Linguistic Synaesthesia of Mandarin Sensory Adjectives: Corpus-Based and Experimental Approaches

Qingqing Zhao[1], Yunfei Long[2], Chu-Ren Huang[3]

[1] Institute of Linguistics, Chinese Academy of Social Sciences
[2] School of Medicine, University of Nottingham
[3] Department of Chinese and Bilingual Studies, The Hong Kong Polytechnic University
zhaoqingqing0611@163.com, Yunfei.Long@nottingham.ac.uk,
churen.huang@polyu.edu.hk

Abstract. This study examines linguistic synaesthesia based on both the corpus distribution and the modality rating of Mandarin synaesthetic adjectives. We find that the tendencies attested through the corpus-based and the experimental approaches are compatible, including: (1) the modality exclusivity is negatively correlated with the usage of Mandarin sensory adjectives in linguistic synaesthesia; and (2) the ratings on sensory modalities of Mandarin synaesthetic adjectives are consistent with the synaesthetic directionality of these adjectives. The paper thus argues for the cognitive reality of linguistic synaesthesia, which can be evidenced by both the language production in the corpus and the language processing in the behavior experiment.

Keywords: Linguistic synaesthesia, corpus-based, modality exclusivity, Mandarin

1 Introduction

Linguistic synaesthesia is a cross-linguistic language usage of lexical items in one sensory domain to conceptualize or describe objects, events, or properties in another sensory domain ([1-5]). Take linguistic synaesthesia in Mandarin Chinese for example. The lexical item 味 *wei4* "taste" originally representing a gustatory perception can be used to denote the odor in the olfactory domain, as in the word 氣味 *qi4-wei4* "odor"; the word 聞 *wen2* "to hear" etymologically conceptualizing the auditory action (as indicated by the radical 耳 *er3* "ear") can be employed to denote smelling, as in the Mandarin expression 聞一聞這個花香不香 *wen2-yi1-wen2 zhe4-ge4 hua1 xiang1-bu4-xiang1* "to smell and judge whether the flower is fragrant or not"; and the Mandarin tactile adjective 冷 *leng3* "cold" can be utilized to describe the visual perception, as in the sentence 色調很冷 *se4-diao4 hen3 leng3* "The color is cold." ([6])

There are basically two approaches adopted to examine linguistic synaesthesia in the literature. One is the corpus-based approach, through which transfer directionalities of sensory items among different sensory modalities in linguistic

synaesthesia are generalized on the basis of the distributions of the items in the corpus. Studies employing the corpus-based approach are such as [4-5] and [7-9]. The other approach is experimental. Research based on experiments (mostly behavior experiments) mainly focuses on the directionality constraints of linguistic synaesthesia in the language processing. More specifically, synaesthetic expressions following transfer directionalities are found to be more easily processed than the expressions violating the directionalities ([3], [10-14]). However, few studies have combined the two different methods to further investigate linguistic synaesthesia. This paper aims to show that the tendencies found through the two approaches are in fact compatible based on the corpus and experimental data of Mandarin synaesthetic adjectives.

2 Methods

[8] utilized a well-built linguistic synaesthesia identification procedure to extract Mandarin sensory adjectives with synaesthetic usages in the Sinica corpus ([15]). In [8], she identified 260 Mandarin synaesthetic adjectives, of which 199 adjectives are composed of morphemes with the etymological origin in the same sensory domain (e.g., 甜 *tian2* "sweet" and 甜美 *tian2-mei3* "tasty") and 61 adjectives are compounded by morphemes originally from different sensory modalities (e.g., 苦澀 *ku3-se4* "bitter-rough [bitter]"). This study employs the corpus data of linguistic synaesthesia from the distributions of Mandarin synaesthetic adjectives in the Sinica corpus collected by [8].

[16] followed [17-18] to ask native speakers to rate the extent (from 0 to 5) to which Mandarin sensory adjectives can be perceived via each of the five sensory modalities (including touch, taste, smell, vision, and hearing). In [16], the modality ratings of 171 Mandarin adjectives consisting of one single sensory concept (e.g., 冷 *leng3* "cold" and 朗朗 *lang3-lang3* "bright") and 61 Mandarin compound adjectives comprising morphemes from different sensory domains (e.g., 濃厚 *nong2-hou4* "of intense taste-thick [dense]") were collected. The experimental data of linguistic synaesthesia this study relies on is from [16]'s modality ratings on Mandarin sensory adjectives.

Based on the two data sources, we extract Mandarin adjectives occurring in both [8] and [16], thus obtaining 72 Mandarin adjectives involving one sensory concept and 61 Mandarin adjectives compounded by different sensory concepts. As the compound adjectives would show different and more complex patterns in linguistic synaesthesia (see [8-9]), we only focus on the 72 adjectives containing one single sensory concept in this paper and leave the compound adjectives for future research.

As shown in Table 1, tactile and visual adjectives are the top two among the extracted 72 Mandarin adjectives, with 32 adjectives (e.g., 脆 *cui4* "crisp" and 暖 *nuan3* "warm") and 26 adjectives (e.g., 暗 *an4* "dark" and 薄 *bao2* "thin") respectively. In contrast, adjectives originally from auditory and olfactory domains are the two least, with only one (i.e., 吵 *chao3* "loud") and two adjective types (i.e., 臭 *chou4* "smelly" and 香 *xiang1* "fragrant") respectively. The gustatory adjectives

are in the middle on the rank of the numbers of adjective types among the extracted adjectives, with 11 adjectives such as 淡 *dan4* "of mild taste" and 酸 *suan1* "sour". The number rank of adjective types is generally in line with the one of the whole Mandarin synaesthetic adjectives attested by [8]. The two differences between this study and [8] are both on the relative order: tactile adjectives are the most and auditory adjectives the least in this study, while visual adjectives are the most and olfactory adjectives the least in [8]. Therefore, the tendency of the extracted 72 Mandarin synaesthetic adjectives employed by this study would be representative of the whole synaesthetic adjectives of Mandarin Chinese.

Table 1. Mandarin adjectives from five sensory modalities

Sensory domain	Number of adjectives	Example
TOUCH	32	脆 *cui4* "crisp"
VISION	26	暗 *an4* "dark"
TASTE	11	酸 *suan1* "sour"
SMELL	2	臭 *chou4* "smelly"
HEARING	1	吵 *chao3* "loud"

In the following, we will demonstrate that: (1) the modality exclusivity scores of Mandarin adjectives calculated based on the sensory ratings show a significantly negative correlation with the usages of these adjectives in linguistic synaesthesia; and (2) the sensory ratings of Mandarin synaesthetic adjectives are consistent with the transfer directionalities of these adjectives in linguistic synaesthesia.

3 Modality Exclusivity of Mandarin Synaesthetic Adjectives

Following [17-18], [16] normalized the mean ratings for 171 Mandarin single-concept adjectives on each of the five sensory modalities. In addition, the study measured the modality exclusivity for these adjectives as the range of the ratings divided by the sum, where the exclusivity scores thus range from 0 to 1: 0 meaning an entirely multimodal, and 1 meaning an entirely unimodal. Based on the exclusivity scores of the 171 Mandarin adjectives, we compare the extracted 72 Mandarin adjectives which have been attested to involve synaesthetic usages (see [8]) and the remaining 99 adjectives which have not been reported with synaesthetic uses. As demonstrated in Table 2, the adjectives with synaesthetic usages have a less exclusivity score on the average than the adjectives without synaesthetic usages. Moreover, the synaesthetic adjectives show a lower standard deviation.

We conduct an ANOVA test on the exclusivity for synaesthetic and non-synaesthetic adjectives. The result shows that the difference is significant ($p < 0.05$). Thus, it can be concluded that the modality exclusivity is negatively correlated with the synaesthetic usage of Mandarin sensory adjectives. In other words, if the adjectives are more multimodal (i.e., with lower exclusivity scores), they are more likely to be used for linguistic synaesthesia.

Table 2. Comparisons between Mandarin adjectives with and without synaesthetic uses

Type of adjectives	Count of adjectives	Sum of exclusivity	Mean of exclusivity	Standard deviation	Confidence interval
Synaesthetic adjectives	72	30.757	0.427	0.165	±0.038
Non-synaesthetic adjectives	99	60.855	0.615	0.208	±0.041

Another interesting pattern can also be observed with respect to the mean modality exclusivity scores of adjectives originally from different sensory domains. That is, adjectives originally from touch and vision show the two least modality exclusivity scores (with 0.366 and 0.450 respectively), while adjectives from hearing and smell have the two most modality exclusivity scores (with 0.841 and 0.525 respectively). The rank of the mean modality exclusivity for adjectives in five sensory domains, i.e., Touch (0.366) < Vision (0.450) < Taste (0.497) < Smell (0.525) < Hearing (0.841), is consistent with the reverse rank of the number of the adjectives in each modality, i.e., Touch (32) > Vision (26) > Taste (11) > Smell (2) > Hearing (1). Therefore, the mean modality exclusivity scores of adjectives from different sensory modalities would also suggest a negative correlation between the modality exclusivity and linguistic synaesthesia.

4 Directionality and Ratings of Mandarin Synaesthetic Adjectives

Table 3 presents three pieces of numeral information for five modalities based on the collected 72 Mandarin synaesthetic adjectives: (1) the first numeral is the synaesthetic transferability (i.e., the percentage of the adjective types in a sensory domain used for a specific modality in all the synaesthetic adjectives in the sensory domain, see [5]); (2) the second is the count of the synaesthetic tokens showing a specific transfer; and (3) the third number in the second line in each cell is the mean rating scores for adjectives in their non-original sensory domains. Take the tactile adjectives with synaesthetic transfers to taste for example. There are five tactile adjectives found with synaesthetic uses for taste among all the 32 tactile adjectives, thus the mapping from touch to taste with the synaesthetic transferability of 15.6% (5/32), where 121 synaesthetic examples are attested to show the transfer. Besides, the 32 tactile adjectives receive the mean rating score of 2.190 in the gustatory domain based on [16]'s exclusivity data.

We follow [8] to generalize the directionality of linguistic synaesthesia, based on both the lexical type and the lexical token (i.e., the two numerals in the first line in each cell in Table 3). Therefore, the synaesthetic transfers from touch to smell, from touch to hearing, and from taste to hearing show the absolute directionality, as the respective transfers with a reverse direction are unattested (cf. Table 3 for the numeral 0 for the transfers from smell to touch, from hearing to touch, and from hearing to taste). The synaesthetic transfers from touch to vision, from taste to smell, and from taste to vision demonstrate the tendencies-based directionality, whose synaesthetic

transferabilities and token examples are both larger than the respective transfers with a reverse direction, i.e., from vision to touch, from smell to taste, and from vision to taste. Although the synaesthetic transferability of the mapping from vision to smell (26.9%) is smaller than the one from smell to vision (50%), the number of the tokens of the former mapping is five times larger than that of the latter one (i.e., 52 vs. 10). In addition, as there are much more visual adjectives than olfactory adjectives used in linguistic synaesthesia (i.e., 26 vs. 2, see Section 2), the tendencies-based directionality could also be figured out for the transfer from vision to smell for the collected Mandarin synaesthetic adjectives. Similarly, the synaesthetic transfer from vision to hearing is also tendencies-based, whose transferability is close to that from hearing to vision, while the token examples of the former transfer are much more frequent than that of the latter one (i.e., 1849 vs. 1). Different from the transfers elaborated above, the number of the transfer tokens from touch to taste are more than that from taste to touch (i.e., 121 vs. 99), while the synaesthetic transferability for the former mapping is less than that of the latter one (i.e., 15.6% vs. 36.4%), both of which, however, are close. Therefore, we assume a bi-directionality for the synaesthetic transfers between touch and taste for the collected Mandarin synaesthetic adjectives. As the synaesthetic transfers from smell to hearing and from hearing to smell for the adjectives are both unattested, there is no synaesthetic directionality between these two sensory modalities. The transfer directionality of the collected 72 Mandarin synaesthetic adjectives is in line with the hierarchy of the whole Mandarin synaesthetic adjectives generalized by [8].

Table 3. The synaesthetic transferability, number of synaesthetic tokens, and rating score of Mandarin synaesthetic adjectives in five modalities

Original Domain	**Target Domain**				
	TOUCH	TASTE	SMELL	VISION	HEARING
TOUCH	-	15.6%|121 2.190	25%|17 0.853	87.5%|1563 2.653	53.1%|605 1.531
TASTE	36.4%|99 0.523	-	72.7%|94 2.867	63.6%|662 1.151	54.5%|87 0.457
SMELL	0|0 0.071	50%|22 3.808	-	50%|10 0.531	0|0 0.051
VISION	30.8%|205 2.757	23.1%|49 1.124	26.9%|52 0.617	-	92.3%|1849 1.899
HERAING	0|0 0.071	0|0 0	0|0 0.010	100%|1 0.860	-

Figure 1 summarizes the transfer hierarchy for the collected 72 Mandarin synaesthetic adjectives, where the mean rating scores for adjectives in their non-original sensory domains are also included. Specifically, the bold numbers are the rating scores of adjectives in the modalities consistent with the transfer directionality. For example, with respect to the mapping from touch to smell, 0.853 (in bold) means the score that tactile adjectives receive on smell, while 0.071 is the olfactory adjectives rated on touch. A closer look at the transfer directionality and the rating scores could suggest an intriguing correspondence between them. ANOVA tests can show significant differences ($p < 0.05$) between tactile adjectives rated on taste and

gustatory adjectives rated on touch (2.190 vs. 0.523) and between tactile adjectives rated on smell and olfactory adjectives rated on touch (0.853 vs. 0.071). Thus, it can be concluded that tactile adjectives are more preferred to be utilized and perceived in smell, rather than vice versa. Although the corpus data shows similar probabilities of synaesthetic transfers between touch and taste, the experimental data of ratings could add new evidence to reveal the preferred directionality from touch to taste in the collected Mandarin synaesthetic adjectives.

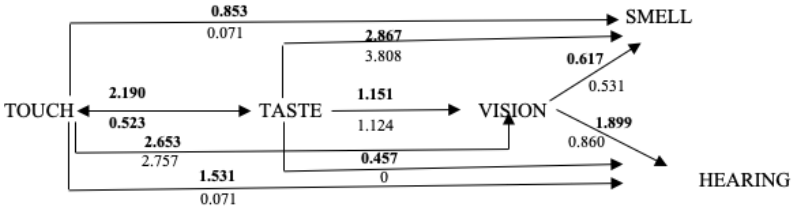

Fig. 1. The synaesthetic hierarchy with the transferred mean rating score

In addition, though the ANOVA test cannot work for the auditory domain with only one adjective, it could still be observed that hearing is more difficult to be rated on other sensory domains than to be perceived through other domains, just analogous to the pattern that hearing is less preferred to serve as the source domain than as the target domain in linguistic synaesthesia. To be more specific, the only one auditory adjective receives a less rating score on touch than tactile adjectives rated on hearing (i.e., 0.071 vs. 1.531), which tendency also appears in the relationship between hearing and taste and between hearing and vision (i.e., 0 vs. 0.457 and 0.860 vs. 1.899 respectively).

Although significant differences have not been shown by the ANOVA test for rating scores between touch and vision, between taste and smell, between taste and vision, and between smell and vision. However, two of them still exhibit a correspondence with the synaesthetic directionality, namely, gustatory adjectives with a higher rating score on vision than visual adjectives rated on taste (i.e., 1.151 vs. 1.124) and visual adjectives rated higher in smell than olfactory adjectives rated on vision (i.e., 0.617 vs. 0.531). It should be noted that the other two pairs, i.e., between touch and vision and between taste and smell, without a strict correspondence to the synaesthetic directionality, are both involving behaviorally and neutrally-integrated modalities, as argued by Winter (2016) (cf. Figure 1 for all the rating scores over 2.5 in the range from 0 to 5). This might override the directionality effect of linguistic synaesthesia. It can therefore be summarized that the transferred ratings of Mandarin synaesthetic adjectives are generally consistent with the transfer directionality of linguistic synaesthesia of these adjectives, which are generalized based on the distributions in the corpus.

5 Conclusion

This study employs both the corpus distributions and the experimental ratings of Mandarin synaesthetic adjectives to investigate linguistic synaesthesia. We have found that the tendencies attested based on the two different kinds of data are compatible, including: (1) the usage of linguistic synaesthesia shows a negative correlation with the modality exclusivity for Mandarin sensory adjectives; (2) the transfer directionality generalized based on corpus distributions are consistent with the sensory ratings of synaesthetic adjectives in their non-original sensory domains. These findings could reveal the cognitive reality of linguistic synaesthesia in both language production and perception. Furthermore, our findings could indicate the conceptual nature of linguistic synaesthesia, but not just a kind of linguistic expression.

Acknowledgments. We would like to thank The National Social Science Fund of China for the financial support on the project (No. 19CYY006).

References

1. Ullmann, S.: The principles of semantics. Basil Blackwell, Oxford (1957)
2. Williams, J. M.: Synaesthetic adjectives: A possible law of semantic change. Language 52(2) (1976) 461-478
3. Shen, Y.: Cognitive constraints on poetic figures. Cognitive Linguistics 8(1) (1997) 33-71
4. Strik Lievers, F.: Synaesthesia: A corpus-based study of cross-modal directionality. Functions of Language 22(1) (2015) 69-95
5. Zhao, Q., Huang, C.-R., Long, Y.: Synaesthesia in Chinese: A corpus-based study on gustatory adjectives in Mandarin. Linguistics 56(5) (2018)1167-1194
6. Zhao, Q., Xiong, J., Huang, C.-R.: Linguistic synaesthesia, metaphor, and cognition: The systematicity and significance of linguistic synaesthesia in Chinese [通感、隱喻與認知——通感現象在漢語中的系統性表現與語言學價值]. Zhuoguoyuwen [中國語文] 2 (2019) 240-253
7. Zhao, Q., Huang, C.-R.: A corpus-based study on synaesthetic adjectives in Modern Chinese. In: Lu, Q., Gao, H. H. (eds.): Chinese lexical semantics (LNAI 9332), Springer International Publishing, Gewerbestrasse (2015) 535-542
8. Zhao, Q.: Synaesthesia, metaphor, and cognition: A corpus-based study on synaesthetic adjectives in Mandarin Chinese. (Doctor of Philosophy), The Hong Kong Polytechnic University, Hong Kong (2018).
9. Zhao, Q. (趙青青), Huang, C.-R. (黃居仁).: Mapping models and underlying mechanisms of synaesthetic metaphors in Mandarin [現代漢語通感隱喻的映射模型與制約機制]. Language Teaching and Linguistic Studies [語言教學與研究] 1 (2018) 44-55
10. Shen, Y., Cohen, M.: How come silence is sweet but sweetness is not silent: A cognitive account of directionality in poetic synaesthesia. Language and Literature 7(2) (1998)123-140

11. Shen, Y., Eisenman, R.: “Heard melodies are sweet, but those unheard are sweeter”: Synaesthetic metaphors and cognition. Language and Literature 17(2) (2008) 107-121
12. Shen, Y., Gil, D.: Sweet fragrances from Indonesia: A universal principle governing directionality in synaesthetic metaphors. In: Auracher, J., Peer, W. (eds.): New beginnings in literary studies. Cambridge Scholars Publishing, Newcastle (2008) 49-71
13. Shen, Y., Gadir, O.: How to interpret the music of caressing: Target and source assignment in synaesthetic genitive constructions. Journal of Pragmatics 41(2) (2009) 357-371
14. Werning, M., Fleischhauer J., Beseoglu H.: The cognitive accessibility of synaesthetic metaphors. In: Proceedings of the 28th annual conference of the cognitive science society (2006) 2365-2370.
15. Chen, K.-J., Huang, C.-R., Chang L.-P., Hsu H.-L. Sinica corpus: Design methodology for balanced corpora. In: Proceedings of the 11th Pacific Asia Conference on Language, Information and Computation (PACLIC 11) (1996) 167-176
16. Chen, I., Zhao, Q., Long, Y., Lu, Q., Huang, C.-R.: Mandarin Chinese Modality Exclusivity Norms. PLOS ONE 14(2) (2019) e0211336
17. Lynott, D., Connell, L.: Modality exclusivity norms for 423 object properties. Behavior research methods 41(2) (2009) 558-564
18. Lynott, D., Connell, L.: Modality exclusivity norms for 400 nouns: The relationship between perceptual experience and surface word form. Behavior research methods 45(2) (2013) 516-526