ATLANTIS PRESS

Advances in Social Science, Education and Humanities Research, volume 480

Proceedings of the 2020 5th International Conference on Modern Management and Education Technology (MMET 2020)

Quantitative Differences of Visual Imagery, Vividness, and Coherence in Memory Between Native English and Mandarin Speakers

Wanxin Zhong

Guangdong Experimental High School, Guangzhou, Guangdong, 510375, China
Corresponding author's e-mail: Vivian.wang@cas-harbour.org

ABSTRACT

The encoding and retrieval of autobiographical memory rely extensively on linguistic cues. Previous researches suggested the critical relationship between memory and language. This study with three experiments seeks to compare the quantitative differences in visual imagery, vividness, and coherence between native English and Mandarin speakers. Based on the statistical results of this study, native Mandarin speakers demonstrate higher accuracy in the task of recalling visual imagery information in memory, with $P <. 01$. More data needs to be collected to draw a sound conclusion on the differences in vividness and coherence in the two language groups.

Keywords: *Visual imagery, vividness, coherence, language affects memory, native Mandarin speakers*

1. INTRODUCTION

Cognitive psychologists have hypothesized that language influences memory profoundly since the 1950s [1], albeit without empirical evidence. Whorf, a pioneer in linguistic relativity, believed that thought and actions are entirely determined by language. This is now referred to as the Whorfian view (or the Whorfian hypothesis), a stronger version of linguistic relativity (the theory that language affects cognition to some extent, as opposed to "universalists' theory", which claims that language does not impact thought).

The lack of data was a primary factor that contributed to the disregard for linguistic relativity in the decades following Whorf's proposed theory. However, after the 1980s, research studies related to how languages affect memory emerged again [2-8] providing sound evidences that endorse linguistic relativity.

Studies on how language affects memory advanced considerably in the past two decades. They unveiled that native speakers of different languages perceive time differently; speakers of left-branching languages outperform speakers of right-branching languages at recalling initial stimuli in working memory tasks and are worse than right-branching languages speakers at recalling final stimuli [2]. Though the theory of linguistic relativity has regained its popularity, the relation between language and memory and the underlying mechanism by which language influences memory is still vaguely understood. Furthermore, most studies on language and memory explore the role which language plays in working memory tasks. Experiments directly related to how language affects various aspects of memory, such as vividness, visual imagery, and coherence, are scarce.

This study aims to shed light upon how language affects memory in the aspects of visual imagery, vividness, and coherence [9] by comparing statistical results of three experiments that test the three categories separately in Mandarin and English native speakers.

2. STRUCTURAL DIFFERENCE OF MANDARIN CHINESE AND ENGLISH AND HYPOTHESIS

According to the World Atlas of Language Structures [10], Mandarin and English differ from each other mainly in the following aspects:

2.1. Differences of written English and Mandarin Chinese and relevant hypothesis

Written Mandarin derived from prehistoric symbols resembling natural figure, while the appearance of English words offers no direct implication to the objects' original form and shape. Therefore, this study hypothesizes that native Mandarin speakers perform better than native English speakers in visual imagery (indicated by the accuracy in Experiment 1), and that Mandarin-English bilinguals show higher level of memory vividness if they memorize images with written Mandarin, compared with written English (indicated by the vividness scale in Experiment 2).

2.2. Grammatical differences and relevant hypothesis

English has a stricter use of chronological grammar, including past tense (past perfect, simple past, past continuous, future past), present tense (simple present, present continuous, present perfect), and future tense (simple future, future continuing). Mandarin does not have strict grammatical norms for tenses. In Mandarin, numerous different words can be used interchangeably to indicate the same tense, and sometimes one phrase could mean several different tenses [10]. Judging from the relatively low awareness of tenses generated by speaking Mandarin as compared to speaking English, this study hypothesizes that native English speakers are more skillful at the corresponding experiment for coherence (shown by the accuracy in Experiment 3).

3. EXPERIMENTS

3.1. Visual Imagery

3.1.1. Participants

Thirty-nine native English speakers and thirty native Mandarin speakers participated in this experiment. All native English speakers were recruited from Amazon Mechanical Turk. Due to the lack of native Mandarin speakers on Amazon Mechanical Turk, all native Mandarin participants were recruited via online platforms in China and with the assistance of graduates from local universities (South China Normal University and Anhui Medical University). Participants of both groups are aged between 22 to 32 years old. The mean age of the native English speakers groups is 28.05 years old (Standard Deviation, SD≈2.39); the mean age of the native Mandarin speakers groups is 24.50 years old (SD≈2.20). All participants were rewarded USD 0.8$ or its equivalent in Chinese currency.

3.1.2. Procedure

Participants completed an online experiment with their own devices. The experiments were posted and conducted on pavlovia.org. The experiment lasts for approximately 3 minutes. Repeated submissions were eliminated. In the experiment, 9 image stimuli, sized 800px × 800px, appear one by one on the screen for 5 seconds each. The stimuli are random pair-ups of shape and color from six different shapes (triangle, rhombus, circle, cross, star, and octagon) and six different colors (blue, purple, red, yellow, green, and grey). The 9 stimuli for each participants are the same. An attention task that requires participants to press ‘m’, ‘e’, and ‘k’ on the key board is placed before the instructions page. Response stage: after a short interval of 30 seconds, based on previous findings that suggested short-term memory is retained for 30 seconds in a person’s mind at most [11]. Next, nine questions appear on the screen, asking which item was displayed on the screen previously. Each questions has four items as options, but only one appeared on the screen. For example, this item was shown previously like fig. 1 and the question is like: Please click on the item in fig. 2 that was on the screen previously:

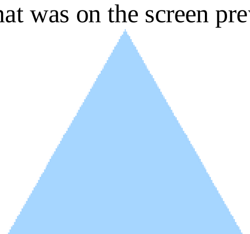

Figure 1. The original picture

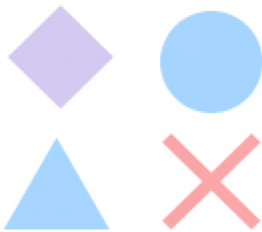

Figure 2. The picture in the question

No limited maximum reaction time was set for each question. Each participant’s number of correct responses and reaction time (in milliseconds) for each question are recorded. The accuracy ranges from 0 to 1.00. All the possible values of accuracy for a single participant are: 0, 0.11, 0.22, 0.33, 0.44, 0.56, 0.67, 0.89, 1.00 (the number of correct responses divided by nine and rounded to the hundredth decimal place).
The design of this particular experiment was adapted from a 1973 paper by Marks [12]. In his experiment, he displayed 15 objects (colored photographs of unrelated objects or complete scenes) and then asked the participants to answer which numbers or objects appeared at a specific position on the photograph. The mean numbers of correct responses for each experimental group were recorded. Experiment 1 used the same approach of testing how accurate participants’ memories about the previous stimuli are, but to limit potential confounding variables, only the color and the shapes of the stimuli needed to be remembered.

3.1.3. Results and Discussion

As previously hypothesized, the native Mandarin speaker group outperformed the native English speaker group in this visual imagery experiment (see Figure 3.). The mean accuracy of the thirty-nine native English speakers, M (Mean) ≈ .48, SD ≈ .24. For the native Mandarin speakers, M ≈ .79, SD ≈ .24.

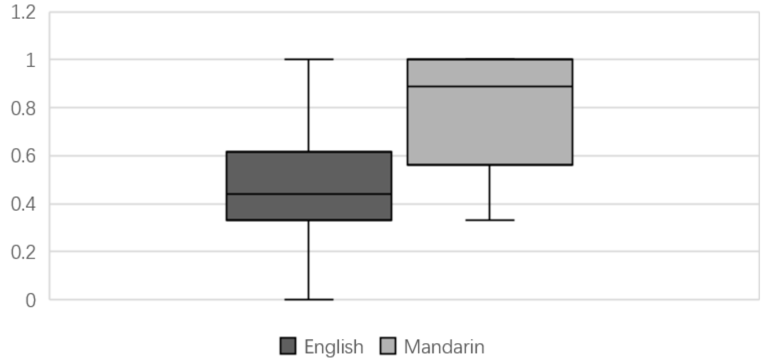

Figure 3. Visual Imagery Accuracy Between Two Native Language Groups

Table 1. One-Way ANOVA of Experiment 1

		Sum of Squares	df	Mean Square	F	*P*
Accuracy	Between Groups	1.538	1	1.538	26.510	.000
	Within Groups	3.887	67	058		
	Total	5.425	68			
Average reaction time (ms)	Between Groups	1989.094	1	1989.094	.000	.992
	Within Groups	1190633536.123	67	17770649.793		
	Total	1190635525.217	68			

The average accuracy and reaction time were analyzed with one-way ANOVA, and were run using IBM SPSS® Statistics (version 27.0). The independent variable (Factor) is native language (Mandarin or English), and the dependent variable is average accuracy and reaction time of the two groups. As shown by the figure above (FIG.2), native language has a significant influence on the accuracy of the two language groups, F (1, 69) = 26.510, $P < 0.01$. The two groups did not show any significant difference in reaction time ($P \approx 0.99$). After the removal of an outlier in the native English speaker group, whose reaction time was 26283 milliseconds (the mean reaction time with this outlier excluded was only 4433 ms), the statistical results of accuracy remains relatively unchanged, F (1, 68) = 27.817, MS (Mean Square) = 1.60, $P < 0.01$. The only notable change was *P* ($P \approx 0.49$).

Table 2. Demographic Information of Participants in Experiment 1

Native Language	**Number of Participant -s**	**Sex**	**Age*: Mean (Range)**	**Country**	**Highest Education Level****	**Second Language: Name**	**Second Language: Length of Acquisition*****	**Second Language: Self-rated Proficiency******
English	39	N(Female)=10	28.05 (24-32)	N(United States)=23	M≈5 Range: 4-6			
				N(India)=15		Tamil (N=1)	6	6
		N(Male)=29				Telugu (N=1)	12	10
				N(Britain)=1				
Mandarin	30	N(Female)=15	24.50 (22-32)	Mainland China	M≈5 Range: 5-6	English (N=17)	M≈11 SD≈5	M≈5.60 SD≈1.80
		N(Male)=15						

*In years.
**Number indicates the education degree: 1-elementary school, 2-junior high school, 3-senior high school, 4-junior college, 5-bachelor's degree, 6-master's degree, 7-doctorate.
***In years.
****On a scale of 1 (least proficient) to 10 (most proficient).

Table.1 shows the participants' basic demographic information collected after the experiment. The participants in the native English speaker group and the native Mandarin speaker group were similar in their mean age, age range, and education level. The main differences between the two experimental groups are the diversity of countries and the reported second language acquisition. Only two of the native English speakers from India reported to acquire second languages (Telugu and Tamil, both of which are among the 22 official languages of India). In the native Mandarin speaker group, more than half of the participants reported to have acquired a second language (all of whom claimed English as their second language). It is possible that some participants declined to report any compulsory second language learned in schools (Presumably, all of the native Mandarin speakers in this study learned English as it is a mandatory subject in Mainland China's educational system).

3.2. Vividness

3.2.1. Participants

Thirty-eight native Mandarin speakers participated in this experiment. All participants were recruited through online platforms in China and with the assistance of graduates from local universities. Participants were rewarded CNY15￥. The mean age of participants is 22.58 (SD ≈ 2.02).

3.2.2. Procedure

Image stimuli appear on the computer screen in a sequence. Each image remains on the screen for 5 seconds. Then, participants observe a short sentence that describes the previous image for 5 seconds. All of the sentences are in the form of "(numeral) + adjective +noun". This stage imitates the linguistic recollecting process. Next, the participants are instructed to close their eyes and recollect the image for 5 seconds and select a picture that most resembles what they had in their minds. The five options are the original image in opacity levels—15%, 30%, 45%, 60%, and 75%.

The descriptions are in English and Mandarin for different trials. Each participant will go through 2 rounds of tasks, with 20 same images in each round. All the images appear in the same order for both rounds to control possible variance in memory rehearsal, and the descriptions appear in two languages alternatively in both rounds For example, first round: image 1-Mandarin, image 2-English, ... image 20-English; second round: image 1-English, image 2-Mandarin, ... image 20- Mandarin.

For example, the original image is like follows:

Figure 4. The original image

Linguistic cue: "a yellow flower" (or its Mandarin translation in simplified Chinese characters)

Please click on the image that most resembles what you had in your mind when you tried to memorize the original image.

Figure 5. The options of the question

The percentage of choosing each of the opacity levels is collected. After the experiment, participants are required to fill out their basic demographic information (the same information as in Experiment 1).

The stage where a short sentence is displayed after the image stimulus was designed to imitate the process of encoding memory through a specific language. After viewing the image, participants recollect the image with the linguistic cues on the screen. This ensures that participants remember the image stimuli in Mandarin for certain trials and in English for other trials. The aforementioned sequence of the images was intended to eliminate unnecessary interference in selected opacity scale due to different time intervals and number of image stimuli that appear between an image's first and second appearance. All images were shown twice (first with English linguistic cue, then with Mandarin linguistic cue) to rule out variance in vividness responses caused by the objects or scenes in the images themselves.

Vividness is defined as the clearness, color (luminance and saturation), and the details of memory. Clearness of an image can be demonstrated by the opacity of images [13].

3.2.3. Results and Discussion

For the image option with 15%, 30%, 45%, 60%, and 75% opacity, the respective data recorded were 1.00, 2.00, 3.00, 4.00, and 5.00. The overall opacity data for the responses following English linguistic cues have a mean of 4.12, SD = .67, range: [2.65, 5.00]. For the responses after Mandarin descriptions of the stimuli, M = 4.16, SD = .66, range: [2.80, 5.00]. However, the language of the linguistic cue did not have significant influence on the opacity scale, with P= .76, t = -.23 in an independent T test analysis where the language of the descriptive sentences were represented with 1 for English, and 2 for Mandarin. When the opacity levels for the first round of images and the second round of the same images (showing up for the second time) were analyzed separately, the results were not significant either, with
P_1 = .49, t_1 = .37; P_2 = .87, t_2 = -.16.

Some adjustments can be made to revise this experiment. Demonstrating vividness solely through the transparency of images might be too limited. Other studies manipulated the color luminance and saturation [14] when testing vividness. Based on the definition of memory vividness, perhaps a systematic combination of trials with options that differ in opacity (or transparency), color saturation and luminance, and image resolution would capture a fuller image of vividness. Additionally, the display time of image stimuli can be shortened. In this way, the images most likely would not have been encoded but only within the participants' perception. This change can serve to make certain that the subjects recollect the image using the designated language while observing the linguistic cue on the screen.

3.3. Coherence

3.3.1. Participants

Forty-three native English speakers and forty-one native Mandarin speakers. All native English speakers were recruited from Amazon Mechanical Turk. Due to the lack of native Mandarin speakers on Amazon Mechanical Turk, all native Mandarin participants were recruited via online platforms in China and with the assistance of graduates from local universities. All participants were rewarded USD 0.8$ or its equivalent in CNY. For the age of the native Mandarin speaker group, M = 22.39, SD ≈ 1.90. For the age of the native English speaker group, M = 26.04, SD ≈ 3.70. The age range for both groups of participants is 20 to 34 years old.

3.3.2. Procedure

An attention task that requires participants to press 's', 'u', and 'v' on the key board in order to proceed is place before the instructions page. After the experiment begins, 12 items, images sized 800px × 800px, with 6 different shapes and 6 colors (paired randomly as in experiment 1) appear in an order on the screen. Some of the items' appearance time overlap that of another item. For example, the yellow star appears on the screen from 7-12 seconds after the start of the video, and the gray cross appears on the screen from 10-15 seconds after the start). The participant takes a brief interval of 30 seconds, based on [11]. Each participant makes responses on the computer to indicate the sequence of the items, and whether a certain action overlapped the other. The options for each question are provided with pictures of the shapes in various colors. The reaction time is not limited for each question.

There are four types of questions in total (Figure 6. 7. 8. 9.).

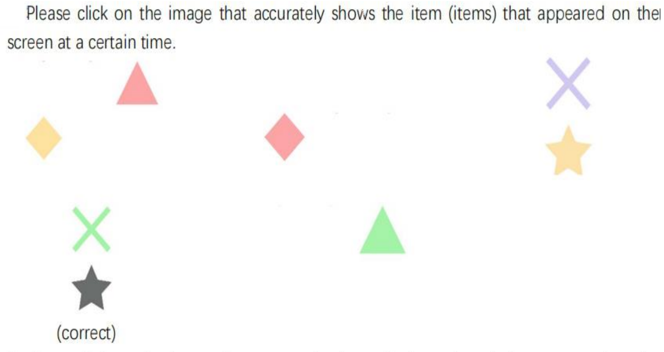

Figure 6. Questions Type 1

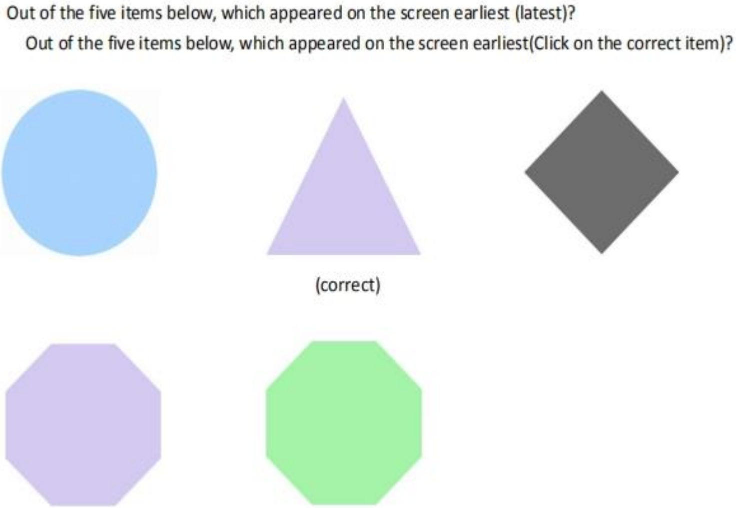

Figure 7. Questions Type 2

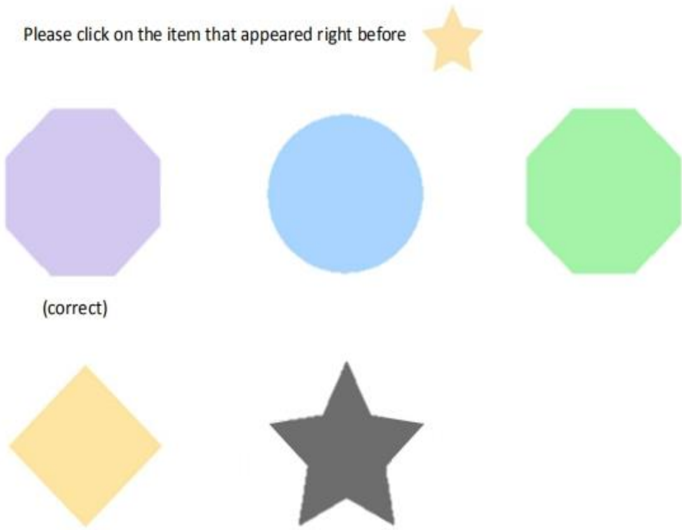

Figure 8. Questions Type 3

Figure 9. Questions Type 4

3.3.3. Results and Discussion

The two groups do not show significant difference in memory coherence, which is not consistent with the hypothesis. The mean accuracy (number of accurate responses divided by the total number of questions, 20) in the native English speaker group is .26, SD =.12. The mean accuracy of the native Mandarin speakers is .29, SD = .13. The accuracy of all questions in the native English and Mandarin speaker groups were analyzed using an independent-samples T Test, *P* = .23, t = -1.26 (as in Experiment 2, 1 represents native language English, and 2 represents native language Mandarin).

The insignificance is most likely caused by the sampling differences between the native English and Mandarin speaker groups. Specifically, the highest education levels in the two groups are dissimilar. In the native Mandarin speaker group, 85.4% (n = 35) of the participants have bachelor's degree, and 14.6% (n = 6) attain master's degree. However, among the native English participants (recruited through Amazon Mechanical Turk with the criteria of between 18-34 years of age), 34.9% (n = 15) have bachelor's degree, 18.6% (n = 8) have a highest degree of senior high school, 16.3% (n = 7) have a junior college degree, 16.3% (n = 7) are with a junior high school degree, and 14.0% (n = 6) are with an elementary school degree. Another reason that may have contributed to the unexpected result is that this experiment mainly uses visual stimuli to test memory coherence (question type 1 in Experiment 3 is especially similar to the method in Experiment 1), and as shown in Experiment 1, native Mandarin speakers are more skillful at remembering visual images. An experiment using auditory stimuli or physical actions in a sequence might be able to rule out the interference of participants' visual imagery in memory.

4. CONCLUSION

Judging from the results of the three experiments in this study, native English and Mandarin speakers differ significantly in the visual imagery aspect of memory. Native Mandarin speakers outperform native English speakers of the same age group and similar demographic background in the accuracy of memory imagery. Mandarin uses hieroglyphic writing system while English uses phonetic writing system. That is, Mandarin characters originated from objects' physical appearances, and modern hieroglyphic writings, like Mandarin Chinese, maintain their pictorial forms. The demanding skills needed to remember the meanings for over four thousand characters (the number of characters recognized by an average native Mandarin speaker) most likely led to the significant advantage in visual imagery for native Mandarin speakers. In Experiment 2, Mandarin-English bilinguals report higher vividness level (indicated by the opacity scale) after observing linguistic cues, which describes the previous image stimuli, in Mandarin compared to those in English, but the results were not significant. In Experiment 3, native Mandarin speakers perform slightly better than native English speakers, contrary to the hypothesis. Experiment 2 and 3 might yield significant results in support of the hypothesis of this study if the experimental procedures of Experiment 2 were refined to measure memory vividness more precisely, and the sampling differences could be better controlled in Experiment 3.

Overall, this study shows that native languages have notable influences on certain aspects of memory, such as the visual imagery. Future studies comparing more groups of languages, and their impacts on other areas of the memory can serve to draw a more comprehensive picture of how different languages affect human memory. Ultimately, these findings can lead to a deeper understanding of how people view their inner world and outer world constructed by memory differently due to their languages.

ACKNOWLEDGMENT

I would like to thank Yuanyuan Dong for her support in the thesis selecting process and her invaluable information on data analyzing. I would also like to thank Yibin Chen for her assistance in participant recruitment.

REFERENCES

[1] Whorf, B. L. (1956) In J. B. Carroll (Ed.), Language, thought and reality: Selected writings of Benjamin Lee Whorf. Cambridge, MA: MIT Press.

[2] Amici, F., Sánchez-Amaro, A., Sebastián-Enesco, C., Cacchione, T., Allritz, M., Salazar-Bonet, J., & Rossano, F. (2019). The word order of languages predicts native speakers' working memory. Scientific Reports, 9(1). https://doi.org/10.1038/s41598-018-37654-9

[3] Boroditsky, L. (2001). Does language shape thought? Mandarin and English speakers' conceptions of time. Cognitive Psychology, 43(1), 1–22. https://doi.org/10.1006/cogp.2001.0748

[4] Chen, Z.-Y., Cowell, P. E., Varley, R., & Wang, Y.-C. (2009). A cross-language study of verbal and visuospatial working memory span. Journal of Clinical and Experimental Neuropsychology, 31(4), 385–391. https://doi.org/10.1080/13803390802195195

[5] Engemann, H., Hendriks, H., Hickmann, M., Soroli, E., & Vincent, C. (2015). How language impacts memory of motion events in English and French. Cognitive Processing, 16(S1), 209–213. https://doi.org/10.1007/s10339-015-0696-7

[6] Francis, W. S. (1999). Cognitive integration of language and memory in bilinguals: Semantic representation. Psychological Bulletin, 125(2), 193–222. https://doi.org/10.1037/0033-2909.125.2.193

[7] MacKay, D., James, L. E., & Abrams, L. (2002). 4 Cross-language facilitation, repetition blindness, and the relation between language and memory: Replications of Altarriba and Soltano (1996) and support for a new theory. Bilingual Sentence Processing, 134, 89–109. https://doi.org/10.1016/s0166-4115(02)80007-3

[8] Marian, V., & Fausey, C. M. (2006). Language-dependent memory in bilingual learning. Applied Cognitive Psychology, 20(8), 1025–1047. https://doi.org/10.1002/acp.1242

[9] Berntsen, D., Hoyle, R. H., & Rubin, D. C. (2019). The Autobiographical Recollection Test (ART): A Measure of Individual Differences in Autobiographical Memory. Journal of Applied Research in Memory and Cognition, 8(3), 305–318. https://doi.org/10.1016/j.jarmac.2019.06.005

[10] Dryer, Matthew S. & Haspelmath, Martin (eds.) 2013. The World Atlas of Language Structures Online.Leipzig: Max Planck Institute for Evolutionary Anthropology. http://wals.info

[11] D., C. P., Ellis, H. C., & Hunt, R. R. (1983). Fundamentals of Human Memory and Cognition. The American Journal of Psychology, 96(4), 579. https://doi.org/10.2307/1422579

[12] Marks, D. F. (1973). Visual imagery differences in the recall of pictures. British Journal of Psychology, 64(1), 17–24. https://doi.org/10.1111/j.2044-8295.1973.tb01322.x

[13] Vandeberg, L., Eerland, A., & Zwaan, R. A. (2012). Out of Mind, Out of Sight: Language Affects Perceptual Vividness in Memory. PLoS ONE, 7(4), e36154. https://doi.org/10.1371/journal.pone.0036154

[14] Cooper, R. A., Kensinger, E. A., & Ritchey, M. (2019). Memories Fade: The Relationship Between Memory Vividness and Remembered Visual Salience. Psychological Science, 30(5), 657–668. https://doi.org/10.1177/0956797619836093

[15] Mata, A., Percy, E. J., & Sherman, S. J. (2013). Adjective–noun order as representational structure: Native-language grammar influences perception of similarity and recognition memory. Psychonomic Bulletin & Review, 21(1), 193–197. https://doi.org/10.3758/s13423-013-0479-z

[16] Matsumoto, A., & Stanny, C. (2006). Language-dependent access to autobiographical memory in Japanese-English bilinguals and US monolinguals. Memory, 14(3), 378–390. https://doi.org/10.1080/09658210500365763