Influences of Language Shift on Speech Fluency in Memory Production of Unbalanced Chinese-English Bilinguals

Lulu Wang
School of English and International Studies, Beijing Foreign Studies University, Beijing, China

Abstract—Language shift at encoding and retrieval has been demonstrated to influence memory. However, previous studies of language shift mostly focused on memory accuracy, leaving its effect on fluency of memory production neglected. The current study sets out to investigate how encoding language and delay interval influence speech fluency in memory production. In the study, 60 Chinese ESL learners were asked to verbally retell two stories in immediate or delayed memory tests. One story was learned in English but recalled in Chinese, and the other was learned and recalled both in Chinese. The fluency of their speech reproduction during story recall was manifested by production fluency and pause rate. The results show that encoding in English and retrieving in Chinese leads to more pauses and lower fluency than encoding and retrieving both in Chinese. However, the language shift effect on speech production diminishes after 24 hours’ delay. The results suggest the key role of delay interval in modulating language shift effect and shed light on bilingual memory and bilingual education.

Index Terms—language shift, story recall, speech fluency, delay interval

I. INTRODUCTION

Immersion education programs are becoming increasingly popular around the world as an efficient approach to teaching foreign languages (Johnson & Swain, 1997). Following the trend, in recent years China has gradually developed English-as-a-medium-of-instruction (EMI) courses to improve students’ English proficiency (Xiong & Feng, 2018; Zhang, 2018). However, although the academic knowledge and concepts in these EMI courses are taught in English, they are usually communicated and put into practice in Chinese.

Would the mismatch between input and output languages affect learning and memory recall? This has been widely debated but remains controversial (Antón, Thierry, Goborov, Anasagasti, & Duñabeitia, 2016). Finding an answer to this question is of critical importance to the understanding of bilingual memory and will provide profound insights into bilingual education. In addition, previous studies have probed into the influence of language shift on memory accuracy and yielded a great number of revealing results (Matsumoto & Stanny, 2006; Marian & Fausey, 2006; Goldstein-Diament & Vakil, 2019). However, speech fluency in memory recollection, as another important indicator for memory retention (Gomez & White, 2006), has drawn a far less amount of attention in the field. Therefore, the current study will focus on the issue of language shift and examine how language shift influences speech fluency in memory recollection and how such effect varies with time.

A. Language Shift Effect on Memory Representation

Language shift occurs when people encode an episode or a text in one language but retrieve the encoded information in another language (Marian & Neisser, 2000). It has been found that language shift would significantly decrease people’s memory performance, which was defined as language shift effect (Marian & Neisser, 2000). A great number of researchers have found evidence in support of language shift effect on memory (Marian & Neisser, 2000; Matsumoto & Stanny, 2006; Marian & Fausey, 2006). For example, Marian and Neisser (2000) examined language shift effect among Russian-English bilinguals. It was shown that their episodic memory was more accessible when the encoding and retrieval languages were matched than when they were mismatched, demonstrating a language-dependent phenomenon. It was explained that language shift at encoding and retrieval would lead to larger processing costs in retrieving past episodes. Similar findings are reported in Matsumoto and Stanny’s study (2006) of Japanese-English bilinguals, that the episodic memory being retrieved at the test was mostly encoded in the same language as the cue words. These studies have provided evidence for language-dependent memory and language shift effect on autobiographic memory retrieval.

However, all of these studies only focused on language shift effect on episodic memory, which is grounded in the experiences of past real-life events. In this regard, Marian and Fausey (2006) stood out by exploring how language shift influences academic learning. In their study, they took use of four short academic-type stories and tested the memory for the stories among balanced and unbalanced Spanish-English bilinguals. Critically, the consistency of the encoding language and retrieval language was manipulated. In line with previous findings, it was shown that memory retrieval of
balanced bilinguals was faster and more accurate when the input language and output language were matched (Marian & Fausey, 2006). Yet such language shift effect was absent for unbalanced bilinguals, which contradicted the fact that the unbalanced bilinguals generally experience larger processing costs during language shift or translation compared with balanced bilinguals, due to lower proficiency in L2. Hence, it is necessary to carry out an experiment to further testify the language shift effect among unbalanced bilinguals.

In addition, it should be noted that previous research of language shift effect on text memory only discussed academic-type stories, while there is little attention to stories of daily-life events (Marian & Fausey, 2006). However, mental representations of the two types of stories should be different (Wolfe & Mienko, 2007). Firstly, processing academic-type stories should rely more on the text-level input or the exact words used in the stories, as the stories usually refer to unfamiliar and novel concepts or abstract scientific knowledge, which can hardly be linked to previous life experiences. However, the representations of daily-life events rely less on the text-level information, as their understanding and interpretation can be constructed based on prior knowledge or experiences. Secondly, academic-type stories usually include certain unfamiliar words or terms, while the vocabulary used in describing daily events is usually simple and frequently used. Therefore, larger processing costs and text-based memory should possibly be associated with memory for academic-type stories, while lower processing costs and bigger influence from previous experiences would be expected for memories of daily-life stories. It leaves us wondering whether different processing mechanisms will influence language shift effect on text memory, which will be the major concern of the current study.

B. The Influence of Delay Interval on Language Shift Effect

Memory is not the exact replica of the original experience and always changes over time (Bartlett, 1932). The initial memory for a story usually retained highly detailed information. However, over time, only the story’s gist and few details could be reserved (Moscovitch, Cabeza, Winocur, & Nadel, 2016; Bartlett, 1932). Then will time and memory decay influence language shift effect on memory recollection? There is only a very limited amount of studies dedicated to this issue.

Gablasova’s (2014) study investigated how delay modulated the effect of encoding languages on memory production. To be specific, it was found that the participants’ memory for academic texts displayed in L2 decayed more noticeably and faster than L1-encoded texts (Gablasova, 2014). However, the findings relating to this issue were inconsistent. In contrast to Gablasova’s (2014) study, Beken, Woumans, and Brysbaert (2018) reported that L2 input does not result in a weaker performance of memory production than L1 input both at the immediate test and at delayed tests (Beken et al., 2018). These contrasting findings suggested that more research needs to be carried out to ascertain whether learning in L2 hinders memory production. In addition, these studies only focused on the condition where no language shift was experienced during memory recall. The present study is going to examine how language shift influences language production fluency immediately and 24 hours after learning. Then, how would language shift effect change over time? Therefore, the second aim of the present study is to examine how language shift effect on memory recollection changes with time.

The answer to this question may lie in the different forgetting rates of different levels of memory representations for narratives. It has been revealed that information is usually encoded and stored in three structures, namely surface structure (the word-by-word text), propositional structure (abstract meanings), and situational structure (story gist) (Schmalhofer & Glavanov, 1986). In addition, it was found that memory for surface structure and propositional structure would quickly fade with time, whereas memory for the situational structure, i.e. the gist of the story, would still be accessible even in four days after learning (Kintsch, Welsh, Schmalhofer & Zimny, 1990).

When it comes to the current study, the Chinese version and the English version of the stories differ in surface structure, as different texts were used in the description, but they were identical in situational structure (story gist). Therefore, when tested immediately after learning, the participants’ memory for the surface structure or word-by-word text will be fresh. Accordingly, retelling an English-encoded story in Chinese would require translation of the original verbatim memory into another language, which will lead to larger processing costs and language shift effect in memory recollection. However, after 24 hours, the surface structure will decay but the situational structure will still retain. At this moment, retelling an English story would be similar to retelling a Chinese story, as memory production was based on situational structure, which is identical between the Chinese story and the English one. Therefore, it is predicted that language shift should influence immediate memory recall, but will make little difference to memory with a longer delay.

C. Fluency of Language Production and Memory

Most previous studies on memory have only concentrated on the amount of produced memory, memory accuracy, recognition speed, and false memory caused by the shift of language at encoding and retrieval (Marian & Neisser, 2000; Matsumoto & Stanny, 2006; Marian & Fausey, 2006; Goldstein-Diament & Vakil, 2019), but speech fluency in memory recall has seldom been explored.

However, language shift may also influence speech fluency in memory recollection. For one thing, language shift may decrease speech fluency during memory recall (Matsumoto & Stanny, 2006; Marian & Fausey, 2006), as further translation was needed in the recall process, which led to enlarged processing costs. For another thing, language shift may also influence speech fluency by lower memory quality at the encoding stage (Hogan et al., 2006; Alexiadou et al.,
which resulted in larger processing difficulty and decreased speech fluency in memory recollection (Gomez & White, 2006). Therefore, the current study resorts to production fluency as an indicator for memory performance and processing speed to examine how speech fluency in memory production for stories differs with language shift.

In general, this study will focus on how language shift influences speech fluency in memory recall among unbalanced Chinese–English bilinguals and discuss the role of delay interval in modulating Chinese–English bilinguals. To do that, we will use two short stories about daily-life events as experimental materials, each having a Chinese version and an English version. Participants will take the story recall test immediately after learning or after one day’s delay. Their speech fluency in story recall was evaluated through the analysis of pause rate and production fluency.

It is predicted that participants who experience the shift of input language and output language would be more fluent in immediate memory recall. However, this language shift effect would become barely noticeable in 24 hours. Besides, we expect that on average Chinese-encoded stories would be reproduced more fluently than English-encoded stories both at the immediate test and the 24-hour-delayed test.

II. Method

A. Participants

We recruited 60 Chinese college students online and randomly assigned half of them to the immediate test group and the other half to the delayed test group. Among them, 16 were males and 44 were females. The participants’ mean age was 21.4. In addition, two participants were removed from data analysis because of poor performance in memory recollection task.

They were all Chinese native speakers who had started learning English from primary school and had kept on learning it in school since then. Information about the participants’ English proficiency was collected, including results of CET-4, CET-6 (Band 4 and Band 6 of College English Test in China, mandatory for Chinese college students), IELTS, TOFEL and other certificates that could serve to identify their English proficiency. We also tested their vocabulary on Lextale, an online platform, with a scale of 100 points. It was shown that the vocabulary level of participants was at an intermediate or advanced level (mean=73.38, SE=9.03). A combined analysis of their vocabulary and other test results revealed no significant difference in English proficiency between the immediate test group and delayed test group (p>0.05). In addition, it should also be noted that none of the participants had the habit of stuttering so that their pauses in the free recall test excluded the possibility of speech impediment.

B. Materials

We created two narrative stories as reading materials and each story had two versions, a Chinese version and an English version. In order to guarantee equivalent semantic meanings and syntactic structure between the two versions of stories, translation was done by three translators independently before they negotiated for the final version of translation. Moreover, polysemy was excluded from the stories to avoid ambiguity and extra cognitive efforts that would come along, and all the English words we chose were in the vocabulary required by National College Entrance Examination, which were highly familiar to our participants. Besides, to minimize the difference between the two stories, they were matched from topic, length, structure, and the amount of conveyed information, to difficulty of vocabulary, grammatical structure, and familiarity with details in the stories.

We also designed a questionnaire to test the participants’ familiarity with the objects and events in the stories. In the online questionnaire, the participants were asked to rate their familiarity with the objects and events occurring in the stories on a seven-point scale. The results showed no significant difference between their familiarity with the two stories (p>0.05). In summary, the two stories were matched in surface structure, difficulty, and familiarity so that the chance of them causing different cognitive efforts was minimum.

C. Experimental Design and Procedure

There were two lists for this experiment. Each list consisted of two stories, with only one version of each story in it. Specifically, List 1 included the Chinese version of Story 1 and the English version of Story 2. List 2 included the English version of Story 1 and the Chinese version of Story 2. An equal number of participants were randomly assigned to List 1 and List 2. The two stories in each list were presented in random order and counterbalanced.

Due to COVID-19, the experiment had to be carried out online. We took use of an online-questionnaire Qualtrics to present stories. At the same time, the experimenter communicated with the participants and gave instructions via video link, through the communication tool Wechat and recorded their voices in the memory recall task using a voice recorder.

The participants were instructed to memorize the stories while reading them aloud at their own speed. They were only given one chance to read through the two stories and the stories were presented one by one in random order. After reading the two stories, participants in the immediate test group took part in the free recall task. At this stage, they were encouraged to orally describe whatever they could remember in Chinese and they could organize their words freely based on their memories and personal preferences. Before retelling each story, a word referring to the theme of it (e.g., shopping and zoo) was shown on the screen as the retrieval cue and the participants were given 30 seconds for preparation. After that, they were asked to describe the learned stories in exquisite detail. There was no time limit for
the recall task so that they could tell the stories at their own speed. For the participants in the delayed test group, instead of being tested immediately after learning, they were asked to maintain their regular schedule and to take part in the free recall task 24 hours after learning, at exactly the same time on the next day.

After the memory test of both groups, the participants would provide some personal information, including age, gender, birth-place, places of growing up, and the aforementioned results of various English tests that the participants had taken part in that could identify their English proficiency. After that, they were required to take the online vocabulary test on Lextale.

D. Data Treatment

All of the recordings of story recall were transcribed. We first resorted to the app Xunfeitingjian to convert those speeches to texts. Then the three coders manually checked the transcription independently to guarantee the accuracy of the transcribed texts. Then they discussed the final version of the transcription. The duration of every pause during story recall was marked by referring to the audio spectrum of the recordings showed by Audacity. To analyze how the fluency of story reproduction differed with delay and language shift, the total number of pauses and the number of words in each story reproduction were coded. In addition, following Esposito, Stejskal, Smekal and Bourbakis (2007) and Kormos (2006), both filled pauses (intervals filled by vocalizations having no lexical meaning, such as “uh…”, “uhm…”, “er…”, etc.) and empty pauses (periods of silence larger than 0.2s) in speech were coded as pauses.

Production fluency and pause rate were further calculated as two indicators of speech fluency: (1) Production fluency is the number of words per second, which is calculated by dividing the number of words by the duration of production; (2) pause rate is the number of pauses per second, which is the number of pauses divided by the duration of production. To analyze how the fluency of story reproduction differed with delay and language shift, the total number of pauses and the number of words in each story reproduction were coded. In addition, following Esposito, Stejskal, Smekal and Bourbakis (2007) and Kormos (2006), both filled pauses (intervals filled by vocalizations having no lexical meaning, such as “uh…”, “uhm…”, “er…”, etc.) and empty pauses (periods of silence larger than 0.2s) in speech were coded as pauses.

Production fluency and pause rate were further calculated as two indicators of speech fluency: (1) Production fluency is the number of words per second, which is calculated by dividing the number of words by the duration of production; (2) pause rate is the number of pauses per second, which is the number of pauses divided by the duration of production.

D. Data Treatment

All of the recordings of story recall were transcribed. We first resorted to the app Xunfeitingjian to convert those speeches to texts. Then the three coders manually checked the transcription independently to guarantee the accuracy of the transcribed texts. Then they discussed the final version of the transcription. The duration of every pause during story recall was marked by referring to the audio spectrum of the recordings showed by Audacity. To analyze how the fluency of story reproduction differed with delay and language shift, the total number of pauses and the number of words in each story reproduction were coded. In addition, following Esposito, Stejskal, Smekal and Bourbakis (2007) and Kormos (2006), both filled pauses (intervals filled by vocalizations having no lexical meaning, such as “uh…”, “uhm…”, “er…”, etc.) and empty pauses (periods of silence larger than 0.2s) in speech were coded as pauses.

Production fluency and pause rate were further calculated as two indicators of speech fluency: (1) Production fluency is the number of words per second, which is calculated by dividing the number of words by the duration of production; (2) pause rate is the number of pauses per second, which is the number of pauses divided by the duration of production.

To examine how production fluency varies with language shift and delay interval, we performed repeated ANOVA with language condition (language shift VS. no language shift), delay interval (immediate test VS. delayed test) as independent variables, and production fluency as the dependent variable. The results showed that there was a significant interaction between delay interval and language condition (F(1, 55)=4.26, p=0.04, see Figure 1). In addition, with further paired-sample T-tests, we found that language shift decreased production fluency only for the immediate group (t(27)=2.88, p=0.01), but not for the delayed group (p>.05). Besides, we failed to observe the main effects of language condition and delay interval.

We also examined how language shift and delay interval influence pause rate. Similarly, we performed repeated ANOVA with pause rate as dependent variable and language condition (language shift VS. no language shift), delay interval (immediate test VS. delayed test) as independent variables. The results showed that there was a significant interaction between delay interval and language condition (F(1, 55)=4.312, p=0.04, see Figure 2). In addition, with further paired-sample T-tests, we found that language shift only increased pause rate when participants were tested immediately (t(27)=3.97, p<0.001), but not when they were tested with 24 hours’ delay (p>.05). However, we did not observe the main effect of language condition nor the main effect of delay interval.
IV. Discussion

The current study aimed to investigate how language shift influences speech fluency of story reproduction and whether the language shift effect will change with time. Consistent with our predictions, we observed the interaction between language shift condition and delay interval, that language shift significantly increased pause rate and decreased fluency, when participants were examined immediately after learning, but not when they were tested after a one-day delay. However, inconsistent with our expectations, we failed to observe the main effects of delay on both production fluency and pause rate.

This study found that language shift decreased fluency in the immediate test, which was consistent with the findings of Marian and Fausey (2006), that bilinguals’ memory for texts is more accessible when the encoding language is matched with the retrieval language. The language shift effect is possibly driven by different language processing mechanisms associated with language shift condition and non-language shift condition. More specifically, when people initially learned the stories, their memory was so vivid and detailed that they could recall the exact words used in the stories. In this condition, their memory for stories should be language-dependent. As a result, if they were asked to recollect the learned stories in a different language, they would have to translate the original text into another language, which increased processing costs and decreased speech fluency during memory recall. However, if the language used at recalling matched the encoding language, no further translation would be needed, and people only need to repeat the text from their memory efficiently. Therefore, language shift exerted influences on the results of the immediate test group.

However, the observed language shift effect was absent in the delayed test, which suggested that delay interval modulated language shift effect on production fluency, which the process of memory consolidation should possibly account for. Specifically, it was suggested that memory for stories would gradually be converted into a gist-based version over time since the surface structure could barely exist in one day, while the propositional structure and situational structure of the story were still retained in their mind at this point (Kintsch et al., 1990). That is to say, people would forget the word-by-word text of the original stories in 24 hours after learning, but could still remember the gist of the stories. In this situation, when participants were asked to retell the stories, they had to organize their words based on the story’s gist at the point of recollection, regardless of the encoding language. In this way, memory for stories was gradually transformed from a language-dependent version into a language-independent version. Therefore, the language shift effect observed at the immediate test was eliminated at the delayed test, owing to the process of memory consolidation.

Nevertheless, it should be noted that we observed the language shift effect among unbalanced bilinguals, while in Marian and Fausey’s study (2006), the language shift effect could only be observed among balanced bilinguals. The reason for this discrepancy was possibly the type of materials used in the two studies. Specifically, in Marian and Fausey’s study (2006), they took use of academic-type texts, which were designed to introduce novel scientific concepts and consisted of a series of abstract and academic terms. Consequently, participants might encounter difficulty in learning the texts, such that their memory for them would be influenced, regardless of the encoding language. However, in the current study, we took use of texts of daily events, which were composed of rather simple vocabulary and could easily be understood. Therefore, we managed to control the interference from the texts. By doing so, we had a higher chance to capture the language shift effect among unbalanced bilinguals.

Moreover, interestingly, inconsistent with our prediction, speech fluency in memory recollection did not change with delay. One possible explanation for this phenomenon is that participants just told the information they remembered, without making efforts to figure out what they had forgotten. This idea can be substantiated by the decreased details in story recall among the delayed group. In addition, group differences may also account for the absence of delay’s influence on speech fluency. As we adopted a between-subject design in this study, the delayed group and the
immediate group were two different groups of people, so that they may vary in production fluency and pause habits. As a result, delay effect on production fluency may have also been interfered with by individual differences.

V. CONCLUSION

In conclusion, this study investigated how language shift and delay interval influence speech fluency during memory recall. It revealed that language shift led to decreased speech fluency only temporarily, and this influence would wane to an insignificant level with a 24 hours’ delay. The results provided further evidence in support of language shift effect among unbalanced bilinguals and suggested the key role of delay interval in modulating language shift effect.

This study provides important insights into bilingual memory and bilingual language processing. Firstly, it is one of the first investigations into language shift effect on speech production fluency, while most previous studies only focused on memory accuracy and retrieval speed. It extends our knowledge of language shift by highlighting its effect on another potential aspect of memory representation. Secondly, it also illustrates the role of delay interval in modulating language shift effect, thus furthering our understanding of how language shift effect alters with time. Thirdly, this study is distinct from most previous studies in that texts of daily events were applied as experimental materials, so that we could eradicate the interference of unfamiliar terms and concepts.

However, the current study is also subject to several limitations. Firstly, the participants were instructed to read the experimental materials aloud to memorize them so that the stories were only encoded through reading. In this condition, only learning outcomes under the circumstance of self-teaching were evaluated, while in most cases school education is performed through verbal communication, in other words, the students usually have to listen to their teachers to learn. However, memory production of information encoded through listening was not tested in this study. Therefore, to some extent, the results of this study cannot account for memory production of knowledge encoded in other ways than reading. In addition, only an immediate recall test and a 24-hour-delayed test were carried out in the current study, but it is unknown whether memory production fluency will change with a longer delay than 24 hours.

Further research should be carried out to investigate whether language shift effect will be modulated when information is encoded not by reading but through other modalities like listening or when memory is not produced verbally but through writing. In addition, it will also be necessary to test memory production fluency with longer delay intervals to check whether language shift effect will work again after disappearing at the 24-hour-delayed test.

APPENDIX

A. Experimental Material: Story 1


English version: S1: Today is Saturday, and Lily's uncle is coming to visit at noon. S2: Lily walked to the supermarket to buy some food at eight o'clock in the morning. S3: She bought some beef, eggs and tomatoes, as well as some oranges and strawberries. S4: Besides, she also bought three bottles of beer. S5: These are all her uncle's favorites. S6: Afterwards, Lily took a taxi home. S7: After arriving at home, she firstly made the bed, then swept the floor, and finally threw away the garbage. S8: However, at this moment, Lily suddenly found that the water supply was cut off, and she could not cook anymore. S9: So she sat in the sofa angrily.

B. Experimental Material: Story 2

Chinese version: S1: 七月份的一天，安娜和朋友们计划去动物园玩。S2: 安娜早上九点开始洗澡化妆。S3: 那天，她选了一件白色T恤和一条黑色短裤以及一双运动鞋。S4: 另外还戴了顶粉色帽子。S5: 她喜欢自己的打扮。S6: 然后，她坐上公交车到动物园。S7: 和朋友们汇合后，她们先去看了大象，然后喂了猴子，最后给熊猫拍了拍照。S8: 但因为老虎馆在整修，她们没有看到老虎。S9: 所以她们伤心地回家了。

English version: S1: One day in July, Anna and her friends planned to go to the zoo. S2: Anna started taking a shower and doing her makeup at nine o’clock in the morning. S3: That day, she chose a white T-shirt, a pair of black shorts and sports shoes. S4: Besides, she also wore a pink hat. S5: She liked the way she looked. S6: And then, she got on the bus to the zoo and met her friends. S7: After meeting her friends, they firstly saw the elephants, then fed the monkeys and finally took pictures of pandas. S8: However, since the Tiger House was under repair, they did not see the tigers. S9: So they went home sadly.

ACKNOWLEDGMENTS

The author wishes to thank all the participants for their time and efforts. My sincere gratitude also goes to Dr. Yaqi Wang from Shandong University for her professional guidance and kind advice on experimental design and data treatment. Besides, Miss Mengmeng Zuo and Mr. Wenzheng Jiang’s assistance in carrying out the experiment is also deeply appreciated.
REFERENCES


Lulu Wang is currently a graduate student in the School of English and International Studies of Beijing Foreign Studies University, China. Her research interests include bilingual memory and translation studies.