

# Eye movements during inspecting pictures of natural scenes for information to verify sentences

Chen Qingrong<sup>1</sup> Jiang Zhijie<sup>2</sup>

(<sup>1</sup>Department of Psychology, Nanjing Normal University, Nanjing 210097, China)

(<sup>2</sup>School of Geography, Nanjing University, Nanjing 210093, China)

**Abstract:** As eye tracking can be used to record moment-to-moment changes of eye movements as people inspect pictures of natural scenes and comprehend information, this paper attempts to use eye-movement technology to investigate how the order of presentation and the characteristics of information affect the semantic mismatch effect in the picture-sentence paradigm. A 3 (syntax) × 2 (semantic relation) factorial design is adopted, with syntax and semantic relations as within-participant variables. The experiment finds that the semantic mismatch is most likely to increase cognitive loads as people have to spend more time, including first-pass time, regression path duration, and total fixation duration. Double negation does not significantly increase the processing difficulty of pictures and information. Experimental results show that people can extract the special syntactic strategy from long-term memory to process pictures and sentences with different semantic relations. It enables readers to comprehend double negation as affirmation. These results demonstrate that the constituent comparison model may not be a general model regarding other languages.

**Key words:** natural scene; semantic mismatch; double negation; eye movement

How do people compare and integrate pictures of natural scenes and texts? One of the interpretations is the constituent comparison model (CCM)<sup>[1]</sup>. The CCM is regarded as a reliable approach in dealing with a great variety of language processes. The underlying premise of the model is that people build a representation of a sentence and a picture that is internally represented in an abstract propositional format, and then the corresponding constituents from the sentences and the pictures are serially compared. Moreover, it suggests that the difficulty of sentence-picture integration depends on their semantic relations: either match or mismatch in the CCM. Furthermore, the mismatch effect is not destroyed by the verification task<sup>[2-4]</sup>.

In recent decades, negation and double negation have been topics of research in linguistics and reasoning<sup>[5-6]</sup>. It is found that negative sentences are difficult to be comprehended and processed in some established literature. Double negation is a special form of negative syntax. If the CCM does hold, we can deduce from the model that readers require longer time to comprehend and verify double negative sen-

tences than affirmative and negative sentences because double negative sentences can be represented as the abstract proposition like {NEG, [NEG, (RED, DOTS)]}. The validity of the CCM has also been challenged by recent eye-movement studies which identify variances such as word orders. The order of sentence-picture presentation influences the integration of sentences and pictures<sup>[7-10]</sup>.

In the current experiment, a complex Chinese picture-sentence task in which the pictures are taken from natural scenes is conducted to address two questions. First, whether the CCM is applicable to other languages and whether the mismatch effect disappears when the picture is first presented in Chinese. Secondly, whether negation, especially double negation, as a special negative form, is always more difficult than affirmation and negation. In contrast to English, in Chinese two negative adverbs can be used together, or be separated only by a particle, such as “mei you bu”, “bu de bu” in the simple sentence. The structure is described as “not-no” or “not-particle-not”.

## 1 Method

### 1.1 Participants

Forty-four undergraduate students of Nanjing Normal University (20 males, 24 females, mean age is 20.3 years) are chosen to participate in the experiment. All the participants are native speakers of Mandarin Chinese and have normal or corrected-to-normal vision. To be sure that the data would not be biased by systematic error, two participants who are reported tracker loss are excluded.

### 1.2 Apparatus

Participants' eye movements are recorded via the SensoMotoric Instruments iView Hi-Speed eye tracker with a high spatial resolution (noise < 0.01°) and a sampling rate of 350 Hz. The left eye is monitored. A chin rest is used to stabilize the head. Participants are seated 65 cm from the computer screen. Each Chinese character subtends approximately a visual angle of 1° both in width and in height.

### 1.3 Materials

The photographs which portray general scenes such as bus stations, shops, restaurants, squares, footpaths and so on are taken in the area on or around a university campus. Each picture depicts some people, objects (e. g., bicycles, desks) and environmental features (e. g., road markings, signs). All of the scenes are readily recognizable, and none contain bizarre events. The pictures subtend angles of 12° and 10°, respectively. Three different kinds of syntax (e. g., affirmation, negation, and double negation) are selected for the experiment. For Chinese character frequency,

Received 2009-09-30.

**Biography:** Chen Qingrong (1980—), male, doctor, jscqr80@sina.com.

**Foundation items:** The National Social Science Foundation of China (No. CBA080236), the Graduate Innovation Project of Jiangsu Province (No. CX08B\_016R).

**Citation:** Chen Qingrong, Jiang Zhijie. Eye movements during inspecting pictures of natural scenes for information to verify sentences[J]. Journal of Southeast University (English Edition), 2010, 26(3): 444 – 447.

the means of the affirmative sentence, the negative sentence, and the double negative sentence are 0.008 3 (SD = 0.012 3), 0.009 2 (SD = 0.012 3), 0.008 6 (SD = 0.011 7) in 1.8 million characters, according to *The Frequency Dictionary of Modern Chinese Words*. The difference of character frequency is not significant.

1.4 Procedure

Each student was tested individually. A 13-point calibration routine was performed by WINCAR. Before inspecting each picture, participants were told to fixate on a cross in the center of the computer screen. Once the participants pressed the space bar, a picture was displayed in the center. After inspecting each picture, the participants pressed the space bar to remove the picture. Then the participants must fixate on a cross on the left of the screen. If the participants pressed the space bar, a sentence appeared and remained on the screen until one of the two response keys was pressed. At the end of the trial, participants were instructed to judge whether the meaning of sentence and picture was identical or not by pressing the “1” or “2” key respectively. The interval between the end of one trial and the onset of the next trial is 1 s.

2 Results

Prior to all the analyses, fixations less than 80 ms in duration are excluded. The overall accuracy is high (>80%), but only correct responses are analyzed. Less than 2% of the fixations are eliminated in subsequent analyses. We analyze two regions in each sentence. The first region is the adverbial phrase including affirmative, negative, and double negative phrases. The second region is a spillover region consisting of the remainder of the sentence. Three reading time measures are computed for each region: total fixation duration, first-pass time, and regression path duration. Total fixation duration is the duration of time of all the fixations made on the sentence or the picture. The first-pass time (FPT) is the sum of all the fixations in a region prior to moving to another region<sup>[11]</sup>. This measure reflects the early stage of processing<sup>[12]</sup>. The regression path duration(RPD) is the duration of time that the reader looks at the target as well as any time spent rereading earlier parts of the sentence before moving ahead to inspect new components of the sentence. Rayner et al.<sup>[13]</sup> argued that regression path duration reflected lexical processing and integration processing.

2.1 Adverbial phrase region

Tab. 1 presents the mean reading time of the first-pass time, the regression path duration, and the total fixation duration (TFD) in the adverbial phrase regions (standard deviation in brackets).

In the analysis of the first-pass time, there is a significant

**Tab.1** Mean reading time of three dependent measures in adverbial phrase regions ms

Measure	Affirmative phrase		Negative phrase		Double negative phrase	
	Match	Mismatch	Match	Mismatch	Match	Mismatch
FPT	206(96)	218(81)	271(79)	279(82)	249(89)	266(90)
RPD	321(99)	331(92)	395(87)	409(98)	332(93)	335(84)
TFD	408(82)	417(64)	493(91)	502(85)	429(57)	426(61)

main effect of the adverbial phrase type,  $F(2,78) = 9.5$ ,  $p < 0.001$ . Post hoc paired comparisons show that there are significant differences between affirmative and negative phrases,  $F(1,39) = 13.96$ ,  $p < 0.01$ , and between affirmative and double negative phrases,  $F(1,39) = 10.73$ ,  $p < 0.01$ , whereas there are no significant differences between negative and double negative phrases. Neither the main effect of the semantic relations, nor the interactions between the phrasal type and the semantic relation is significant.

The main effects of the adverbial phrase type remained in the regression path duration,  $F(2,78) = 10.81$ ,  $p < 0.001$ . Likewise, not only the negative and affirmative phrases  $F(1,39) = 19.81$ ,  $p < 0.001$ , but also the negative and double negative phrases ( $F(1,39) = 10.83$ ,  $p < 0.01$ ) show statistically significant differences. The difference between affirmative and double negative phrases is not significant, nor is it significant for the main effect of semantic relations and for the interaction between the phrase types and the semantic relations.

Regarding the total fixation duration measure, the main effect of the adverbial phrase type is significant,  $F(2,78) = 33.63$ ,  $p < 0.001$ . There are significant differences between negative and affirmative phrases,  $F(1,39) = 59.18$ ,  $p < 0.001$ , between negative and double negative phrases,  $F(1,39) = 37.97$ ,  $p < 0.001$ . However, the difference between affirmative and double negative phrases is not significant, nor is it significant for the semantic relation or for the interaction between the phrase type and the semantic relation.

2.2 Spillover region

Tab. 2 presents the first-pass time, the regression path duration, and the total fixation duration made in the spillover region (standard deviation in brackets).

**Tab.2** Mean reading time of three dependent measures in spillover region ms

Measure	Affirmative sentence		Negative sentence		Double negative sentence	
	Match	Mismatch	Match	Mismatch	Match	Mismatch
FPT	409(65)	427(71)	422(80)	449(63)	413(90)	435(75)
RPD	606(76)	649(64)	704(75)	810(57)	619(89)	654(86)
TFD	510(66)	529(74)	556(83)	600(81)	508(67)	522(78)

In the analysis of the first-pass time, there is no significant main effect of the sentence type, but a significant main effect of the semantic relation,  $F(1,39) = 7.06$ ,  $p < 0.05$ . The interaction between the sentence type and the semantic relation does not approach significance. The first-pass time is increased by 22 ms under the mismatch condition.

In general, the regression path duration in the spillover region shows a main effect of sentence type,  $F(2,78) = 21.3$ ,  $p < 0.001$ . Post hoc paired comparisons show that the differences between the negative sentence and the other two sentences are significant,  $F_1(1,39) = 38.03$ ,  $p < 0.001$ ,  $F_2(1,39) = 21.98$ ,  $p < 0.001$ , but there is no significant difference between the affirmative sentence and the double negative sentence. Furthermore, the main effect of the semantic relation is significant,  $F(1,39) = 14.1$ ,  $p < 0.01$ , but there is no significant difference between the sentence type and the semantic relation. The regression path duration is decreased by 61 ms under the congruous condition.

On the total fixation duration measure, the main effect of the sentence type is significant,  $F(2, 78) = 19.35$ ,  $p < 0.001$ . Also, the differences between the negative sentence and the other two types of sentences are significant,  $F_1(1, 39) = 23.44$ ,  $p < 0.001$ ,  $F_2(1, 39) = 34.14$ ,  $p < 0.001$ , but no significant difference between affirmative and double negative sentences. Furthermore, the main effect of the semantic relation is significant,  $F(1, 39) = 7.66$ ,  $p < 0.01$ , but the interaction between the sentence type and the semantic relation is not significant. The total fixation duration is increased by 26 ms in the mismatch condition.

### 3 Discussion

As hypothesized, affirmative, negative, and double negative sentences have different syntactic complexities in the experiment, which might influence the eye fixations on the collection of information from the picture and on the reading comprehension. Compared with previous studies, this paper enlarges the debate about CCM in three ways. First, research demonstrates that the double negative effect disappeared. It signifies that the degree of difficulty verifying the double negative sentence is equivalent to the affirmative sentence (see Fig. 1). At this point, the effect of double negation provokes questions about the general validity of the CCM in the integration of picture and sentence.

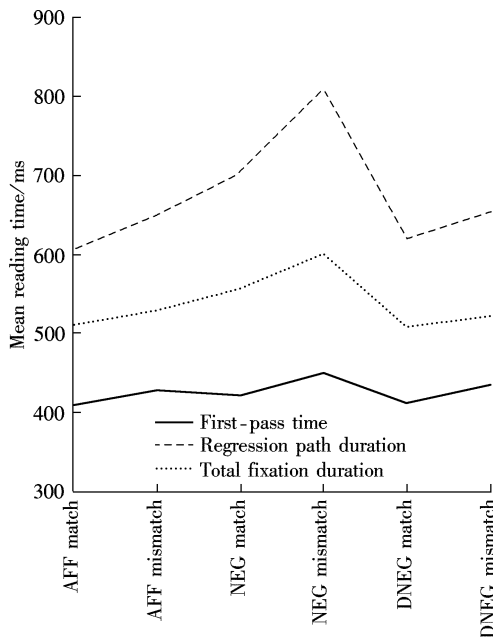


Fig. 1 Mean reading times in the spillover region

Carpenter et al.<sup>[1]</sup> suggested that verification times should be a direct function of the number of constituent comparisons. Therefore, negative sentences are more difficult to be verified than affirmative sentences. As a special negation, double negative sentence can express positive information by using two negative adverbs. According to the CCM, it is the most difficult to confirm a double negative sentence of the three kinds of sentences. In our study, there are no significant differences between double negative and affirmative adverbial phrases on the total fixation duration as well. Interestingly, there is no significant difference on the regression path duration.

Does this reflect the effect of a special strategy? We argue that there is an automatic syntactic strategy that can be activated by some lexical information. Many people know and use the syntactic rule that double negation is equal to affirmation, as they are usually told that if two negative adverbs are put together in a simple sentence, you can regard the sentence as an affirmative sentence. Therefore, when they comprehend the double negative sentence, they will need some time to activate and use the rule that the two negative adverbs should be canceled out and be processed as an affirmative syntax. Results show that the difference between the double negative and the affirmative adverbial phrases is significant in the first-pass time, but there are no significant differences in the regression path duration and the total fixation duration. We also find that there are no differences between double negation and affirmation in the spillover region. It signifies that the double negative adverbial phrase can be processed as affirmation in late processing. From these findings, one can tentatively conclude that there is a separation between the initial processing and the late processing. Double negative sentences should not be represented as the abstract proposition like {NEG, [NEG, (RED, DOTS)]} on the basis of the information from bottom to top, and we think that the mental representation of double negative sentence might be the same as an affirmative sentence.

Secondly, the present study demonstrates that negation is more difficult to be comprehended and verified than double negation in the later processing. Caplan et al.<sup>[14]</sup> found that features of the syntactic structure influenced the process of sentence-picture matching. The double negation effect disappeared on the regression path duration and the total fixation duration in the adverbial phrase region. Therefore, it is more difficult for participants to verify the negative picture-sentence than the double negative picture-sentence in the spillover region. The differences are significant on the regression path duration and the total fixation duration.

Finally, the data of eye movements supported the mismatch effect in Chinese. Knoeferle et al.<sup>[15]</sup> found that inspection times for individual sentence regions revealed a mismatch effect at the very sentence constituent for which the corresponding picture constituent mismatches. In the present study, it is obvious that there is a mismatch effect in the Chinese sentence comprehension task. Semantic mismatch invariably increases readers' cognitive load. However, Underwood et al.<sup>[8]</sup> failed to find the mismatch effect when the sentence and picture were presented one after another. It is possible that the total sentence time measured by Underwood et al. failed to disclose the mismatch effect.

In summary, there must be some other factors not considered in the study. For example, that what kind of mental representation can be used to verify the negative sentence is still a question. Some results show that the principles of mental representation used in the CCM do not apply to Chinese. Although the double negative sentences contain two negative adverbs, it is easier to be verified than the sentence in which there is only one negative adverb. We suggest that readers can use the syntactic strategy to comprehend the double negative sentences. How will people process the sen-

tences that contain three negative words? Secondly, we need to investigate the influence of other factors on sentence-picture integration in the comprehension of spoken Chinese sentences.

## 4 Conclusion

The present study provides strong evidence that the mismatch effect is also displayed under the condition that the pictures of natural scenes appear first. Nevertheless, double negation, as a special negation, is easier to be verified than negation. This may be attributable to the fact that a special strategy, namely the double negative verbs cancel each other out to construct mental representation in sentence comprehension, is adopted by readers. Therefore, we think the constituent comparison model may not be a more general model of sentence-picture tasks in other languages.

## References

- [1] Carpenter P A, Just M A. Sentence comprehension: a psycholinguistic processing model of verification [J]. *Psychological Review*, 1975, **82**(1): 45–73.
- [2] Clark H H, Chase W G. On the processing of comparing sentences against pictures [J]. *Cognitive Psychology*, 1972, **3**(3): 472–517.
- [3] Reichle E D, Carpenter P A, Just M A. The neural bases of strategy and skill in sentence-picture verification [J]. *Cognitive Psychology*, 2000, **40**(4): 261–295.
- [4] Kaup B, Zwaan R. Effects of negation and situational presence on the accessibility of text information [J]. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 2003, **29**(3): 439–446.
- [5] Schroyens W, Verschueren N, Schaeken W, et al. Conditional reasoning with negations: implicit and explicit affirmation or denial and the role of contrast classes [J]. *Thinking and Reasoning*, 2000, **6**(3): 221–251.
- [6] Tramow D, Finlay K A. The prediction of attitudes from be-

- liefs and evaluations: the logic of the double negative [J]. *British Journal of Social Psychology*, 2002, **41**(1): 77–86.
- [7] Underwood G, Jebb L, Roberts K. Inspecting picture for information to verify a sentence: eye movements in general coding and in focused search [J]. *Quarterly Journal of Experimental Psychology*, 2004, **57**(1): 165–182.
- [8] Underwood G, Crundall D, Hodson K. Confirming statements about pictures of natural scenes: evidence of the processing of gist from eye movements [J]. *Perception*, 2005, **34**(9): 1069–1082.
- [9] Knoeferle P, Crocker M W, Pickering M, et al. The influence of the immediate visual context on incremental thematic role-assignment: evidence from eye-movements in depicted events [J]. *Cognition*, 2005, **95**(1): 95–127.
- [10] Chen Q R, Deng Z, Tan D L. Sentence-picture integration in Chinese: evidence from the measurement of eye movement [J]. *Acta Psychologica Sinica*, 2008, **40**(5): 543–551.
- [11] Traxler M, McElree B, Williams R S, et al. Context effects in coercion: evidence from eye movements [J]. *Journal of Memory and Language*, 2005, **53**(1): 1–25.
- [12] Calvo M G. Working memory and inferences: evidence from eye fixations during reading [J]. *Memory*, 2001, **9**(4): 365–381.
- [13] Rayner K, Warren T, Juhasz B, et al. The effect of plausibility on eye movement in reading [J]. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 2004, **30**(6): 1290–1301.
- [14] Caplan D, DeDe G, Brownell H. Effects of syntactic features on sentence-picture matching in Broca's aphasics: a replay to Draai and Grodzinsky [J]. *Brain and Language*, 2006, **96**(2): 129–134.
- [15] Knoeferle P, Crocker M W. Incremental effects of mismatch during picture-sentence integration: evidence from eye-tracking [C]//*Proceedings of the 26th Annual Conference of the Cognitive Science Society*. Mahwah, NJ, USA: Lawrence Erlbaum, 2005: 1166–1172.

# 自然场景信息提取和验证加工的眼动研究

陈庆荣<sup>1</sup> 蒋志杰<sup>2</sup>

(<sup>1</sup> 南京师范大学心理系, 南京 210097)

(<sup>2</sup> 南京大学地理学院, 南京 210093)

**摘要:** 由于眼球追踪可以用来记录人们观看自然场景图片和理解信息时的实时眼球运动, 因此, 采用眼动技术考察“句子-图片”范式中, 呈现顺序和信息特征对语义失配效应的影响. 实验采用 3(句法) × 2(语义关系) 因子设计, 句法和语义关系是被试内变量. 实验发现, 语义失配会增加人们的认知负荷, 即第 1 遍通过时间、回视路径时间和总注视时间显著增加; 双重否定没有显著增加图片和信息的加工难度. 实验结果表明, 人们能够从长时记忆中提取特定的句法策略来处理不同语义关系的图片和句子. 这种句法加工策略使得读者能够将双重否定转换成肯定. 研究结果证明成分比较模型可能无法适用于其他类型的语言.

**关键词:** 自然场景; 语义失配; 双重否定; 眼球运动

**中图分类号:** TP391