The Sensorimotor Norms for the Chinese Classifiers

Yimei Shao, Yu-Yin Hsu, Chu-Ren Huang
Department of Language Science and Technology, The Hong Kong Polytechnic University
PolyU Hung Hom Bay Campus, 8 Hung Lok Road, Hung Hom, Kowloon, Hong Kong
yimei.shao@connect.polyu.hk; yu-yin.hsu@polyu.edu.hk; chu-ren.huang@polyu.edu.hk

Abstract

Sensorimotor information plays a crucial role in the conceptual representation of linguistic knowledge. While previous studies have established sensorimotor norms for nouns and adjectives, little is known about how Chinese numeral classifiers encode perceptual and action-based experiences. The present study constructs the first large-scale sensorimotor norms for Chinese classifiers, collecting perceptual and action ratings for 357 classifiers from 288 native Chinese speakers. Participants evaluated each classifier along six perceptual modalities (vision, hearing, taste, smell, touch, and interoception) and five action effectors (foot/leg, hand/arm, mouth/throat, head, and torso). The resulting dataset provides detailed sensorimotor profiles for each classifier and reveals systematic mappings between classifier semantics and embodied dimensions. The findings demonstrate that Chinese classifiers are not purely syntactic markers but encode distinct sensorimotor features grounded in perceptual and motor systems, highlighting the embodied foundation of the classifier system and offering valuable resources for future psycholinguistic and computational modelling studies of Chinese semantics.

Keywords: Embodiment, Chinese Classifiers, Sensorimotor Norms

1. Introduction

Human beings conceptualize and categorize the world in diverse ways, and these conceptual distinctions are often encoded through language. Different languages adopt distinct linguistic devices to classify entities. For instance, Indo-European languages typically employ grammatical gender to distinguish nouns, whereas Mandarin Chinese relies on numeral classifiers—morphemes that appear between numerals and nouns—to categorize referents.

In Mandarin, classifiers typically occur in the structure "*numeral + classifier + noun*" (e.g., 一支笔 *yi4 zhi1 bi3* 'one CL-pen', 两把椅子 *liang3 ba3 yi3zi* 'two CL-chair'). The relation between classifiers and nouns is systematic rather than arbitrary. From a combinatorial perspective, this relationship exhibits one-to-many and many-to-one mappings: a single classifier can apply to a class of semantically related nouns (e.g., 支 *zhī* for long, slender objects such as *pens*, *guns*, and *candles*), while a single noun may co-occur with multiple classifiers that highlight different perceptual or functional aspects (e.g.,一朵花 *yi4 duo3 hua1* 'a flower', 一束花 *yi2 shu4 hua1* 'a bunch of flowers', 一盆花 *yi4 pen2 hua1* 'a pot of flowers').

The mechanisms underlying classifier–noun selection have been debated along two main lines: syntactic and semantic accounts. From a syntactic perspective, classifier usage is governed by grammatical structure or morphosyntactic agreement features (e.g., Cheng & Sybesma, 1999), and some neurocognitive findings have been interpreted as supporting this view—for example, ERP studies showing N400 in response to classifier violations have been taken to indicate structural processing costs (Chan, 2019).

In contrast, the semantic-driven hypothesis posits that classifier choice primarily depends on the semantic properties of the modified noun (e.g., Wu & Bodomo, 2009). This view has been strongly supported by linguistic (Tai, 1994; Huang & Ahrens, 2003) and psycholinguistic (Frankowsky et al., 2022; Wang et al., 2025; Wang et al., 2025a, 2025b) analyses showing that appropriate classifier selection relies on features such as animacy, shape, size, and material/texture. For instance, 支 *zhi1* tends to occur with long, slender artifacts (e.g., *pens*, *arrows*), whereas 把 *ba3* is used for manipulable objects that afford hand action (e.g., *knives*, *chairs*).

The semantic-driven hypothesis has also been supported by recent neurocognitive studies. For example, Wang et al. (2025) investigated whether the visual shape semantics encoded by shape classifiers are engaged during production of classifier–noun phrases and at what processing stage. Using a blocked cyclic naming paradigm with native speakers, they orthogonally manipulated classifier congruency (same vs. different classifier across items within a block) and classifier shape similarity (similar vs. dissimilar). Electrophysiologically, shape-dissimilar relative to shape-similar blocks produced more negative amplitudes beginning ~300ms (N400-like), localizing both effects to post-lemma, lexical stages. The authors conclude that classifiers are automatically activated during lemma retrieval in Chinese noun phrase production and that the visual shape features embedded in classifiers are processed at the lexical level.

However, most existing studies focus on isolated semantic dimensions—such as animacy or shape—without offering a systematic characterization of the entire classifier system. Although qualitative research (Ahrens & Huang,

Proceedings of the Fifteenth Language Resources and Evaluation Conference (LREC 2026), pages 11439–11450
11–16 May 2026. ©ELRA Language Resources Association (ELRA), 2026

2016) has provided detailed descriptions of the specific usage contexts of individual classifiers, there is still a lack of quantitative characterization of the classifier system as a whole.

According to embodied cognition theory, language is not an abstract, amodal symbol system but is grounded in the perceptual and motor systems of the human body (Barsalou, 1999, 2008; Zwaan & Madden, 2005). Comprehending linguistic meaning involves partial reactivation of the same sensory and motor representations that arise during perception and action. From this view, the semantic distinctions encoded by classifiers should reflect embodied experiences of perception and interaction with the world. If classifiers indeed derive their meanings from sensorimotor experience, their semantic profiles ought to be quantifiable along perceptual and effector-based dimensions.

Furthermore, embodied accounts emphasize that conceptual representations are graded rather than categorical. Instead of activating discrete symbolic features, conceptual processing is assumed to involve varying degrees of sensorimotor simulation depending on the experiential associations of a given concept (Myachykov et al., 2014; Truman & Kutas, 2024). As a result, different linguistic units may evoke perceptual and motor information to different extents. Extending this view to classifier systems, classifiers should not be treated as encoding fixed semantic categories but rather as exhibiting graded sensorimotor profiles reflecting the perceptual and interactional properties of the entities they typically classify.

Recent advances in sensorimotor norming provide an effective method for quantifying embodied semantics. Norming studies collect human ratings on multiple sensory (e.g., visual, auditory, tactile, gustatory, olfactory, interoception) and motor dimensions (e.g., hand, leg, mouth, head, trunk), thereby capturing fine-grained sensorimotor profiles of lexical items (Lynott et al., 2020; Speed & Majid, 2020). This approach has proven fruitful for nouns (Zhong et al., 2022), adjectives (Chen et al., 2019), and verbs (Britton, 2024) in Mandarin, enabling large-scale, continuous, and comparable measurement of embodied content. Yet, no existing database has extended this framework to grammatical categories such as classifiers, despite their clear conceptual grounding.

The present study addresses this gap by developing a comprehensive sensorimotor norm database for Mandarin classifiers, profiling each classifier across multiple sensory and motor dimensions. This project pursues two major goals:

1. To present a quantitative overview of the semantic landscape of the Chinese classifier system from an embodied perspective; and
2. To establish a reliable, fine-grained database that supports future behavioral, neurocognitive, and computational research on classifier processing and embodied semantics.

Through this endeavor, the study aims to bridge grammatical theory and embodied cognition, offering data-driven insights into how grammatical knowledge in Chinese reflects human perceptual and motor experience.

2. This study

Building upon the framework of embodied cognition, the present study aims to establish a large-scale sensorimotor norm for Chinese classifiers. This study addresses two central research questions (RQs):

RQ1: To what extent do the semantics of Chinese classifiers reflect their dominant sensory and motor dimensions? If classifier meanings are grounded in embodied experience, we expect a systematic correspondence between their semantic categories and sensorimotor strengths.

RQ2: What are the overall patterns of sensorimotor distribution across the classifier system? We hypothesize that visual and manual dimensions will be the most prominent, given the perceptually and functionally rich nature of classifiers that categorize entities based on shape, size, or manipulability.

3. Methodology

3.1 Participants

We recruited 523 participants (244 males) from various universities and colleges in mainland China. Their ages ranged from 17 to 52 years (Mean = 22.24, *SD* = 9.49). They were all native speakers of Mandarin Chinese. All participants reported themselves to be mentally and physically healthy. All participants gave informed consent before the experiment and were explicitly informed that their participation was voluntary and that they could withdraw at any time without any consequences.

3.2 Materials

The following steps were taken to select 357 Chinese classifiers for the present study. First, we compiled an initial list of classifiers from three authoritative Chinese linguistic resources: *Xiandai Hanyu Cidian* (Chinese Academy of Social Sciences, Institute of Linguistics, 2016), Hanyu Liangci Xiangjie yu Suyuan Cidian (Chen, 2024) and Center for Chinese Linguistics (CCL) Corpus of Peking University (2003-present). Next, the list was cross-checked and refined by three students who are native speakers of Mandarin Chinese and major in linguistics.

In selecting classifiers, we adopted an inclusive criterion. Specifically, we included all items that serve a classificatory or quantificational function in the canonical "Numeral + Classifier + Noun" structure, as well as verbal classifiers used to quantify actions or events. Both currently productive classifiers and those that primarily appear in classical texts were retained. Furthermore, we did not exclude classifiers based on the flexibility of their collocations: items that combine with a wide range of nouns and those that occur in fixed expressions (e.g., 一叶扁舟 *yi2 ye4 pian1zhou1* "a small boat," a fixed idiomatic phrase) were all included in our dataset. We excluded classifiers that merely indicate units of measurement (e.g., 毫升 *hao2sheng1* 'milliliter', ml) or transliterated classifiers (e.g., 打 *da2* 'dozen').

3.3 Procedure

Online questionnaires were generated using Tencent platform (http://wj.qq.com). The 357 Chinese numeral classifiers were randomly divided into 24 lists, each containing between 14 and 15 classifiers. Each participant was assigned to one list only to minimize fatigue and maintain data quality. The overall questionnaire design followed the procedures and rating framework of Chen et al. (2019) and Zhong et al. (2022).

Each item contained 11 sensorimotor dimensions: six perceptual modalities (vision, hearing, taste, smell, touch, and interoception) and five action effectors (foot/leg, hand/arm, mouth/throat, head, and torso). Participants were asked to rate how strongly each classifier evokes perceptual or motor experiences along these 11 dimensions on a 6-point Likert scale (0 = no feelings at all, 5 = very strong feelings).

We adopted the 11-dimension sensorimotor framework (six perceptual modalities and five action effectors) for three reasons. First, this set provides a widely used, theory-motivated "core" description of embodied experience and ensures direct comparability with existing Mandarin sensorimotor norms for nouns and adjectives (Chen et al., 2019; Zhong et al., 2022). Second, given that each participant rated ~160–170 cells per questionnaire (11 dimensions × 14–15 classifiers), expanding the framework with additional fine-grained features (e.g., material, rigidity, elasticity) would substantially increase time-on-task and fatigue, potentially compromising data quality. Third, our goal in this first release is to establish a reliable, reusable backbone of perceptual-motor profiles for the classifier system. Future work can build on this backbone by adding targeted subdimensions or context-sensitive feature judgments for specific semantic contrasts (see Limitations).

To ensure data quality, the questionnaire included four quality control questions assessing participants' fundamental knowledge of Chinese. These items covered multiple linguistic domains, including orthographic, morphological, and semantic processing abilities. The specific questions were as follows: 1) Please write the Chinese characters corresponding to "xiāo shēng nì jì"; 2) What is the radical of the character "灿"? 3) Among the following four options, please select the one with the correct character form. "火伴; 队员; 姿式; 脾汽"; 4) Please choose the option whose meaning does not match that of the word "包袱 *bao1 fu*". "A bundle of clothes wrapped in cloth; A metaphor for a psychological or emotional burden; A piece of cloth inside a backpack; A humorous element or punchline in a comic dialogue or rhythmic storytelling performance." These questions examined participants' basic competence in Chinese orthography, radical identification, and semantic understanding.

Before starting the formal task, participants were presented with detailed rating instructions illustrating each perceptual and motor dimension in Chinese version: *"Visual sense allows us to see this beautiful world; auditory sense permits us hearing all the sound; gustatory sense considers tastes; olfactory sense accounts for odours; tactile sense perceives temperature, pain, and textures of objects; interoceptive sense detects the feelings of hungry, exhausted, disgusting, etc. We also use our body parts to complete these perceptions. For example, we need to move our head or torso when we view the surroundings, we will open our mouth and articulate with our throat when we talk, we use our hands (and arms) to reach and grasp for objects, and we also resort to our feet (and legs) to walk or to kick. In what follows, you will see some CLASSIFIERS in Chinese. We would like you to judge how much you experience these concepts through each of the perceptions and actions we ask, on a scale from 0 = no feelings at all, to 5 = very strong feelings."*

Participants were instructed to rate according to their own experience, proceeding at their own pace, with the option to pause at any time. Each questionnaire took approximately 10 minutes to complete, and participants received a monetary reward for their participation.

4. Data Preparation and Analysis

4.1 Data Cleaning

A total of approximately 1,190,000 potential data points were collected from 523 participants across 24 lists. Each list contained between 14 and 15 unique classifiers, resulting in an average of 160–170 rating items per participant.

Prior to statistical analysis, all questionnaire data underwent a rigorous multi-step screening process to ensure data quality and reliability. First, four instructed-response items were embedded to assess attentiveness. Participants who provided incorrect answers to any of these items were excluded. Second, two additional response-quality indices were examined within each list: (a) the missing response rate, where participants with more than 20% unanswered items were excluded and (b) straightlining, defined as selecting the same rating option for ⩾90% of items. Third, completion time was examined within each list. Participants whose total completion time fell outside ±2 *SD*s of the list mean were excluded. Applying these criteria led to the exclusion of 174 participants (37.4%) based on attention-check failure or implausible completion times (Steps 1 & 3), and a further 3 participants (0.6%) due to straight lining (Step 2). Consequently, 288 participants (61.9%) remained for subsequent analyses.

At the data-point level, these exclusions resulted in a reduction from 1,190,000 potential response entries to approximately 56,000 valid data points, corresponding to a retention rate of 4.7%. Although this percentage appears small, it reflects the large potential response matrix (i.e., all possible participant–item combinations) and the stringent participant-level exclusion criteria. Importantly, the 4.7% retention rate is computed against the full potential response matrix (all participant–item combinations), whereas each participant actually completed only one list; thus the final ~56,000 data points reflect the intended sparse design combined with conservative participant-level exclusions. Such a conservative screening procedure ensures that the final dataset is composed solely of valid, attentive, and representative responses. All data cleaning and statistical analyses were conducted using R version 4.3.1 (R Core Team, 2023).

4.2 Computation of Indices

After the ratings were collected, we examined the sensorimotor norms by assigning the following measurements to each word (Zhong et al., 2022). In addition to reporting mean ratings for each of the 11 sensorimotor dimensions (vision, hearing, taste, smell, touch, interoception; hand/arm, leg/foot, mouth/throat, head, torso; 0–5 scale), we derived several summary indices. The Perceptual Mean is the average of the six sensory ratings, and the Action Mean is the average of the five effector ratings. The Dominant Modality and Dominant Effector are the highest values within the sensory and motor sets, respectively; the Dominant Sensorimotor is the highest value across all 11 dimensions. To quantify concentration, we computed exclusivity as the range divided by the sum of ratings within a set (Connell & Lynott, 2009): Modality Exclusivity for the six senses, Effector Exclusivity for the five effectors, and Sensorimotor Exclusivity across all 11 dimensions (higher values indicate a profile concentrated on a single channel). The processed classifiers and their corresponding values for each sensorimotor dimension can be accessed via the following anonymous link: https://osf.io/zgj57/overview?view_only=9e3d9f2ff4144098a18e3b0ea869bd03.

In the second step, the intercorrelations among the sensorimotor dimensions were examined to explore how perceptual and motor experiences are interrelated across classifiers. This analysis helped uncover the underlying structure of sensorimotor organization in the classifier system.

Finally, to illustrate the validity and interpretability of the norms, we selected several pairs of near-synonymous classifiers and visualized their ratings using radar charts. These plots highlighted both shared and distinctive sensorimotor features, demonstrating that the derived norms capture meaningful experiential distinctions among classifiers.

4.3 Reliability of the Measures

To assess inter-rater reliability, we conducted a split-half reliability analysis following Yao et al. (2017). For each dimension, raters were randomly divided into two equal subgroups, and the mean rating of each classifier in each subgroup was computed; correlations were then Spearman–Brown corrected. The corrected split-half coefficients indicated overall acceptable reliability, ranging from 0.575 to 0.834 across the 11 dimensions. Reliability was highest for hand/arm (shoulder) action ratings (r_SB = .834), followed by tactile (r_SB = .799) and olfactory (r_SB = .780) ratings. The lowest reliability was observed for interoception/internal sensation (r_SB = .575), suggesting greater variability in how participants interpreted or applied this dimension compared with externally anchored perceptual and action dimensions. Overall, the reliability pattern indicates that ratings tied to more concrete sensorimotor experiences tend to be more consistent across raters.

4.4 Rating Consistency

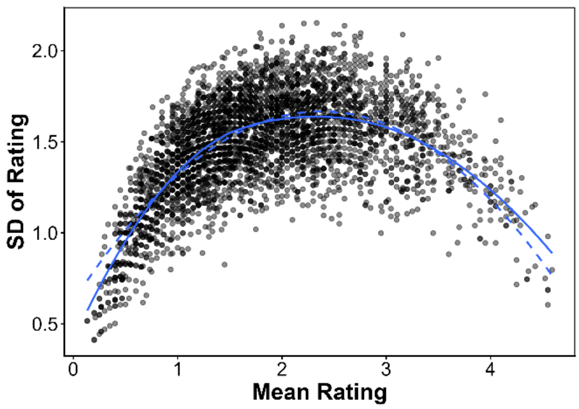

Figure 1: Relationship between mean rating and rating variability (*SD*) across classifiers and dimensions.

To assess the consistency of participants' ratings, we examined the relationship between the mean and standard deviation (*SD*) of ratings for all classifiers across dimensions (as shown in Figure 1). A moderate positive correlation was observed between the mean and SD of ratings (r = .37), indicating that classifiers receiving higher mean ratings tended to exhibit greater variability across participants.

To further characterize this pattern, a quadratic regression was conducted with *SD* as the dependent variable and the mean rating as the predictor ($SD = \beta_0 + \beta_1 \times \text{Mean} + \beta_2 \times \text{Mean}^2$). The model accounted for approximately 50% of the variance in *SD* ($R^2 = .4999$, $F(2, 3848) = 1923$, $p < .001$). Both the linear ($\beta_1 = 0.876$, $p < .001$) and quadratic ($\beta_2 = -0.184$, $p < .001$) terms were significant, revealing a clear inverted-U relationship: variability was highest for items with intermediate mean ratings and lower for items rated as either very weak or very strong on a given dimension. This pattern suggests that rating variability tends to decrease as consensus becomes stronger at the extremes of the scale.

5. Results

5.1 Descriptive Statistics

5.1.1 Overview of the sensorimotor strength

Across all 357 classifiers, the overall mean perceptual and action strength was 1.82 (*SD* = 0.45) on a six-point scale, indicating that Chinese classifiers generally evoke a moderate level of perceptual and motor experience. Figure 2 (see Appendix) visualizes the overall distribution of the sensorimotor dimensions across classifiers. It demonstrates broad variability within and across dimensions, with visual and hand/arm ratings showing the highest median values and widest ranges. This pattern again underscores the central role of perception through vision and action through manipulation in classifier meaning.

Among the six perceptual modalities, the perceptual mean for the six senory modalities is 1.87 (*SD*=0.45). The visual modality showed the highest mean strength (M = 2.90, *SD* = 0.69), followed by touch (M = 2.25, *SD* = 0.77) and interoception (M = 2.07, *SD* = 0.49). In contrast, auditory (M = 1.83, *SD* = 0.67), olfactory (M = 1.09, *SD* = 0.63), and gustatory (M = 1.01, *SD* = 0.65) modalities received considerably lower ratings. This pattern indicates that visual and tactile information are the primary sensory bases of classifier meaning, whereas smell and taste contribute minimally.

Furthermore, the interoception dimension exhibited a mean comparable to the tactile and visual domains, implying that many classifiers encode not only external perceptual properties but also internal bodily or affective states. The overall sensorimotor strength across all 11 dimensions (including motor-related ones) was 1.82 (*SD* = 0.45), slightly lower than the sensory-only average, reinforcing the relatively stronger perceptual grounding of classifiers compared to their motor associations.

In addition to sensory ratings, action mean for the five action effectors is 1.75 (*SD*=0.54). The hand/arm effector obtained the highest strength (M = 2.20, *SD* = 0.81), reflecting the strong manual interaction and manipulability associated with classifier semantics. This was followed by torso (M = 1.81, *SD* = 0.68), foot/leg (M = 1.66, *SD* = 0.73), head (M = 1.62, *SD* = 0.57), and mouth/throat (M = 1.46, *SD* = 0.62). These results suggest that manual and whole-body actions are the most prominent motor dimensions encoded in the classifier system.

Overall, these descriptive patterns reveal that Chinese classifiers are strongly grounded in visual and manual experiences, reflecting their role in categorizing entities according to shape, manipulability, and bodily interaction. This supports the view that the classifier system embodies perceptual and motor structures of human experience, rather than serving as a purely syntactic device.

5.1.2 Dominant Sensorimotor Dimensions and Exclusivity

Figure 3 (see Appendix) presents the distribution of dominant sensorimotor across the 357 classifiers. The majority of classifiers were visually dominant (N = 215, 60.2%), indicating that visual information plays a central role in classifier semantics. This is followed by classifiers dominated by hand/arm actions (N = 41, 11.5%) and interoception (N = 27, 7.6%), the latter typically linked to abstract or introspective categories (e.g., 种 *zhong3* "type", 类 *lei4* "category"). Auditory (N = 18, 5.0%) and tactile (N = 16, 4.5%) dominance occurred less frequently, while gustatory, olfactory, mouth/throat, head, and trunk dominances were rare (<2%). These distributions again suggest that visual and manual experiences constitute the most salient embodied dimensions encoded by Chinese classifiers, consistent with their function in distinguishing entities based on shape and manipulability (Ahrens & Huang, 2016).

Within the sensory modalities, a strong bias toward the visual modality was observed, which accounted for 247 classifiers (69.2%), indicating that the majority of Chinese classifiers primarily evoke visual experience. The tactile modality ranked second, covering 36 classifiers (10.1%), followed by interoception with 31 classifiers (8.7%), suggesting that certain classifiers also engage somatosensory or interoceptive experiences. By contrast, auditory (22 classifiers, 6.2%), gustatory (4 classifiers, 1.1%), and olfactory (4 classifiers, 1.1%) modalities were much less represented, reflecting the relatively peripheral role of these sensory channels in the classifier system.

The hand/shoulder effector overwhelmingly dominated the distribution, accounting for 213 classifiers (59.7%), indicating that most classifiers are associated with manual or upper-limb actions such as grasping, holding, or manipulating objects. The leg/foot effector ranked second (N = 40, 11.2%), followed by mouth/throat (N = 36, 10.1%), suggesting that a smaller subset of classifiers is related to actions involving locomotion or oral articulation. Additionally, trunk (excluding limbs) and head (excluding mouth) accounted for 35 classifiers (9.8%) and 26 classifiers (7.3%), respectively, indicating limited but discernible associations with bodily movements centered in the torso or head.

The sensorimotor exclusivity values, which quantify the extent to which a classifier is associated with a single sensorimotor modality, were generally low across the dataset. As shown in Figure 4 (see Appendix), the mean exclusivity was 11.61% (*SD* = 4.19%), with values ranging from 0.00% to 24.39%. The distribution was slightly right-skewed, indicating that most classifiers are moderately multimodal rather than being dominated by a single dimension. The interquartile range (IQR = 8.93 - 14.48) further suggests that the majority of classifiers show low to medium degrees of modality concentration. This pattern implies that Chinese classifiers typically integrate information from multiple perceptual and action domains, consistent with their semantic functions of categorizing nouns according to various perceptual or functional attributes rather than a single sensory source.

The modality exclusivity had a mean of 19.72% (*SD* = 7.30%), indicating that most classifiers were associated with moderately distributed rather than highly modality-specific sensory profiles. This moderate exclusivity suggests that classifier concepts tend to evoke multiple perceptual experiences simultaneously instead of being tied to a single sensory channel.

Additionally, the effector exclusivity averaged 13.62% (*SD* = 7.03%), notably lower than that of the sensory dimensions. This lower exclusivity indicates that motor-related associations were more diffusely distributed across different body effectors, reflecting overlapping sensorimotor imagery rather than sharply localized action representations.

Taken together, the results suggest that Chinese classifiers are more perceptually grounded than motorically grounded, with visual and tactile imagery contributing more strongly to their conceptual representation than bodily movement.

5.2 Interrelations among Sensorimotor dimensions

Figure 5 presents the intercorrelations among the

11 sensorimotor dimensions. Overall, the pattern of correlations reveals a coherent organization of sensory and motor experiences, in line with findings reported for other word classes (Lynott et al., 2020; Zhong et al., 2022).

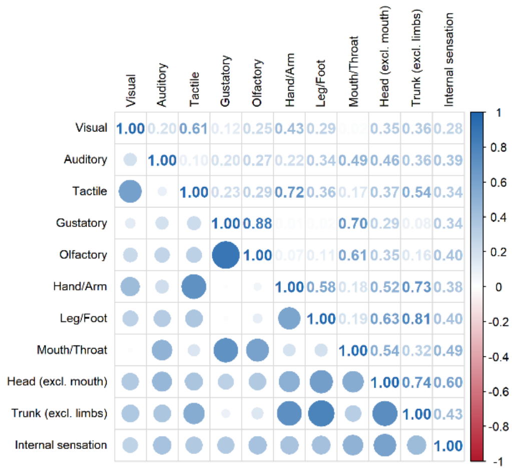

Figure 5: Intercorrelations among the 11 sensorimotor dimensions.

Among the perceptual modalities, the strongest positive relationship was observed between gustatory and olfactory dimensions (r = 0.88), confirming their close experiential and neural linkage. Visual and tactile strengths were also moderately correlated (r = 0.61), reflecting the frequent co-occurrence of visual and manual information in representing object. By contrast, auditory strength showed only weak correlations with other sensory modalities (all rs < 0.30), suggesting that hearing plays a relatively independent role in classifier semantics.

Regarding action effectors, the hand/arm dimension exhibited positive correlations with both visual (r = 0.43) and tactile (r = 0.72) modalities, consistent with the dominance of visually guided manual interaction in classifier use (e.g., 把 *ba3*, 支 *zhi1*, 根 *gen1*). The leg/foot effector was modestly related to vision and touch (rs = 0.29 - 0.36), while mouth/throat correlated more strongly with auditory (r = 0.49) and gustatory (r = 0.70) modalities, reflecting their shared articulatory or ingestion-related experiences.

Finally, the interoceptive dimension showed small-to-moderate correlations with multiple domains (mean r = 0.35), suggesting that abstract or internal concepts associated with classifiers often integrate low-level sensorimotor information with introspective experience.

Taken together, these results reveal a structured but differentiated sensorimotor network underlying Chinese classifier meanings: strong convergence between gustatory and olfactory modalities, moderate coupling among vision, touch, and manual action, and relative independence of auditory and interoceptive dimensions.

5.3 Validation of the Norms through Near-Synonymous Classifier Comparison

To assess the validity of the established sensorimotor norms, we compared the multidimensional profiles of near-synonymous classifiers that have been discussed as semantically related in previous linguistic research (e.g., Ahrens & Huang, 2016). These sets of classifiers share similar referential domains but differ subtly in semantic nuance or usage contexts. For instance, 条 *tiao2* and 根 *gen1* all apply to individual objects with the prototypical attributes of long, thin, and cylindrical, yet 条 *tiao2* typically denotes flexible entities (e.g., 一条蛇 *yi4 tiao2 she2* "a snake"), 根 *gen1* emphasizes rigidity (e.g., 一根钢管 *yi4 gen1 gang1guan3* "a steel tube"). Similarly, 片 *pian4* and 块 *kuai4* are used to classify concrete nouns in typical piece form. 片 refers to thin slices or fragments (e.g., 一片树叶 *yi2 pian4 shu4ye4* "a leaf"), whereas 块 *kuai4* emphasizes a more three-dimensional-like state (e.g., 一块乳酪 *yi2 kuai4 ru3lao4* "a piece of cheese"). Comparing their radar-shaped sensorimotor profiles revealed that these classifiers share overlapping sensorimotor patterns consistent with their semantic domains, while also displaying fine-grained divergences that reflect their distinct conceptual emphases and usage conventions.

5.3.1 Near-synonymous set 1: 条 and 根

Both classifiers display a visual-tactile sensorimotor profile as shown in Figure 6, with Visual as the dominant dimension and Tactile as the secondary peak. Visual salience is strongest for 根 *gen1*, followed by 条 *tiao2*, consistent with their shared association with long and slender objects that are easily visualized in terms of shape. Tactile strength is likewise higher for 根 *gen1*, reflecting its emphasis on rigidity and solid materiality (e.g., 一根钢管 *yi4 gen1 gang1guan3* "a steel tube"), whereas 条 typically denoting flexible and elongated entities (e.g., 一条蛇 *yi4 tiao2 she2* "a snake") elicits relatively weaker tactile associations.

Interestingly, the two classifiers also diverged in their motion effector profiles. 根 *gen1* showed the highest mean value on the Leg/Foot dimension (M = 1.70), whereas 条 *tiao2* reached its peak on the Mouth/Throat dimension (M = 1.95). The pattern for 根 *gen1* is consistent with previous descriptions that trace its prototype to the feature [+component: having roots] (Wang, 2022). In this sense, the leg/foot effector metaphorically represents the notion of "rootedness" or extension from a base. In contrast, the elevated mouth/throat association for 条 *tiao2* lacks a clear theoretical account to date, as no existing framework has linked this classifier to articulatory or oral actions in human experience.

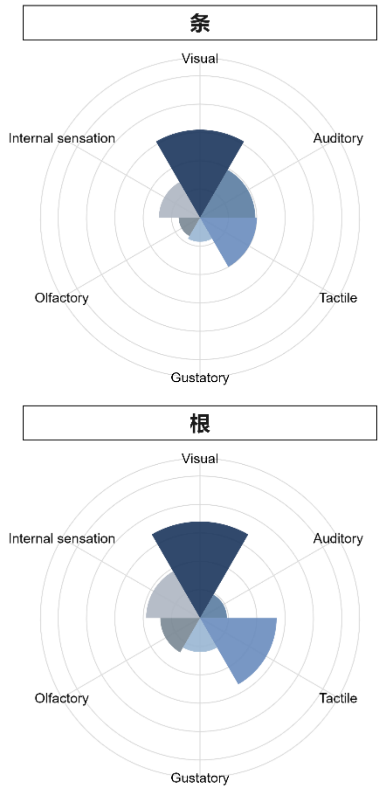

Figure 6. Sensory profiles of the classifiers 条 and 根. Bar length represents the average rating on a 0 – 5 scale.

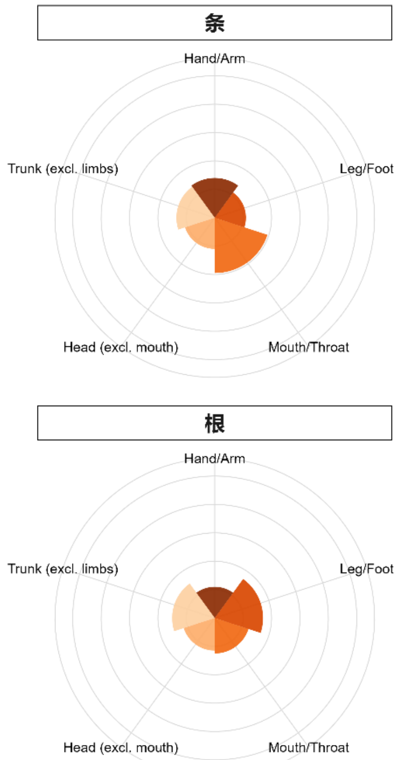

Figure 7. Motor profiles of the classifiers 条 and 根. Bar length represents the average rating on a 0 – 5 scale.

5.3.2 Near-synonymous set 2: 片 and 块

片 and 块 share a visual–tactile core, consistent with their flat, sheet-like construal. Within this common perceptual profile, however, a subtle gradation emerges. From the perceptual dimensions, the two classifiers show no marked differences. Both are strongly visual and moderately tactile, reflecting their shared association with planar or surface-like objects.

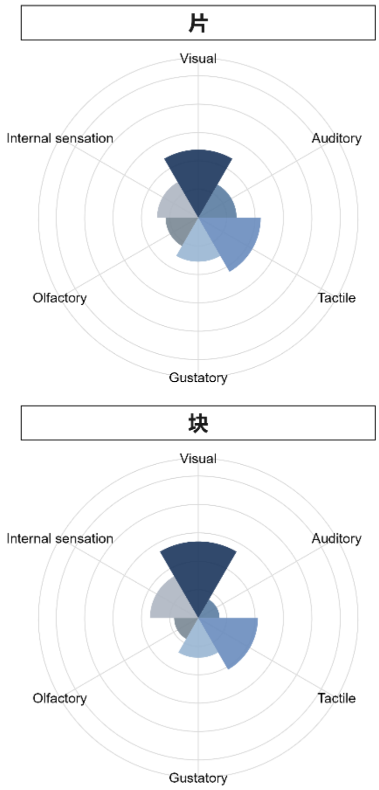

Figure 8. Sensory profiles of the classifiers 片 and 块. Bar length represents the average rating on a 0–5 scale.

In contrast, their motor effector profiles reveal a systematic divergence. 块 *kuai4* exhibits higher mean ratings than 片 *pian4* across all effector dimensions, particularly on Hand/Shoulder, suggesting that 块 *kuai4* evokes a stronger sense of body-object interaction and physical manipulation. This pattern accords with the semantic distinction that 块 *kuai4* typically refers to three-dimensional (Meng and Li, 2014; Zhou & Shao, 2014), solid masses, whereas 片 *pian4* denotes two-dimensional (Meng & Li, 2011), thin surfaces which is a contrast that maps neatly onto greater embodied engagement and volumetric representation for 块 *kuai4*.

6. Discussion

The present study provided a large-scale quantitative characterization of Chinese classifiers, based on perceptual and motor ratings across 11 embodied dimensions. Specifically, we systematically examined how classifiers encode sensory and action-based experience, and how these embodied dimensions structure the classifier system. The results revealed robust visual–manual dominance, meaningful tactile and interoceptive contributions, and distinctive profiles among near-synonymous classifiers, together supporting an embodied view of classifier semantics.

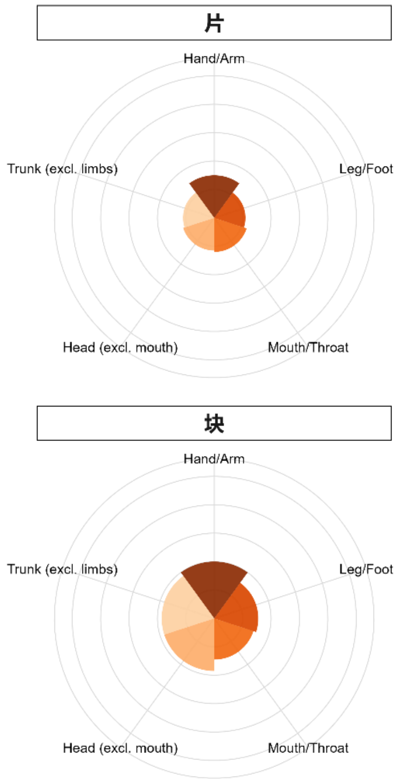

Figure 9. Motor profiles of the classifiers 片 and 块. Bar length represents the average rating on a 0–5 scale.

6.1 Embodied Semantics of Chinese Classifiers

Within the framework of embodied cognition (Barsalou, 1999, 2008), conceptual knowledge is grounded in the sensorimotor systems that support real-world perception and action. Words referring to manipulable objects or actions elicit simulation of bodily interactions, engaging perceptual and motor representations that mirror actual experience. The strong visual and hand/arm dominance observed across classifiers is consistent with this view: classifiers frequently encode how entities are seen, grasped, or manipulated, reflecting their function as linguistic markers of shape, size, and affordance.

From this perspective, the classifier system can be understood as a lexical interface between language and embodied experience, where semantic distinctions are structured by differences in potential bodily interaction. Classifiers that encode greater affordance for manual contact (e.g., 块 *kuai4*, 根 *gen1*) exhibit higher ratings on motor effectors, particularly the hand/shoulder dimension, whereas classifiers referring to less manipulable or abstract referents rely more heavily on visual or interoceptive grounding. This alignment between linguistic function and bodily engagement underscores that classifier semantics are not arbitrary conventions but emergent reflections of perceptual–motor organization in human cognition.

This embodied mapping also resonates with findings from other lexical domains. For instance, sensorimotor norming studies on Chinese nouns, adjectives, and verbs (Britton, 2024; Chen et al., 2019; Zhong et al., 2022) similarly identify vision as the dominant modality across linguistic categories. However, compared with nouns and adjectives, the classifier system exhibits a more pronounced contribution from motor dimensions, particularly those associated with manual and arm actions. This distinctive sensorimotor profile highlights the inherently embodied nature of Chinese classifiers, as they reflect not only how humans perceive the world but also how they physically interact with it. In this sense, the systematic distribution of classifier semantics reveals a deep structural linkage between language and bodily experience, providing novel evidence for the embodied foundations of the Chinese semantic system.

6.2 Systemic Patterns of Sensorimotor Distribution

At the systemic level, two major trends emerge: (1) the predominance of visual and manual dimensions, and (2) the relatively weak involvement of auditory, olfactory, and gustatory modalities. The visual dominance (≈ 60–70% of classifiers) reflects the centrality of shape and spatial configuration in classifier semantics. Since classifiers categorize nouns primarily by perceptual features such as form, size, and arrangement, visual experience naturally constitutes their strongest grounding domain. The prominence of hand/arm effector (≈ 60%) further emphasizes the role of manipulability in classifier meaning, consistent with the human bias toward object interaction and tool use. Classifiers thus encode not only how entities are perceived but also how they are acted upon—linking perceptual recognition with motor routines.

Conversely, auditory, gustatory, and olfactory modalities were infrequent and weakly correlated with other dimensions. This asymmetry can be attributed to both linguistic and ecological factors. Classifiers rarely categorize entities by sound, taste, or smell, as these properties are less reliable for individuating referents in visual contexts. The low exclusivity scores across the dataset reinforce this multimodal yet perception-heavy architecture: most classifiers integrate overlapping visual, tactile, and motor features rather than being dominated by a single sensory channel. In this sense, the classifier system mirrors the perceptual–functional organization of human experience, privileging domains that are most relevant for identifying and manipulating objects.

A notable pattern concerns interoception. Interoceptive ratings were non-trivial on average and constituted the dominant modality for a subset of classifiers. One interpretation is that interoception provides an embodied grounding route for more abstract classifier meanings (Zhong et al., 2022): items used for categorization or taxonomy (e.g., 种 *zhong3* “type”, 类 *lei4* “category”) may evoke internal-state metaphors such as a sense of “mental sorting” or affective evaluation rather than concrete external properties. At the same time, interoception also showed the lowest split-half reliability among the dimensions, suggesting greater variability in how participants mapped this label to their experience. Accordingly, we treat interoceptive effects cautiously and view them as a promising but less stable channel that would benefit from additional clarification in instructions and independent validation (e.g., targeted follow-up ratings with richer examples, or behavioral/neurocognitive tasks testing whether high-interoception classifiers yield measurable processing differences).

7. Conclusion

This study provides the first comprehensive sensorimotor norms for Chinese numeral classifiers, demonstrating that classifier meanings are systematically grounded in perception and action. Classifiers predominantly evoke visual and manual experiences, reflecting human categorization based on how entities are seen and handled. The results reveal a multimodal yet perception-biased structure: most classifiers integrate visual, tactile, and motor features, while auditory, olfactory, and gustatory channels play minor roles. The two-cluster pattern further distinguishes perception-based from action-enriched classifiers, underscoring the embodied continuum that links sensory imagery and motor affordance in grammatical semantics. Taken together, these findings show that even grammatical markers such as classifiers encode embodied schemata derived from perceptual and motor experience. They highlight the classifier system as a grammatical reflection of how humans perceive and interact with the world. Beyond descriptive characterization, these norms provide continuous, reusable predictors for future behavioral, neurocognitive, and computational studies, enabling direct tests of whether classifier-specific sensorimotor strengths systematically modulate online comprehension and production.

8. Limitations

First, the present study is constrained by the dimensional framework used for rating. Although the 11 perceptual and motor dimensions (e.g., vision, touch, hand/arm) capture major channels of embodied experience and enable comparability with existing Mandarin sensorimotor norms, they may not fully reflect fine-grained semantic properties that are often crucial for classifier choice (e.g., material, rigidity, elasticity, curvature, or specific affordance types). We adopted the core 11-dimension framework as a principled compromise between coverage and participant burden given the dense rating structure within each questionnaire. Future work could extend the current backbone by adding targeted subdimensions or feature-based judgments for subsets of classifiers where finer distinctions are theoretically important.

Second, the norms are based on subjective introspective ratings. While such norms are standard in embodied-semantics research and we implemented stringent quality screening and reliability checks, ratings may still be influenced by individual differences in imagery ability, linguistic exposure, and task interpretation—particularly for dimensions that are less externally anchored, such as interoception. Complementary behavioral and neurocognitive validation (e.g., priming, lexical decision, sentence-level comprehension, EEG/MEG) would help establish how these normative strengths map onto online processing.

Third, our inventory intentionally includes both highly productive classifiers and items that are low-frequency or primarily attested in classical or fixed expressions. Because we do not currently provide corpus-based frequency information for each classifier, it is difficult to quantify how usage frequency might shape the stability of ratings or the generalizability of particular patterns. This limits the interpretability of comparisons between widely used and rarer items, and constrains the ability to control for frequency in downstream analyses. A valuable next step would be to add standardized frequency metadata from a clearly specified corpus (and, where appropriate, to annotate items by contemporary vs. historical usage), allowing users to filter items or include frequency as a covariate in modelling.

9. Ethics Statement

This research was reviewed and approved by the ethics committee of the authors' affiliated institution. All procedures involving human participants were conducted in accordance with recognized ethical research standards. Before participating, individuals were fully informed about the purpose and procedures of the study and provided informed consent via an online form. Participation was entirely voluntary, and participants could withdraw at any time without penalty. All responses were collected anonymously, and no personally identifiable information was recorded or disclosed. The study data was used exclusively for academic research purposes.

10. Bibliographical References

Ahrens, K., & Huang, C.-R. (2016). Classifiers. In C.-R. Huang & D. L. Shi (Eds.), *A reference grammar of Chinese* (pp. 169–198). Cambridge University Press.

Barsalou, L. W. (1999). Perceptions of perceptual symbols. *Behavioral and Brain Sciences*, *22*(4), 637-660. https://doi.org/10.1017/s0140525x99532147

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology*, *59*, 617-645. https://doi.org/10.1146/annurev.psych.59.103006.093639

Chan, S. H. (2019). An elephant needs a head but a horse does not: An ERP study of classifier-noun agreement in Mandarin. *Journal of Neurolinguistics*, *52*, 100852. https://doi.org/10.1016/j.jneuroling.2019.100852

Chen, I. H., Zhao, Q. Q., Long, Y. F., Lu, Q., & Huang, C. R. (2019). Mandarin Chinese modality exclusivity norms. *Plos One*, *14*(2), Article e0211336. https://doi.org/10.1371/journal.pone.0211336

Cheng, L. L. S., & Sybesma, R. (1999). Bare and not-so-bare nouns and the structure of NP. *Linguistic Inquiry*, *30*(4), 509-542. https://doi.org/10.1162/002438999554192

Frankowsky, M., Ke, D., Zwitserlood, P., Michel, R., & Bölte, J. (2022). The interplay between classifier choice and animacy in Mandarin-Chinese noun phrases: an ERP study. *Language Cognition and Neuroscience*, *37*(7), 866-882. https://doi.org/10.1080/23273798.2022.2026420

Huang, C. R., & Ahrens, K. (2003). Individuals, kinds and events: classifier coercion of nouns. *Language Sciences*, *25*(4), 353-373. https://doi.org/10.1016/s0388-0001(02)00021-9

Britton, J. R. (2024). *Embodiment and modality exclusivity of Chinese verbs* [Doctoral dissertation, The Hong Kong Polytechnic University]. PolyU Institutional Research Archive. https://theses.lib.polyu.edu.hk

Lynott, D., Connell, L., Brysbaert, M., Brand, J., & Carney, J. (2020). The Lancaster Sensorimotor Norms: multidimensional measures of perceptual and action strength for 40,000 English words. *Behavior Research Methods*, *52*(3), 1271-1291. https://doi.org/10.3758/s13428-019-01316-z

Meng, F., & Li, R. (2011). 量词“片”的语法化 [The grammaticalization of the classifier “pian”]. *语言研究 [Yuyan Yanjiu / Language Research]*, **31**(3), 66–70.

Meng, F., & Li, Y. (2014). 量词“块”的产生及其发展演变 [The emergence and development of the classifier “kuai”]. *宁夏大学学报（人文社会*

科学版）[Journal of Ningxia University (Humanities & Social Sciences Edition)], (3), 37–43.

Myachykov, A., Scheepers, C., Fischer, M. H., & Kessler, K. (2014). TEST: A tropic, embodied, and situated theory of cognition. *Topics in Cognitive Science*, 6(3), 442–460. https://doi.org/10.1111/tops.12024

R Core Team. (2023). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Speed, L. J., & Majid, A. (2020). Grounding language in the neglected senses of touch, taste, and smell. *Cognitive Neuropsychology*, *37*(5-6), 363-392. https://doi.org/10.1080/02643294.2019.1623188

Tai, J. H.-Y. (1994). Chinese classifier systems and human categorization. In M. Chen, & O. Tseng (Eds.). *In honor of professor william S-Y. Wang: Interdisciplinary studies on language and language change* (pp. 479–494). Taipei: Pyramid Publishing Company.

Truman, A., & Kutas, M. (2024). Flexible Conceptual Representations. *Cognitive Science*, 48(6), e13475. https://doi.org/10.1111/cogs.13475

Wang, J., Witteman, J., & Schiller, N. O. (2025). Processing of visual shape information in Chinese classifier-noun phrases. *Cognitive Neuropsychology*. https://doi.org/10.1080/02643294.2025.2485974

Wang, X. (2022). 量词"根"的词典释义探讨 [On the lexicographic definition of the classifier "gen"]. *辞书研究 [Cishu Yanjiu / Lexicographical Studies]*, (2), 48–54.

Wang, Y. F., Witteman, J., & Schiller, N. O. (2025a). Activation of classifiers in word production: insights from lexico-syntactic probability distributions. *Language Cognition and Neuroscience*, *40*(7), 987-1001. https://doi.org/10.1080/23273798.2025.2498017

Wang, Y. F., Witteman, J., & Schiller, N. O. (2025b). The role of animacy in language production: evidence from bare noun naming. *Quarterly Journal of Experimental Psychology*, *78*(7), 1461-1473. https://doi.org/10.1177/17470218241281868

Wu, Y. C., & Bodomo, A. (2009). Classifiers ≠ Determiners. *Linguistic Inquiry*, *40*(3), 487-503. https://doi.org/10.1162/ling.2009.40.3.487

Yao, Z., Wu, J., Zhang, Y. Y., & Wang, Z. H. (2017). Norms of valence, arousal, concreteness, familiarity, imageability, and context availability for 1,100 Chinese words. *Behavior Research Methods*, *49*(4), 1374-1385. https://doi.org/10.3758/s13428-016-0793-2

Zhong, Y., Wan, M. Y., Ahrens, K., & Huang, C. R. (2022). Sensorimotor norms for Chinese nouns and their relationship with orthographic and semantic variables. *Language Cognition and Neuroscience*, *37*(8), 1000-1022. https://doi.org/10.1080/23273798.2022.2035416

Zhou, S., & Shao, J. (2014). 近义量词"条、根、道"的三维解释与组合机制 [A three-dimensional account and combinational mechanism of the near-synonymous classifiers "tiao," "gen," and "dao"]. *语言教学与研究 [Yuyan Jiaoxue yu Yanjiu / Language Teaching and Linguistic Studies]*, (1), 81–88.

Zwaan, R. A., & Madden, C. J. (2005). Embodied Sentence Comprehension. In D. Pecher & R. A. Zwaan (Eds.), *Grounding Cognition: The Role of Perception and Action in Memory, Language, and Thinking* (pp. 224-245). Cambridge University Press. https://doi.org/10.1017/CBO9780511499968.010

11. Appendix

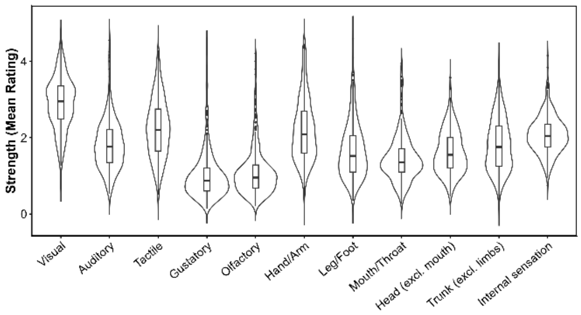

Figure 2: Distribution of sensorimotor strength across the 11 dimensions. Each violin plot represents the distribution of mean strength ratings for one sensorimotor dimension across the 357 classifiers. Boxplots embedded within the violins indicate the median and interquartile range.

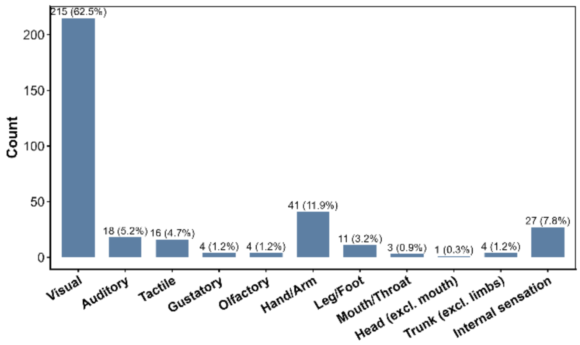

Figure 3: Distribution of dominant sensorimotor modalities across classifiers.

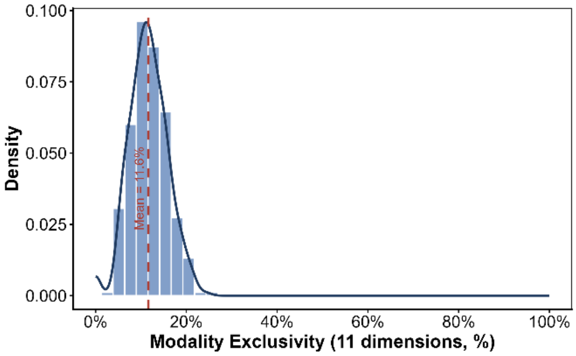

Figure 4: Distribution of modality exclusivity across classifiers (11 dimensions). The solid curve represents the kernel density estimate, and the dashed vertical line indicates the mean exclusivity.

12. Language Resource References

Xiandai Hanyu Cidian / Modern Chinese Dictionary (7th ed.) Dictionary Editing Office, Institute of Linguistics, Chinese Academy of Social Sciences. 2016. Xiandai Hanyu Cidian (7th ed.). The Commercial Press. ISBN 9787100124508.

Chen, F. (Ed.). 2024. Hanyu Liangci Xiangjie yu Suyuan Cidian 汉语量词详解与溯源词典 [Chinese dictionary of quantifiers/classifiers]. Yuwen Chubanshe.

CCL Corpus (Center for Chinese Linguistics, Peking University) Center for Chinese Linguistics, Peking University. 2003–present. CCL Corpus (Modern & Classical Chinese). Hosted at Center for Chinese Linguistics (Peking University). Available at: http://ccl.pku.edu.cn/