

Auditory Synaesthesia and Near Synonyms: A Corpus-Based Analysis of sheng1 and yin1 in Mandarin Chinese

Qingqing Zhao¹Chu-Ren Huang²Hongzhi Xu³

The Department of Chinese and Bilingual Studies
The Hong Kong Polytechnic University

¹zhaoqingqing0611@163.com²churen.huang@polyu.edu.hk³hongz.xu@gmail.com

Abstract

This paper explores the nature of linguistic synaesthesia in the auditory domain through a corpus-based lexical semantic study of near synonyms. It has been established that the near synonyms 聲 sheng “sound” and 音 yin “sound” in Mandarin Chinese have different semantic functions in representing auditory production and auditory perception respectively. Thus, our study is devoted to testing whether linguistic synaesthesia is sensitive to this semantic dichotomy of cognition in particular, and to examining the relationship between linguistic synaesthesia and cognitive modelling in general. Based on the corpus, we find that the near synonyms exhibit both similarities and differences on synaesthesia. The similarities lie in that both 聲 and 音 are productive recipients of synaesthetic transfers, and vision acts as the source domain most frequently. Besides, the differences exist in selective constraints for 聲 and 音 with synaesthetic modifiers as well as syntactic functions of the whole combinations. We propose that the similarities can be explained by the cognitive characteristics of the sound, while the differences are determined by the influence of the semantic dichotomy of production/perception on synaesthesia. Therefore, linguistic synaesthesia is not a random association, but can be motivated and predicted by cognition.

1 Introduction

Synaesthesia is a phenomenon of one sensation connecting to another, which has been studied in two distinct disciplines. One is neuroscience,

which characterizes synaesthesia as a neural disorder (Cytowic, 1993), and the other one is linguistics that widely describes synaesthesia as a metaphor (Williams, 1976; Geeraerts, 2010). This paper is focused on the linguistic synaesthesia.

Synaesthesia occurs commonly and naturally in languages (Huang, 2015), such as “sweet voice” in English and 高音 gao-yin “high pitch” in Chinese, of which “sweet” and 高 gao “high” are normally perceived through gustation and vision respectively, while “voice” and 音 yin “sound” both belong to the auditory domain, therefore, the whole combinations exhibit an association of different sensations, namely synaesthesia.

In terms of research on linguistic synaesthesia, most previous studies are concentrated on the directionality of synaesthetic transfers, that is, which sensation usually acts as the source domain and which sensation is the target domain. For instance, Ullmann (1957) studied creative usages of synaesthesia on 2000 examples from poems in the 19th century, and proposed a tendency of hierarchical distribution for synaesthesia. He concluded that synaesthetic transfers are usually from much “lower domains” (such as touch and taste) to much “higher domains” (such as vision and hearing)¹, and the acoustic field emerges as the main recipient. Williams (1976) also claimed that the diachronic meaning change of synaesthetic adjectives in ordinary English obeys a strict rule that can be universal for all human languages, as shown in Figure 1.

¹The terms, “lower domains” and “higher domains”, are copied from Ullmann (1957), where the former refers to touch, taste and smell, and the later includes hearing and vision.

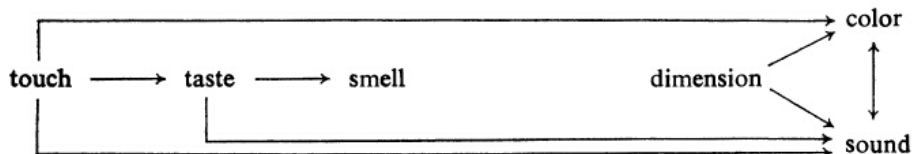


Figure 1: The Meaning Change of Synaesthetic Adjectives (Source: Williams, 1976).

In response to Williams (1976), Zhao and Huang (2015) carried out a study on synaesthesia in the Chinese language and demonstrated that the hierarchy proposed based on English is not applicable to Chinese. Similarly, Strik Lievers (2015) also pointed out that the “directionality principle” reflects the frequency of association types, rather than representing universal constraints on synaesthetic transfers. Thus, linguistic synaesthesia should be investigated specifically for different languages in depth. Unfortunately, fine-grained studies on synaesthesia in Chinese are still scarce.

Given the gap on synaesthesia in Chinese, we think that it could be a good way to start off with examining synaesthesia in near synonyms, because the interaction between synaesthesia and subtle semantic differences in near synonyms could provide much deeper and more fine-grained clues for synaesthesia. The benefits of studies on near synonyms to explain and predict linguistic phenomena have been widely recognized. For example, Chief et al. (2000) suggested that some semantic features obtained by comparing near synonyms can be useful on the prediction of syntactic performances. Similarly, Hong and Huang (2004), Hong and Huang (2005) also pointed that semantic/cognitive features embedded in perceptual near synonyms can influence their usages in the language.

As previous studies show that the auditory domain is the main target of synaesthetic transfers (Ullmann, 1957; Williams, 1976), the near synonyms, 聲 sheng “sound” and 音 yin “sound” in Mandarin Chinese, should be good candidates for our research, which have been established to represent different semantic/cognitive focuses on the sound. Specifically speaking, 聲 is concentrated on the auditory production, while 音 is focused on the auditory perception, although both can refer to the sound (Hong and Huang, 2004; Hong and Huang, 2005). Therefore, through studying the synaesthetic usages of

this pair of near synonyms, not only can we test whether synaesthesia is sensitive to the semantic dichotomy of production and perception, but also we can examine the relationship between synaesthesia and cognition.

The present paper is organized as follows: we will introduce the methods of collecting and annotating data in Section 2, and figure out the similarities as well as differences between 聲 and 音 on synaesthesia in Section 3, which will be followed by some explanations in Section 4. In the last Section, we will summarize our main findings and propose our future work.

2 Methodology: Corpus Selection and Data Annotation

2.1 Corpus Selection and Measurement Criteria

In order to make our study more tenable, we attempt to exhaust the data as possible as we can, and hence rely on the corpus for data collection. We select four widely-used huge corpora as our data sources, of which Sinica Corpus² (Chen et al., 1996) is the main corpus and other three corpora, including Chinese GigaWord 2 Corpus (Mainland, simplified)³; Chinese GigaWord 2 Corpus (Taiwan, traditional)⁴; and the journal corpus of BCC Corpus⁵, act as the complement.

The reason for giving the priority to Sinica Corpus is on the consideration that the corpus is a well-recognized balanced and high-quality corpus with

²Accessed at: <http://app.sinica.edu.tw/kiwi/mkiwi/>

³Accessed at: https://the.sketchengine.co.uk/bonito/run.cgi/first_form?corpname=preloaded/cgw2_sc

⁴Accessed at: https://the.sketchengine.co.uk/bonito/run.cgi/first_form?corpname=preloaded/cgw2_tc

⁵Accessed at: <http://bcc.blcu.edu.cn/index.php?corpus=2>

tagging information (Chen et al., 1996), which we think can facilitate our research.

Besides, usages in newspapers are normally well-established and the controversial issues about grammaticality for some specific examples could be decreased to a degree, and we hence select three journal corpora as the complement (Refer to Hong and Huang, 2006 for the introduction of two GigaWord corpora).

Regarding the data from different corpora, we set some measurement criteria to make them comparable. To be specific, we depend on the statistical information, which includes the frequency information of examples per million word tokens in their respective corpora and also percentage information of each example in its source corpus. For instance, 大聲 da-sheng “loudly” occurs 407 times in Sinica Corpus (4.0 edition) as a “stative and intransitive verb”⁶ and word tokens in the corpus is around 10 million, so its calculated frequency is 40.7 and the percentage is 0.00407% for the analysis in this paper. For some complementary examples, which do not appear in Sinica Corpus, we also utilize the criterion to calculate its frequency in their relevant corpora.⁷ For example, the calculated frequency of 長聲 chang-sheng “in prolonged voice” is 0.01 and hence its percentage is 0.000001%, because it occurs 10 times in the journal corpus of BCC with around 1000 million word tokens.

2.2 Data Collection and Annotation

We exhaust all the examples with two characters in the modifier-head relation in Sinica Corpus, whose heads are either 聲 or 音. If the modifier does not belong to audition but can be classified into any of four other sensations, including touch, gustation, olfaction and vision, we think that this example involves synaesthesia.

In terms of determining the sensory domain of modifiers, we refer to its original meaning in 說文解字 shuo-wen-jie-zi and 漢典 han-dian⁸. Take the meaning of 雜 za “mixed” in 說文解字 for example, it is paraphrased as 五彩相會 wu-cai-xiang-hui “five colors mixing together”, therefore, we classify

⁶The term is used in Sinica Corpus, and marked as VH.

⁷For more detailed information about the size and design of these corpora, please refer to the websites mentioned above.

⁸Accessed at: <http://www.zdic.net/>

it into the visual domain.

Moreover, we divide some sensory domains further into several sub-domains. For instance, we distinguish vision into size, dimension and so on.

All the synaesthetic examples we find in Sinica Corpus are words⁹ with the information of part of speech, so we calculate the frequency of each example according to different part of speech labels. For some examples that do not appear in Sinica Corpus, we annotate them manually with the part of speech information compared to Sinica Corpus, and utilize the measurement rule mentioned above to obtain their frequencies and percentages for discussion. If one complementary example has different frequencies and percentages in additional corpora, we take the most frequent one and the respective corpus into consideration.

Therefore, our data is summarized as follows in Table 1 and Table 2, with the part of speech¹⁰, frequency and percentage information¹¹.

3 Similarity and Difference on Synaesthesia

3.1 Synaesthetic Similarity for the Near Synonyms

As shown in Table 1 and Table 2, we can see that there are some similarities on synaesthesia between the near synonyms.

General speaking, as 聲 and 音 both belong to the auditory sensation, other modalities can transfer frequently to describe these two near synonyms, of which morphemes from both visual and tactile domains can be used to modify 聲 and 音. Therefore, hearing is also a productive target domain of synaesthetic transfers in Mandarin Chinese. This is in line with the observation for English (Ullmann, 1957; Williams, 1976).

⁹We do not plan to involve the controversial issue about how to determine a sequence is a word or not, for the discrimination will not influence the discussion below.

¹⁰Word/POS/Freq/Perc in Table 1 and Table 2 refers to word, part of speech, frequency and percentage. Please note that we group VH and D, although VH denotes “stative and intransitive verbs” and D denotes “adverbs” in Sinica Corpus, for both of them are related to events, rather than referring to an entity.

¹¹Please also note that although there are some polysemous examples in our data, some senses are irrelevant to synaesthesia, such as 美音 mei-yin “pronunciation in American English”. We exclude this type of usages.

Source Domain	聲 sheng1 “sound”		音 yin1 “sound”	
	Word/POS/Freq/Perc	Word/POS/Freq/Perc	Word/POS/Freq/Perc	
VISION	size	大聲/VH/40.7/0.00407% “loudly”		
		小聲/VH/4.3/0.00043% “in a low voice”		
	dimension	高聲/D/5.4/0.00054% “loudly”		高音/Na/4.6/0.00046% “high pitch”
		低聲/D/6.3/0.00063% “lowly”		低音/Na/4.6/0.00046% “low pitch”
				中音/Na/2.9/0.00029% “mediant”
	light	朗聲/D/0.1/0.00001% “in a clear voice”		
		陰聲/D/0.1/0.00001% “in a deep voice”		
	shape	尖聲/D/1.2/0.00012% “in a sharp voice”		尖音/Na/0.2/0.00002% “sharp sound”
			平聲/Na/0.4/0.00004% “level tone”	
	evaluation	齊聲/D/4/0.0004% “in chorus”		
			美聲/Na/0.4/0.00004% “bel canto”	美音/Na/0.2/0.00002% “beautiful sound”
	length	長聲/D/0.01/0.000001% “in a prolonged voice”	長聲/Na/0.03/0.000003% “prolonged sound”	長音/Na/0.2/0.00002% “prolonged sound”
			短聲/Na/0.013/0.0000013% “short sound”	短音/Na/0.018/0.0000018% “short sound”
	color	雜聲/Na/0.014/0.0000014% “noise”		雜音/Na/3.5/0.00035% “noise”
transparency	清聲/D/0.001/0.0000001% “in a clear voice”	清聲/Na/0.002/0.0000002% “clear sound”	清音/Na/0.3/0.00003% “voiceless sound”	
			濁音/Na/0.01/0.000001% “voiced sound”	

Table 1: Synaesthetic Examples of 聲 and 音 from Vision in Mandarin Chinese

Specifically, for 聲 and 音, different modalities have different transferability. Vision transfers more frequently to the auditory domain than touch, which can be reflected in both types and frequencies. There are 14 morphemes from the visual domain that can

modify 聲 and 10 morphemes from the tactile domain. Similarly, 10 morphemes from the vision can be used to describe 音 and only 6 morphemes from touch. In terms of frequencies of synaesthetic examples, the most frequent modifier for 聲 is 大 da “big”

Source Domain		聲 sheng1 “sound”	音 yin1 “sound”
		Word/POS/Freq/Perc	Word/POS/Freq/Perc
TOUCH	weight	輕聲/D/8/0.0008% “in a soft voice”	輕音/Na/0.006/0.000006% “light tone”
		沉聲/VH/1.4/0.00014% “in a heavy voice”	重音/Na/0.6/0.00006% “stress”
	texture	粗聲/VH/0.5/0.00005% “raucously” 細聲/D/0.4/0.00004% “in a soft voice” 柔聲/D/2/0.0002% “in a soft voice” 軟聲/VH/0.001/0.000001% “in a soft voice” 硬聲/D/0.001/0.000001% “in a hard voice”	細音/Na/0.001/0.000001% “tiny sound” 柔音/Na/0.002/0.000002% “soft sound” 軟音/Na/0.001/0.000001% “soft sound” 滑音/Na/1/0.0001% “smooth sound”
temperature	寒聲/VH/0.1/0.00001% “in a cold voice” 溫聲/D/0.007/0.000007% “in a warm voice” 冷聲/D/0.003/0.000003% “in a cold voice”	溫聲/Na/0.001/0.000001% “warm sound”	
TASTE	taste	甜聲/D/0.001/0.000001% “in a sweet voice”	

Table 2: Synaesthetic Examples of 聲 and 音 from Touch and Taste in Mandarin Chinese

from the visual domain. Similarly, 高 gao “high” and 低 di “low” modify 音 most frequently with the same frequency, which are both from vision. If we calculate the ratio of all the frequencies of morphemes from vision to those from touch, the ration for 聲 is around 5 times and that for 音 is about 10 times. It seems that vision is much easier to transfer into hearing than touch in Mandarin Chinese, which has not been mentioned in previous research yet.

In summary, the similarities on synaesthesia between 聲 and 音 exist in two aspects. The first one is that both 聲 and 音 can be productive recipients

of synaesthetic transfers, which can be described by visual words and tactile words. The other similar property is concerned with the priority of vision to touch when modifying these two synonyms. In other words, vision is used more frequently to characterize 聲 and 音 than touch.

3.2 Synaesthetic Difference for the Near Synonyms

Besides similarities, there are also two differences on synaesthesia for 聲 and 音, one of which is reflected in selective constraints and the other exists in the syntactic function of the whole combinations.

3.2.1 Selective Constraints

As we see in Table 1 and Table 2, there are more domains and more modifiers that can be used for 聲 than 音. 聲 can be naturally described by gustation, but 音 cannot. There are totally 25 morphemes modifying 聲, including 14 from vision, 10 from touch and 1 from taste, while only 16 morphemes modifying 音, including 10 from vision and 6 from touch. In terms of the frequencies of words, words containing 聲 occur more frequently in corpora than those containing 音. Therefore, 聲 seems to be much easier and more common to be described than 音 through synaesthesia.

Interestingly, there is another difference between 聲 and 音 on synaesthesia about the selection constraint. That can be instantiated in the symmetry of modifiers selection for sub-domains. Polar items in gradable antonymous relations (Lyons, 1977) in each sub-domain are usually selected for synaesthetic usages, such as 大 da “big” and 小 xiao “small” for the size and 清 qing “clear” and 濁 zhuo “turbid” for the transparency. For some sub-domains, there exists the symmetry between modifiers selection. For examples, in terms of the dimension domain for 聲, the modifiers selection is symmetrical, because the polar items 高 gao “high” and 低 di “low” are both selected and there is no other in-between item on the gradable axis that has been selected. However, in terms of the dimension domain for 音, the selection is asymmetrical, for 中 zhong “middle” is also used besides the polar items 高 and 低. In this way, we can calculate the symmetry of modifiers selection for 聲 and 音 in each sub-domain, which is shown in Table 3. Thus, the whole symmetry ratio for 聲 is 50%, while that for 音 is only 38%.

Therefore, there is a noteworthy difference on selective constraints for the near synonyms on synaesthesia. Specifically speaking, 聲 receives more synaesthetic modifications than 音, and also employs higher symmetry on modifiers selection in each sub-domain.

3.2.2 Syntactic Function

Another difference on synaesthesia between 聲 and 音 is related to the whole syntactic function of the synaesthetic combinations. All the examples containing 音 refer to the entity, that is, a specific kind of

Sub-domains	聲 sheng1 “sound”	音 yin1 “sound”
size	+	NA
dimension	+	-
light	+	NA
shape	-	-
evaluation	-	-
length	+	+
color	-	-
transparency	-	+
weight	+	+
texture	+	-
temperature	-	NA
taste	-	NA

Table 3: The Synaesthetic Symmetry of 聲 and 音 in Mandarin Chinese.

sound, such as 長音 chang-yin “prolonged sound”. However, most examples containing 聲 is concerned with the event, such as 大聲 da-sheng “loudly”, for 大聲 can only be used to modify a predicate (e.g., 大聲唱歌 da-sheng-chang-ge “sing loudly”) or express a state (e.g., 他說話很大聲 ta-shuo-hua-hen-da-sheng “his talking is loud”), whereas it cannot denote an entity (e.g., *一個大聲 yi-ge-da-sheng “a loud sound”).

As Table 1 and Table 2 shows, only 4 adjective morphemes exclusively form a noun when combining with 聲. In terms of these 4 exceptional examples, we can see that two of them are terminological words in some specific areas, namely 美聲 mei-sheng “bel canto” and 平聲 ping-sheng “level tone”. In addition, the other two are used in a very low frequency, and 短聲 duan-sheng “short sound” as well as 雜聲 za-sheng “noise” neither occur in Sinica Corpus, with the calculated frequencies of only 0.013 and 0.014 respectively in the journal corpus of BCC. Therefore, the synaesthetic combinations of 聲 and 音 have different syntactic functions, of which combinations with 聲 are normally related to an event, whereas combinations with 音 always refer to an entity.

In conclusion, although 聲 and 音 are near synonyms and both can denote the sound; there are some noteworthy differences on synaesthesia. 聲 is much easier to be described through synaesthesia than 音, and also employs higher symmetry on synaesthetic selection of modifiers. Moreover, the combinations of 聲 or 音 have different syntactic

functions, of which combinations with 聲 are normally related to an event, while combinations with 音 always refer to an entity.

4 Production and Perception in Sound

As a pair of near synonyms, 聲 and 音 have both similar properties and different characteristics on synaesthesia in Mandarin Chinese, which can be explained by combining the framework of Generative Lexicon (Pustejovsky, 1995) and the semantic/cognitive dichotomy of production and perception in the sound (Hong and Huang, 2004).

As Pustejovsky (1995) proposed that lexical items can be represented in Lexical Inheritance Structure and some semantic information encoded in lexical items are inherited from upper concepts, which determines their performances in languages. Both 聲 and 音 are in the auditory domain, which can denote the sound, thus, both of them can inherit common semantic information from the upper concept, namely sound. Therefore, we think that the similarities on synaesthesia between 聲 and 音 are determined by the whole synaesthetic characteristics for hearing, that is, hearing is a productive recipient of synaesthetic transfers and is much easier to be the target domain of vision than touch.¹²

In terms of synaesthetic differences between 聲 and 音, the semantic/cognitive dichotomy of production and perception really play a role. Hong and Huang (2004) suggested that in the whole transmission process of sounds it should include the starting point, the process and the ending point, of which some concepts focus on the starting point, namely the production of the sound, such as 聲, and some concepts concentrate on the ending point, namely perception and evaluation of the sound, such as 音. According to this proposal, the selective differences on synaesthesia between 聲 and 音 can be reasonably explained. From the cognitive point of view (Gibbs, 2006), embodied action can influence conceptualization in languages. Compared with the ending point of the sound, humans can impose much more active actions on the starting point, whereas the perception and evaluation can be received much

¹²Frankly speaking, we have not figure out the reason why vision transfers more easily into hearing than touch. However, we think that this issue will not influence the hypothesis proposed in this paper, and should be involved in our future work.

more passively. Therefore, the production of the sound is conceptualized and characterized more frequently than the perception through synaesthesia. Also, we speculate that the volition may result in the higher symmetry on sound production for 聲, which needs further research.

The difference on syntactic functions of synaesthetic combinations for 聲 and 音 can be explained by the combination of Generative Lexicon and the distinction between production and perception. In Generative Lexicon, qualia structure of a lexical item, including formal role, constitutive role, telic role and agentive role, can predict their performances in languages (Pustejovsky, 1995). As 聲 focuses on the production of the sound, the salient qualia role in its semantic information is the agentive role, which is usually related to an event or predicate that brings something into being. Conversely, synaesthetic modifiers always contribute to the formal role of 音, for 音 focuses on perception and evaluation of the sound.

5 Conclusion

This paper is devoted to a fine-grained study on the interaction between synaesthesia and the near synonyms 聲 and 音 of the auditory domain in Mandarin Chinese. We take a corpus-based approach and at the same time employ a combination method of qualitative and quantitative analyses. Eventually, some interesting findings are obtained.

In particular, 聲 and 音, as a pair of near synonyms in the auditory domain, have both similar properties and different characteristics on synaesthesia. In terms of similarities, both 聲 and 音 can be productive recipients of synaesthetic transfers and both can be described by vision and touch. Moreover, vision is observed to act as the most predominant source domain for these near synonyms. In addition, the differences on synaesthesia between 聲 and 音 are also apparent. 聲 receives much more synaesthetic modifications than 音, and also has a higher symmetry on synaesthetic modifiers selection. On the other hand, the combinations of synaesthetic modifications with 聲 are normally related to an event (i.e., production of the sound), while those with 音 always refer to an entity (namely the perception or evaluation of the sound).

The similarities and differences mentioned above can be explained by semantic information and cognitive motivations. Both 聲 and 音 can denote the sound, so they can inherit common semantic information from the sound, which result in the similarity on synaesthesia. However, the different cognitive focuses on the sound, namely the distinction between production and perception, make 聲 and 音 have different synaesthetic performances.

In general, we can conclude that auditory synaesthesia is sensitive to the cognitive dichotomy of production/perception in Mandarin Chinese. In other words, synaesthesia can be predicted and determined by the cognitive modelling. Therefore, besides the transfer tendency recognized by previous studies, linguistic synaesthesia exhibit another regularity, that is, it can be determined by cognitive motivations. Therefore, linguistic synaesthesia is not a kind of random associations of different sensations, but of many principles and regularities, which deserves much more attention and deeper studies.

In the following study, we will expand the research scope and test more interactions between synaesthesia and the cognitive modelling, through which we hope that our studies could contribute to bridging the research on synaesthesia in linguistics and that in neuroscience eventually.

Acknowledgments

We would like to give thanks to Dennis Tay from the Hong Kong Polytechnic University for his insightful comments on this work.

References

- Chu-Ren Huang. 2015. Towards a Lexical Semantic Theory of Synaesthesia in Chinese. *Keynote Speech in the 16th Chinese Lexical Semantics Workshop (CLSW-16)*. Beijing.
- Dirk Geeraerts. 2010. *Theories of Lexical Semantics*. Oxford University Press, New York.
- Francesca Strik Lievers. 2015. Synaesthesia: A Corpus-Based Study of Cross-Modal Directionality. *Sensory Perceptions in Language and Cognition*: 69-95
- Jia-Fei Hong and Chu-Ren Huang. 洪嘉馥, 黃居仁. 2004. A Comparative Analysis on the Near Synonyms sheng1 and yin1: the Relation between Lexical Semantics and Concepts 「聲」與「音」的近義辨析: 詞義與概念的關係. *In the International Conference on the Situations and Trends about Chinese Lexical Semantics 漢語詞彙語意研究的現狀與發展趨勢國際學術研討會*. Peking University 北京大學.
- Jia-Fei Hong and Chu-Ren Huang. 洪嘉馥, 黃居仁. 2005. A Comparative Analysis on Perceptual Verbs with Near Synonyms: the Relation between Lexical Semantics and Concepts, 感官動詞的近義辨析: 詞義與概念的關係. *In the 6th Chinese Lexical Semantics Workshop (CLSW-6)*: 20-24. Xiamen.
- Jia-Fei Hong and Chu-Ren Huang. 2006. Using Chinese GigaWord Corpus and Chinese Sketch Engine in Linguistic Research. *In the Proceeding of the 20th Pacific Asia Conference on Language, Information and Computation (PACLIC-20)*: 183-190. Wuhan.
- John Lyons. 1977. *Semantics (1)*. Cambridge University Press, New York.
- James Pustejovsky. 1995. *The Generative Lexicon*. The MIT Press, Cambridge.
- Joseph M. Williams. 1976. Synaesthetic Adjectives: A Possible Law of Semantic Change. *Language* 52(2): 461 – 478.
- Keh-jiann Chen, Chu-Ren Huang, Li-ping Chang, and Hui-Li Hsu. 1996. Sinica Corpus: Design Methodology for Balanced Corpora. *In the Proceeding of the 11th Pacific Asia Conference on Language, Information and Computation (PACLIC-11)*:167-176. Seoul.
- Lian-Cheng Chief, Chu-Ren Huang, Keh-Jian Chen, Mei-Chih Tsai and Li-li Chang. 2000. What Can Near Synonyms Tell Us. *Computational Linguistics and Chinese Language Processing* 5(1): 47-60.
- Qingqing Zhao and Chu-Ren Huang. 2015. A Corpus-Based Study on Synaesthetic Adjectives in Modern Chinese. *In the 16th Chinese Lexical Semantics Workshop (CLSW-16)*. Beijing.
- Richard E. Cytowic. 1993. *The Man Who Tasted Shapes*. MIT Press, Massachusetts.
- Raymond W. Gibbs. 2006. *Embodiment and Cognitive Science*. Cambridge University Press, New York.
- Stephen Ullmann. 1957. *The Principles of Semantics*. Basil Blackwell, Oxford.