoceptive, are more strongly associated with extreme valence and higher arousal.

Findings from English also indicate potential systematic modality–affect correspondences: gustatory and olfactory words tend to be more affectively rich and emotionally flexible (Winter, 2016), while auditory and haptic words are often linked to higher arousal (Lynott et al., 2020). However, other studies argue that these associations may be context-dependent or mediated by conceptual categories (Citron et al., 2014; Lynott and Connell, 2013). The extent to which such correspondences generalize across languages, and across word classes such as adjectives versus nouns, remains an open question.

The relationship between emotion and bodily sensation is not entirely separable: sensory experience serves as an essential, and in some cases sufficient, condition for emotional activation. Empirical evidence supports this link. Dagaev and Terushkina (2014) employed a property verification task to test whether emotional concepts involve embodied somatosensory components. Participants judged whether a given property applied to a concept (e.g., CLOWN–funny), with each trial pair consisting of a context trial and a target trial. A modality switch occurred when the two trials belonged to different modalities (e.g., somatosensory → emotional). Their results revealed an asymmetric effect between emotional and bodily sensation channels: when switching from emotional to somatosensory trials, no significant cost was observed; When switching from somatosensory to emotional trials, reaction times were significantly longer. The findings provide direct evidence for the embodied nature of emotional knowledge, indicating that emotional understanding involves the partial re-enactment of bodily states rather than purely symbolic operations.

Nevertheless, it remains unclear how different types of sensory experiences contribute to emotional responses, and whether their influence varies across sensory modalities. Moreover, in the modality-switch paradigm as mentioned above, researchers often select strongly unimodal words to control material consistency, using modality exclusivity as an index of the breadth and focus of perceptual engagement (e.g., Dagaev and Terushk-

ina, 2014; Vermeulen et al., 2007). This paradigm assumes that strongly unimodal words, compared with multimodal ones, more effectively activate specific sensory experiences, thereby producing faster reaction times. From an affective perspective, this raises an important question: Do the dimensionality and breadth of bodily sensory experience differentially influence emotional activation? In other words, the extent to which different words evoke emotions may depend on both their sensory modality (the type of embodied experience) and their modality exclusivity (the degree of perceptual specificity).

Despite increasing interest in the interface between sensory and affective dimensions, Mandarin research remains fragmented in three ways. First, there is no integrated dataset that combines balanced sensory modality ratings and affective ratings for adjectives—a lexical class rich in perceptual and emotional meaning. Second, modality interrelationships, such as gustatory–olfactory coupling or auditory independence, have not been systematically tested in Mandarin adjectives. Third, it remains unclear whether modality exclusivity—perceptual focus on a single dominant modality—predicts emotional valence and arousal in Chinese, and whether such effects align with or diverge from English findings. Addressing these gaps will advance embodied accounts of Chinese lexical semantics and provide a cross-linguistic perspective on how perception and emotion jointly shape word meaning.

3 Methodology

3.1 Participants

A total of 160 native speakers of Mandarin Chinese participated in the study. All participants were recruited in mainland China, held at least a college-level education, and reported normal or corrected-to-normal vision and hearing. The sample was balanced for gender (80 male, 80 female) with a mean age of 24.7 years (SD = 3.2, range = 18–35). None reported a history of neurological or psychiatric disorders. Written informed consent was obtained prior to participation, and the study adhered to ethical guidelines for human subjects research.

3.2 Materials

The stimulus set was constructed from two principal sources to maximize lexical coverage and ensure representativeness across sensory modalities.

1. Existing norms – Sensory modality ratings for 133 monosyllabic Chinese adjectives were obtained from the database developed by Chen et al. (2019), which provides perceptual strength ratings across five modalities: visual, auditory, haptic, gustatory, and olfactory.

2. Newly identified items – To expand the coverage of perceptually salient adjectives, we adopted the selection by Peng et al. (2024), who identified 165 additional monosyllabic adjectives through expert linguistic judgment. These items were specifically chosen for their potential sensory relevance, based on semantic transparency, frequency, and morphological characteristics.

Affective ratings (valence and arousal) for all items were obtained from Peng et al. (2024), ensuring that both sensory and emotional dimensions were available for each word and no scale conversions were necessary.

3.3 Procedure

For the 165 newly identified adjectives, sensory ratings were collected in an online questionnaire format. Participants rated each word on five sensory dimensions (visual, auditory, haptic, gustatory, olfactory) in response to the prompt: "To what extent do you think the word 清 (clear) can be used to describe the following sensory experiences?" Ratings were made on a 6-point Likert scale from 0 ("no association") to 5 ("strong association").

The 165 adjectives were evenly distributed across eight questionnaire lists, each containing 21 or 22 items. Assignment of words to lists was pseudo-randomized to balance modality coverage across lists, and participants were randomly assigned to lists. All instructions were presented in Mandarin, and no time limit was imposed on responses.

3.4 Data Preparation

For each word, mean sensory ratings across participants were calculated for all five modalities. The dominant modality of each word was defined as the modality with the highest mean rating. Modality exclusivity was computed following Lynott and Connell (2009) as the range of the five modality ratings divided by their sum, yielding values from

0 (completely multimodal) to 1 (completely uni-modal).

The final dataset comprised 298 adjectives (133 from Chen et al., 2019, ; 165 newly rated), each annotated with five sensory ratings and two affective ratings.

To visualize the overall distribution of ratings, Figure 1. presents histograms for all sensory and affective dimensions, with dashed lines marking the mean (red) and median (black) of each variable. These plots show that visual strength was rated highest on average, whereas gustatory and olfactory strengths were relatively low. Valence and arousal exhibited near-normal distributions, indicating a balanced coverage of positive–negative and low–high arousal adjectives.

3.5 Data Analysis

Data analysis proceeded in three stages. First, descriptive visualizations were generated to display the distribution of ratings across sensory and affective dimensions, enhancing transparency in the data.

Second, to examine the internal structure of sensory modality ratings, Spearman's rank correlation coefficients were computed among the five modalities (visual, auditory, haptic, gustatory, and olfactory). This approach allowed us to identify modality pairs that tend to co-occur in perceptual profiles, as well as those that remain relatively independent (Lynott and Connell, 2009). Significance levels were reported for each pairwise comparison, and the results were visualized using a correlation heatmap with significance markers.

Finally, two multiple linear regression models were constructed to assess how well affective evaluations could be predicted from sensory profiles. Valence and arousal served as the dependent variables, while the five sensory modalities and modality exclusivity were entered as predictors. To facilitate coefficient comparison, predictors were standardized prior to analysis. To ensure the validity and robustness of regression results, several standard diagnostic and correction procedures were employed. Multicollinearity was assessed using Variance Inflation Factors (VIF; Long and Ervin, 2000), which quantify how strongly predictors are linearly related. Heteroscedasticity was tested via the Breusch–Pagan test (Breusch and Pagan, 1979), and residual normality was examined using the Shapiro–Wilk test (Shapiro and Wilk, 1965). Influential observations were identi-

fied based on Cook's distance (Cook, 1977), which measures each data point's impact on model estimates. When heteroscedasticity was detected, HC3 heteroskedasticity-consistent standard errors (White, 1980; Long and Ervin, 2000) were applied to correct inferential tests. To further evaluate coefficient stability under potential outliers, robust regression using M-estimators (Huber, 1964) was conducted. These procedures provide a comprehensive evaluation of model reliability and ensure that the reported effects are not driven by violations of linear model assumptions. All analyses were
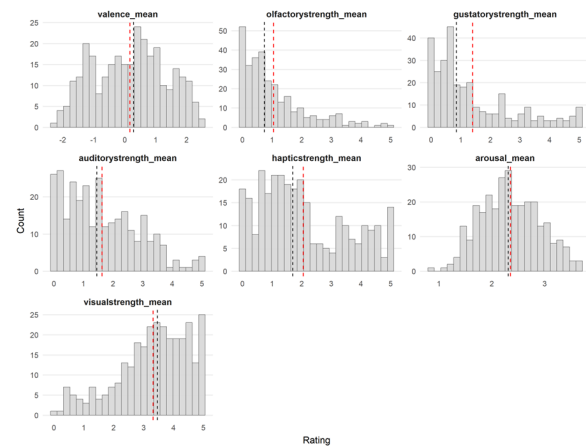

Figure 1: Distributions of sensory and affective ratings.

conducted in R (R Core Team, 2023) using the *stats*, *Hmisc* (Harrell, 2024), *car* (Fox and Weisberg, 2019), *lmtest* (Zeileis and Hothorn, 2002), sandwich (Zeileis, 2004), and MASS (Venables and Ripley, 2002) packages, with the significance level set at $\alpha = .05$.

4 Results

Figure 1 illustrates the distributions of sensory and affective ratings. Visual strength showed the highest overall ratings, gustatory and olfactory strengths were comparatively low, and both valence and arousal exhibited near-normal distributions, indicating balanced affective coverage and clear differentiation across sensory modalities.

4.1 Correlation analysis of sensory modality ratings

The correlation analysis among the five sensory modality ratings revealed both well-documented and novel patterns of interrelationships, offering a detailed picture of the sensory structure underlying Mandarin monosyllabic adjectives. As shown in Figure 2, which presents the full correlation matrix

using Spearman's rank coefficients (ρ), the most prominent positive association was found between gustatory and olfactory strength ($\rho = .81$, $p < .001$). This strong coupling is consistent with robust psycholinguistic and perceptual evidence that taste and smell are closely linked in both neural processing and conceptual representation (Auvray and Spence, 2008; Lynott and Connell, 2009). In practical terms, this means that adjectives highly associated with taste (e.g., 甜 "sweet") are also very likely to be strongly associated with smell, reflecting the multisensory integration that underlies flavor perception. The size of this correlation is comparable to, or even slightly higher than, that reported in English adjective norms (Lynott et al., 2009), suggesting that this gustatory–olfactory pairing may be a cross-linguistic and potentially universal feature of sensory semantics.

A weak but noteworthy positive correlation emerged between visual and haptic strength ($\rho = 0.11$, $p = .057$). Although this effect was marginally significant, its direction aligns with findings in English and other languages where visual attributes (e.g., 光滑 "smooth") often co-occur with tactile impressions. This pattern may reflect the fact that many surface properties—such as smoothness, roughness, or texture—are accessible both visually and through touch. The magnitude of this association in Mandarin appears weaker than in some English datasets (e.g., Speed and Brybaert, 2022), potentially reflecting differences in how adjectives encode surface descriptors in the two languages.

In contrast, visual strength exhibited significant negative correlations with both gustatory strength ($\rho = -0.30$, $p < .001$) and olfactory strength ($\rho = -0.29$, $p < .001$). This suggests a form of mutual exclusivity, where words highly tied to visual imagery tend to have weaker connections to taste and smell. Such negative associations may arise from the fact that visual descriptors in Mandarin often refer to properties of objects that are not inherently linked to flavor or scent (e.g., 亮 "bright," 暗 "dark"), and that taste- or smell-related words tend to evoke internal, bodily-oriented experiences rather than external, visually observable properties.

Auditory strength showed weak to moderate positive correlations with haptic ($\rho = .15$, $p = .011$) and olfactory strength ($\rho = .21$, $p < .001$), and a marginal association with gustatory ($\rho = .11$, $p = .053$). These patterns indicate occasional overlap between sound-related and other sensory imagery, perhaps in words depicting intensity or dynamic

qualities (e.g., 响 "loud"). Nevertheless, the overall independence of auditory ratings from most other modalities supports the idea that auditory experience relies on distinct temporal and acoustic representations (Farmer et al., 2006; Lynott and Connell, 2009).

A weaker but statistically significant positive correlation was observed between haptic and olfactory strength ($\rho = 0.14$, $p = .017$). Although the effect size is small, it is intriguing because it suggests occasional overlap between tactile and olfactory imagery. One possible explanation is that some adjectives describe materials or substances (e.g., 腥 "fishy," 滑 "slippery") that are jointly characterized by both texture and smell.

Overall, the results reveal a perceptual organization in which gustatory–olfactory coupling forms a tightly integrated cluster, visual–haptic links remain marginal, auditory is largely distinct, and visual–taste/smell interactions are mutually exclusive. These findings establish a structural baseline for exploring how specific sensory channels contribute to affective meaning in subsequent analyses.
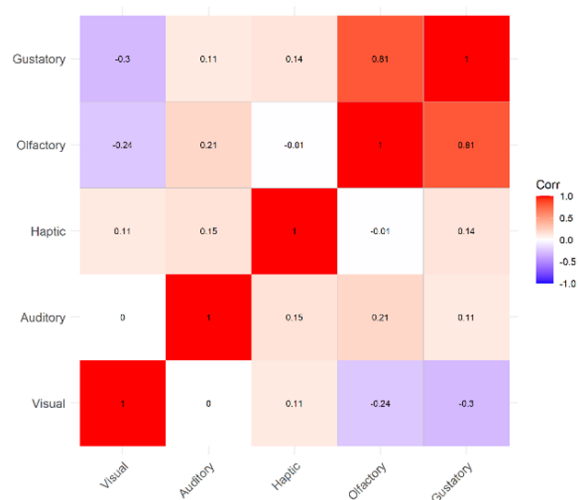

Figure 2: Correlation matrix of sensory modalities.

4.2 Regression analysis of sensory modality ratings and emotional dimensions

To examine how sensory modality profiles are associated with affective meaning, two multiple regression models were fitted, with valence and arousal as outcome variables, respectively. All predictors were standardized to facilitate coefficient comparison.

Comprehensive diagnostic checks were conducted to evaluate model validity. Variance In-

flation Factors (VIFs) were all below 3.3, suggesting no substantial multicollinearity among the predictors. The Breusch–Pagan test indicated heteroscedasticity in the valence model (BP = 37.80, $p < .001$), but not in the arousal model (BP = 4.62, $p = .59$). The Shapiro–Wilk test suggested slight deviations from normality for residuals (valence: W = 0.98, $p < .001$; arousal: W = 0.99, $p = .011$), though Q–Q plots confirmed that the distributions were approximately normal overall.

Inspection of Cook's distance identified 10 potential influence points in the valence model and 19 in the arousal model (Cook's D > 0.0135). Removing these points did not change the pattern or direction of the results; thus, all data points were retained.

To address potential heteroscedasticity and influential observations, we additionally computed HC3 robust standard errors and performed M-estimator regressions (MASS::rlm). The results were consistent across estimation methods, indicating that the observed relationships were statistically reliable and not driven by outliers.

The valence regression model was statistically significant (F(6, 290) = 2.60, $p = .018$; adjusted R^2 = .031), indicating that sensory modality ratings collectively explained about 3% of the variance in valence evaluations.

After applying HC3 robust standard errors, visual strength remained positively associated with valence ($\beta = 0.20$, $p = .005$), suggesting that adjectives evoking strong visual experiences tend to be perceived as more pleasant. In contrast, modality exclusivity showed a marginally negative association ($\beta = –0.19$, $p = .07$), indicating that adjectives tied to a single sensory channel may be evaluated as slightly less pleasant. Other modalities (auditory, haptic, gustatory, olfactory) did not show reliable associations (all $ps > .10$).

The arousal regression model accounted for a larger portion of variance (F(6, 290) = 5.80, $p < .001$; adjusted R^2 = .089). Under the HC3 correction, auditory strength ($\beta = 0.18$, $p < .001$), haptic strength ($\beta = 0.11$, $p = .007$), and modality exclusivity ($\beta = 0.18$, $p < .001$) each showed significant positive associations with arousal. This pattern suggests that adjectives evoking auditory and tactile experiences, or dominated by one sensory channel, are more likely to be judged as activating or intense.

The M-estimator regression confirmed the robustness of these findings: auditory ($\beta = 0.18$, $p <$ .001), haptic ($\beta = 0.11$, $p = .007$), and modality exclusivity ($\beta = 0.19$, $p < .001$) remained significant.

Taken together, these analyses reveal that valence is modestly associated with visual and multimodal characteristics, whereas arousal is more strongly linked to auditory and haptic experiences, as well as modality dominance. Importantly, all interpretations are correlational in nature and should not be construed as evidence of direct causal influence between sensory and affective dimensions.

5 Discussion

The present study set out to examine how sensory and affective dimensions jointly shape the semantics of Mandarin monosyllabic adjectives, addressing three key questions: (1) the interrelationships among sensory modalities, (2) the extent to which valence and arousal can be predicted from sensory profiles, and (3) the role of modality exclusivity in affective evaluations. By integrating perceptual ratings across five modalities (visual, auditory, haptic, gustatory, olfactory) with affective ratings (valence, arousal) in a unified dataset, we provide a systematic account of how perceptual experience contributes to emotional word meaning in a lexical class that is both perceptually rich and affectively salient.

The correlation analysis revealed a perceptual organization broadly consistent with cross-linguistic findings, yet with certain language-specific nuances. As expected, gustatory and olfactory modalities formed a tightly coupled cluster, reflecting their shared experiential and neural basis in flavor perception (Lynott and Connell, 2009; Lynott et al., 2020). This strong coupling suggests that Mandarin adjectives related to taste often simultaneously evoke olfactory sensations, mirroring the multisensory integration that underlies flavor perception more generally.

A weak but noteworthy overlap was observed between visual and haptic modalities, implying that surface-related features—such as texture, smoothness, or glossiness—may be accessible both visually and through touch, though to a lesser degree than reported in English norms (Speed and Brybaert, 2022). This cross-linguistic difference suggests that Mandarin adjectives may encode surface properties more functionally or contextually rather than visually.

Interestingly, visual imagery tended to exclude internal sensations such as taste and smell, implying a conceptual divide between outwardly observable properties and inward bodily experiences. This asymmetry aligns with the general observation that adjectives describing visual appearance (e.g., 亮 "bright", 暗 "dim") often capture spatially external, object-based qualities, while taste and smell adjectives (e.g., 酸 "sour", 臭 "smelly") pertain to embodied, interoceptive sensations. Meanwhile, auditory modality remained largely independent from the others, reinforcing the idea that auditory representation is grounded in temporally dynamic and acoustically distinct processing. Overall, Mandarin sensory adjectives exhibit a structured but asymmetrical organization—one dominated by taste–smell integration, with vision, touch, and sound occupying more specialized and partially independent roles.

Regression analyses revealed that different sensory channels contribute unequally to emotional meaning. Visual imagery emerged as a modest yet consistent predictor of pleasantness, suggesting that adjectives evoking vivid visual experiences tend to be evaluated as more pleasant. This tendency may reflect cultural and cognitive associations in Mandarin between brightness, clarity, and aesthetic harmony and positive emotional tone, consistent with the symbolic role of light and visual clarity in Chinese idioms, poetry, and moral metaphors.

In contrast, arousal was primarily linked to auditory and haptic modalities, consistent with the idea that sounds and physical sensations carry strong activation potential. This finding highlights how certain sensory experiences—particularly those involving abrupt or proximal stimuli—tend to evoke stronger physiological engagement. Adjectives describing auditory intensity (e.g., 吵 "noisy") or tactile extremity (e.g., 滑 "slippery") appear to embody dynamic or forceful qualities that naturally enhance arousal.

Notably, gustatory and olfactory modalities showed no reliable associations with either valence or arousal, diverging from English results where taste and smell often carry strong emotional valence (Winter, 2016). This discrepancy may be attributable to lexical and cultural factors: Mandarin adjectives describing taste and smell are fewer in number and often restricted to physical or contextual descriptions (e.g., describing food or environment) rather than abstract emotional evaluation, thereby limiting their contribution to affective meaning at the lexical level.

Modality exclusivity further clarified how the breadth of sensory experience shapes affective interpretation. Adjectives that evoke multiple sensory modalities were associated with higher pleasantness, whereas those tied to a single dominant modality were perceived as more activating. This dual pattern indicates that sensory focus intensifies emotional arousal—likely by engaging vivid and detailed mental simulations—but may narrow the hedonic scope, leading to less positive evaluations overall.

This finding complements embodied cognition accounts (Barsalou, 1999, 2008) by showing that the extent of sensory engagement, not merely the presence of sensory content, modulates affective meaning. Words restricted to one sensory channel may evoke more vivid and focused simulations, enhancing arousal; in contrast, multimodal words integrate diverse perceptual cues, producing richer, more balanced experiential representations that align with positive affect.

From an embodied cognition perspective, these findings reinforce that conceptual representation is grounded in perceptual and affective systems, with different modalities contributing unequally to distinct emotional dimensions. Vision appears more closely tied to hedonic evaluation (valence), while audition and touch contribute to physiological activation (arousal). This division may reflect underlying neurocognitive specializations: the visual system is linked to appraisal and aesthetic judgment, whereas auditory and tactile systems are connected to proximity, urgency, and survival-relevant reactions.

Furthermore, the role of modality exclusivity underscores the importance of sensory diversity and focus in shaping emotional meaning. Words evoking multiple modalities may reflect more integrative and experiential processing, while modality-specific words capture sharper but narrower affective tones. Taken together, these findings suggest that sensory profiles—rather than isolated modalities—form the foundation for affective meaning in lexical semantics.

6 Conclusion

The present study examined how sensory modalities and affective dimensions jointly shape the semantics of Mandarin monosyllabic adjectives. Integrating newly collected sensory ratings with affective evaluations, we constructed a dataset of 298 adjectives for the first systematic analysis of sensory–affective mappings in this lexical class. The results revealed an asymmetrical sensory organization: gustatory and olfactory modalities were strongly coupled, visual–haptic overlap was weak, and auditory imagery remained largely independent. Regression analyses showed that visual strength predicted more positive valence, whereas auditory and haptic strength predicted higher arousal. Modality exclusivity further influenced affective meaning—multimodal adjectives were more pleasant, while unimodal ones were more arousing. These findings support embodied cognition accounts, showing that sensory experience systematically contributes to affective semantics, with both universal and language-specific patterns in Mandarin.

7 Limitations and future directions

While the present study provides a comprehensive dataset and robust statistical analyses, several limitations should be noted. First, although the dataset covers a large number of monosyllabic adjectives, it does not include disyllabic adjectives or other word classes that may exhibit different modality–affect patterns. Future work could broaden the scope to these forms and compare results across lexical categories. Second, the study relies on explicit ratings, which capture conscious associations but may not fully reflect automatic or context-dependent processing. Incorporating time-sensitive methods such as EEG, eye-tracking, or priming paradigms could reveal how modality–affect mappings unfold during real-time comprehension. Third, cultural and contextual factors were not directly examined; cross-linguistic comparisons, for instance between Mandarin and English, could clarify universal versus language-specific effects.

Beyond theoretical contributions, the findings have implications for language education, computational linguistics, and affective computing. Understanding which sensory modalities align with positive or high-arousal meanings could inform vocabulary teaching and sentiment modeling, while the dataset may aid multimodal systems in interpreting and generating emotionally nuanced language. In sum, this study shows that sensory and affective dimensions are systematically linked in Mandarin adjectives, reflecting both universal and language-specific patterns. The integrated sensory–affective norms offer a theoretical contribution to embodied cognition and a practical resource for future psycholinguistic and cross-linguistic research.

References

Malika Auvray and Charles Spence. 2008. The multisensory perception of flavor. *Consciousness and Cognition*, 17(3):1016–1031.

Lawrence W. Barsalou. 1999. Perceptions of perceptual symbols. *Behavioral and Brain Sciences*, 22(4):637–660.

Lawrence W. Barsalou. 2008. Grounded cognition. *Annual Review of Psychology*, 59:617–645.

Margaret M Bradley and Peter J Lang. 1999. Affective norms for english words (anew): Instruction manual and affective ratings. Technical report, Technical report C-1, the center for research in psychophysiology . . . .

T. S. Breusch and A. R. Pagan. 1979. A simple test for heteroscedasticity and random coefficient variation. *Econometrica*, 47(5):1287–1294.

I-H Chen, Q Zhao, Y Long, Q Lu, and C-R Huang. 2019. Mandarin Chinese modality exclusivity norms. *PLOS ONE*, 14(2):e0211336.

Francesca M.M. Citron, Brendan S. Weekes, and Evelyn C. Ferstl. 2014. Arousal and emotional valence interact in written word recognition. *Language, Cognition and Neuroscience*, 29(10):1257–1267.

R. Dennis Cook. 1977. Detection of influential observation in linear regression. *Technometrics*, 19(1):15–18.

Nikolay I. Dagaev and Yulia I. Terushkina. 2014. Conceptual knowledge of emotions includes somatosensory component: Evidence from modality-switch cost effect. *Journal of Cognitive Psychology*, 26(3):322–332.

A. R. Damasio. 1994. *Descartes' Error: Emotion, Reason, and the Human Brain*. G. P. Putnam's Sons, New York, NY.

Thomas A. Farmer, Morten H. Christiansen, and Padraic Monaghan. 2006. Phonological typicality influences on-line sentence comprehension. *Proceedings of the National Academy of Sciences*, 103(32):12203–12208.

John Fox and Sanford Weisberg. 2019. *An R Companion to Applied Regression*, third edition. Sage, Thousand Oaks, CA.

Curtis Hardin. 1993. The influence of language on thought. *Social Cognition*, 11(3):277–308.

Frank E. Harrell, Jr. 2024. *Hmisc: Harrell Miscellaneous*. R package version 5.1-2.

Peter J Huber. 1964. Robust estimation of a location parameter: Annals mathematics statistics, 35. *Ji, S., Xue, Y. and Carin, L.(2008),'Bayesian compressive sensing', IEEE Transactions on signal processing*, 56(6):2346–2356.

Boris Kogan, Enrique García-Marco, Agustina Birba, Camila Cortés, Margherita Melloni, Agustín Ibáñez, and Adolfo M. García. 2020. How words ripple through bilingual hands: Motor-language coupling during l1 and l2 writing. *Neuropsychologia*, 146:107563.

S.-T. Kousta, G. Vigliocco, D. P. Vinson, M. Andrews, and E. Del Campo. 2011. The representation of abstract words: Why emotion matters. *Journal of Experimental Psychology: General*, 140(1):14–34.

J. Scott Long and Laurie H. Ervin. 2000. Using heteroscedasticity consistent standard errors in the linear regression model. *The American Statistician*, 54(3):217–224.

Dermot Lynott and Louise Connell. 2009. Modality exclusivity norms for 423 object properties. *Behavior research methods*, 41(2):558–564.

Dermot Lynott and Louise Connell. 2013. Modality exclusivity norms for 400 nouns: The relationship between perceptual experience and surface word form. *Behavior research methods*, 45(2):516–526.

Dermot Lynott, Louise Connell, Marc Brysbaert, James Brand, and James Carney. 2020. The lancaster sensorimotor norms: multidimensional measures of perceptual and action strength for 40,000 english words. *Behavior research methods*, 52(3):1271–1291.

Cheng Peng, Xu Xu, and Zhen Bao. 2024. Sentiment annotations for 3827 simplified chinese characters. *Behavior Research Methods*, 56(2):651–666.

Jesse J. Prinz. 2005. Passionate thoughts. In Rolf Zwaan and Diane Pecher, editors, *The Grounding of Cognition: The Role of Perception and Action in Memory, Language, and Thinking*, pages 93–114. Cambridge University Press, Cambridge, MA.

R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Giacomo Rizzolatti, Luciano Fadiga, Vittorio Gallese, and Leonardo Fogassi. 1996. Premotor cortex and the recognition of motor actions. *Cognitive Brain Research*, 3(2):131–141. Mental representations of motor acts.

Eleanor Rosch. 1974. Linguistic relativity. In Albert Silverstein, editor, *Human communication: Theoretical perspectives*, pages 95–121. Lawrence Erlbaum Associates, Hillsdale, NJ.

S. S. Shapiro and M. B. Wilk. 1965. An analysis of variance test for normality (complete samples). *Biometrika*, 52(3/4):591–611.

Laura J Speed and Marc Brybaert. 2022. Dutch sensory modality norms. *Behavior research methods*, 54(3):1306–1318.

W. N. Venables and B. D. Ripley. 2002. *Modern Applied Statistics with S*, fourth edition. Springer, New York. ISBN 0-387-95457-0.

Nicolas Vermeulen, Paula M Niedenthal, and Olivier Luminet. 2007. Switching between sensory and affective systems incurs processing costs. *Cognitive Science*, 31(1):183–192.

L. S. Vygotsky. 2000. *Thought and Language*. MIT Press, Cambridge, MA.

Amy Beth Warriner, Victor Kuperman, and Marc Brysbaert. 2013. Norms of valence, arousal, and dominance for 13,915 english lemmas. *Behavior research methods*, 45(4):1191–1207.

Halbert White. 1980. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica*, 48(4):817–838.

Benjamin Lee Whorf. 1956. *Language, thought, and reality*. MIT Press, Cambridge, MA.

Bodo Winter. 2016. Taste and smell words form an affectively loaded and emotionally flexible part of the english lexicon. *Language, Cognition and Neuroscience*, 31(8):975–988.

Zhao Yao, Jia Wu, Yanyan Zhang, and Zhenhong Wang. 2017. Norms of valence, arousal, concreteness, familiarity, imageability, and context availability for 1,100 chinese words. *Behavior research methods*, 49(4):1374–1385.

Wei Yi, Haitao Xu, and Kaiwen Man. 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.

Achim Zeileis. 2004. Econometric computing with hc and hac covariance matrix estimators. *Journal of Statistical Software*, 11(10):1–17.

Achim Zeileis and Torsten Hothorn. 2002. Diagnostic checking in regression relationships. *R News*, 2(3):7–10.

Yin Zhong and Kathleen Ahrens. 2023. The emotion code in sensory modalities. In *Chinese Lexical Semantics*, pages 183–192, Cham. Springer Nature Switzerland.

Rolf A. Zwaan. 2003. The immersed experiencer: Toward an embodied theory of language comprehension. volume 44 of *Psychology of Learning and Motivation*, pages 35–62. Academic Press.

Rolf A Zwaan, Robert A Stanfield, and Richard H Yaxley. 2002. Language comprehenders mentally represent the shapes of objects. *Psychological science*, 13(2):168–171.

A   Supplementary Material

The data for the experiment is available at `https://osf.io/yzuhr/`