科技部補助專題研究計畫報告

再探閱讀中文文句時的斷詞是否為一種基於統計學習的結果

報 告 類 別 ： 成果報告
計 畫 類 別 ： 個別型計畫
計 畫 編 號 ： MOST 108-2410-H-003-043-
執 行 期 間 ： 108年08月01日至110年01月31日
執 行 單 位 ： 國立臺灣師範大學華語文教學系

計 畫 主 持 人 ： 陳振宇

計畫參與人員： 學士級-專任助理：林佩玲
博士班研究生-兼任助理：王贊育

本研究具有政策應用參考價值：□否 ■是，建議提供機關教育部
（勾選「是」者，請列舉建議可提供施政參考之業務主管機關）
本研究具影響公共利益之重大發現：■否 □是

中 華 民 國 110 年 01 月 27 日

中文摘要：本研究採用Fiser and Aslin的實驗派典繼續探討受試者處理一串視覺語言符號時，能否從中擷取隱含的統計資訊，實驗的結果顯示，初學的華二語受試者在處理一串中文字的時候展現了統計學習效果，華母語的受試者在處理一串西伯來文字母時，也展現了統計學習效果。這些結果表示統計學習可以發生在以視覺文字為材料的閱讀上。本研究進一步使用與閱讀較相近的RSVP實驗派典探討閱讀中的統計學習。實驗的結果顯示，初學的華二語受試展現了顯著的統計學習效果(.57)，效果的大小與我們先前以Saffran等人的實驗派典和聽覺呈現的方式所得到的效果相當(.57)，也與Saffran等人早先以英語音節為材料所得到的效果相近(.58)。本研究的結果說明了(1) 統計學習需有序列處理的要求才會發生，(2) 閱讀中文無詞間空格的句子時，所必須進行的斷詞處理，可能是建立在統計學習的機制上。本研究的發現對華語的閱讀教學有一定的啟發。

中文關鍵詞：統計學習、華語、中文、閱讀

英文摘要：The present study applied the Fisher and Aslin paradigm to investigating statistical learning in reading a string of linguistic symbols (Korean letter, Hebrew letters, and Chinese characters). Statistical learning was observed regardless of the type of orthography and orthographic similarity between the participants’ native language and the experimental stimuli. The results indicate that statistical learning can be induced and observed with visual linguistic stimuli. More interestingly, the present study also succeeded in inducing and observing statistical learning with the RSVP paradigm, which is more similar to normal reading than the Fiser and Aslin paradigm. The statistical learning effect (.57) observed in the RSVP of Chinese characters was comparable to the effect (.57) observed with the auditory version of the Saffran et al. paradigm using Chinese characters in our previous study. It was also similar to the effect (.58) originally reported by Saffran et al. with auditory English syllables. Two conclusions can be drawn from the current and our previous investigations of statistical learning in reading Chinese. First, in reading unspaced Chinese sentences, Chinese readers compute statistical information contained in the sentence and use that information to find word boundaries and identify words. Second, statistical learning requires sequential processing of input, and can be observed only if a task can effectively induce sequential processing. The results of the present study bear interesting implications for the teaching of reading in Chinese.

英文關鍵詞：Statistical learning, reading, Chinese, visual linguistic input, word spacing

科技部補助專題研究計畫成果報告

（□期中進度報告/■期末報告）

再探閱讀中文文句時的斷詞是否為一種基於統計學習的結果

計畫類別：☑個別型計畫　□整合型計畫
計畫編號：MOST108-2410-H-003-043-
執行期間：108 年 08 月 01 日至 110 年 01 月 31 日

執行機構及系所：國立臺灣師範大學華語文教學系

計畫主持人：陳振宇
共同主持人：
計畫參與人員：王贊育、林佩玲

本計畫除繳交成果報告外，另含下列出國報告，共 0 份：
□執行國際合作與移地研究心得報告
□出席國際學術會議心得報告
□出國參訪及考察心得報告

本研究具有政策應用參考價值：　□否　■是，建議提供機關 教育部
(勾選「是」者，請列舉建議可提供施政參考之業務主管機關)
本研究具影響公共利益之重大發現：■否　□是

中　華　民　國 110 年　01 月　日

目　錄

摘要

本研究採用Fiser and Aslin的實驗派典繼續探討受試者處理一串視覺語言符號時，能否從中攫取隱含的統計資訊，實驗的結果顯示，初學的華二語受試者在處理一串中文字的時候展現了統計學習效果，華母語的受試者在處理一串西伯來文字母時，也展現了統計學習效果。這些結果表示統計學習可以發生在以視覺文字為材料的閱讀上。本研究進一步使用與閱讀較相近的RSVP實驗派典探討閱讀中的統計學習。實驗的結果顯示，初學的華二語受試展現了顯著的統計學習效果(.57)，效果的大小與我們先前以Saffran等人的實驗派典和聽覺呈現的方式所得到的效果相當(.57)，也與Saffran等人早先以英語音節為材料所得到的效果相近(.58)。本研究的結果說明了(1) 統計學習需有序列處理的要求才會發生，(2) 閱讀中文無詞間空格的句子時，所必須進行的斷詞處理，可能是建立在統計學習的機制上。本研究的發現對華語的閱讀教學有一定的啟發。

關鍵詞：統計學習、華語、中文、閱讀

Abstract

The present study applied the Fisher and Aslin paradigm to investigating statistical learning in reading a string of linguistic symbols (Korean letter, Hebrew letters, and Chinese characters). Statistical learning was observed regardless of the type of orthography and orthographic similarity between the participants' native language and the experimental stimuli. The results indicate that statistical learning can be induced and observed with visual linguistic stimuli. More interestingly, the present study also succeeded in inducing and observing statistical learning with the RSVP paradigm, which is more similar to normal reading than the Fiser and Aslin paradigm. The statistical learning effect (.57) observed in the RSVP of Chinese characters was comparable to the effect (.57) observed with the auditory version of the Saffran et al. paradigm using Chinese characters in our previous study. It was also similar to the effect (.58) originally reported by Saffran et al. with auditory English syllables. Two conclusions can be drawn from the current and our previous investigations of statistical learning in reading Chinese. First, in reading unspaced Chinese sentences, Chinese readers compute statistical information contained in the sentence and use that information to find word boundaries and identify words. Second, statistical learning requires sequential processing of input, and can be observed only if a task can effectively induce sequential processing. The results of the present study bear interesting implications for the teaching of reading in Chinese.

Keywords: Statistical learning, reading, Chinese, visual linguistic input, word spacing

A further investigation of visual statistical learning for word segmentation during reading of Chinese sentences

Introduction

It is well known that Chinese sentences consist of strings of characters with no explicit spaces between words. How the readers of Chinese segment and identify words in a sentence remains a mystery. Past psycholinguistic research on this issue has not addressed it directly, but, instead, turned the question around and investigated the psychological reality of a word in Chinese (Cheng, 1981; Hue, 1989; Hoosain, 1992; J.-Y. Chen, 1998; Peng & Chen, 2004) or asked whether inserting word spaces or other physical cues could facilitate reading (Liu, Yeh, Wang, & Chang, 1974; Yang, 1998; Bai, Yan, Liversedge, Zang, & Rayner, 2008; Bai, Liang, Blythe, Zang, Yan, & Liversedge, 2013; Li, Rayner, & Cave, 2009; Shen et al., 2018). Research in computational linguistics has investigated the statistical meaning of a word and proposed a few algorithms for automatic word segmentation by machines (Huang, Chen, Chen, & Chang, 1997; Tang, Geva, Xu, & Trotman, 2009; Shu, Wang, Shen, & Qu, 2017). Few studies have examined whether Chinese readers represent words statistically and whether they use the statistical and distributional information to segment and identify words while reading. The lack of such studies may be due to the fact that the questions have been (mis)construed in the context of skilled reading while it is really learning that the questions are about. That is, the questions to ask are whether beginning readers of Chinese can learn to capture the statistical and distributional information in a string of characters and use that information to find the words therein.

There is now ample evidence to suggest that the answers to the above questions are positive. Saffran, Aslin, and Newport's seminal work (1996) demonstrated the potential of 8-month-old infants in capturing the transitional probabilities (TPs) between syllables in a continuous flow of speech-like sounds and used them to distinguish syllable strings with high TPs ("words") from those with low TPs ("nonwords"). Subsequent research replicated this statistical learning effect in infants and adults with a natural language (Saffran, Newport, Aslin, Tunick, & Barrueco, 1997; Hay, Pelucchi, Graf Estes, & Saffran, 2011; Kittleson, Aguilar, Tokerud, Plante, & Asbøjrnsen, 2010). Additional research shows that statistical learning extends to visual, tactile, and kinesthetic modalities as well as nonlinguistic stimuli, pointing to its domain-general nature (Fiser & Aslin, 2002; Conway & Christiansen, 2005; Baldwin, Anderson, Saffran, & Meyer, 2008; Aslin, Saffran, & Newport, 1999; Frost, Armstrong, Siegelman, & Christiansen, 2015). Moreover, it is a type of implicit learning, requiring little attention (Baker, Olson, & Behrmann, 2004; Toro, Sinnett, & Soto-Faraco, 2005; Perruchet & Pacton, 2006; Kim, Seitz, Feenstra, & Shams, 2009; Kittleson et al., 2010; Hamrick & Rebuschat, 2012). For recent reviews, see Armstrong, Frost, and Christiansen (2017) and Cunillera and Guilera (2018).

Although statistical learning has been examined in the visual modality, existing research has all used visual nonlinguistic stimuli. Visual linguistic stimuli such as the script of a language have not been investigated. The Chinese script, with its characters and no-word-spacing sentences, offers a perfect testing ground for such an investigation. Our laboratory has made a few attempts to investigate whether statistical regularity between characters could be captured by beginning learners of Chinese. We briefly summarize our findings below. A detailed report is available at the Ministry of Science and Technology website.

Our initial attempt repeated Saffran, Newport, Aslin, Tunick, and Barrueco's (1997) experiment but manipulated the transitional probabilities between six Chinese syllables (gu3, ke4, xi4, qu1, pi2, xian1). Six "words" (gu3ke4, xi4ke4, gu3qu1, pi2qu1, ke4xian1, xian1pi2) were defined where the TPs between characters ranged from .46 to 1. These "words" were judged to be not real words whether read forwards or backwards. Each was repeated 300 times and randomly concatenated to form a continuous string of 3600 syllables. The TPs between syllables across "words" ranged from 0 to .29. The syllable string was artificially pronounced by a speech synthesizer developed by the Industrial Technology Research Institute (ITRI) of Taiwan, R.O.C. The syllable string was played back to 20 foreign learners of Chinese while they performed a paper coloring task. In the subsequent two-alternative-forced-choice (2AFC) task, they could distinguish "words" from "nonwords" with a mean accuracy of .57 (s.e. = .08), which was significantly greater than .5 ($p < .05$). The accuracy was similar to that observed by Saffran et al. (1997), which was .58. The result demonstrated a successful replication of Saffran et al. (1997).

Our next attempt was to turn the above experiment into the visual mode, presenting the syllable string as a character string. The six characters were 古，文，唐，詩，仙，人. The six words were 古文，唐詩，仙人，古詩，詩仙，人文. The string of 3600 characters were divided into 36 consecutive blocks with each block of 100 characters presented in one single screen. In a screen, the 100 characters was presented one at a time at a rate of .5 second per character. Sporadically (0~.03), the presentation rate doubled for two consecutive characters and the participants were required to press a button to indicate that they had noticed the change. In the subsequent 2AFC task, they could barely distinguish words from nonwords with a mean accuracy of .53 (s.e. = .08), which was not significantly greater than the .5 chance expectation. Although the result could be shown to be significant if the accuracy rate for each participant was analyzed first and the p values pooled meta-analytically, the mean accuracy of .53 was obviously lower than the .57 accuracy in the previous auditory experiment.

To boost the statistical learning effect in reading Chinese, the subsequent attempts (1) required the participants to read out loud each character rather than detecting a change in presentation rate, (2) doubled the exposure, or (3) reduced the complexity of characters. The mean accuracy rates for these attempts were .48, .52, and .52, respectively, and not significant. The results suggest that the Saffran paradigm did not work well with the visual stimuli.

In a further experiment, we adopted the Fiser and Aslin's (2002) task for presenting abstract geometric shapes. A rectangular bar was present on the screen. A geometric shape appeared from behind the bar, moved to one edge of the screen, returned, and disappeared into the back of the bar. The next shape then appeared but moved to the other edge of the screen. In our experiment, we used Korean characters instead of geometric shapes. Twelve characters (ㄷ、ㅎ、ㄹ、ㅋ、ㅊ、ㅡ、ㅗ、ㅁ、ㅠ、ㅂ、ㅏ、ㄴ) were used to construct four "words", each consisting of three characters. Each "word" was repeated 24 times and randomly concatenated into a string of 288 characters, with the restriction that the same 'word' does not appear twice in a row. After the exposure/learning, the participants took a 2AFC test, in which they indicated which of the two character strings (shown in the same way as in the learning phase) was the one they saw before. "Words" and "nonwords" were defined and distinguished by co-occurrence probabilities (which actually correlate highly with TPs; TPs for "words" were 1 and TPs for "nonwords" were 0). The result from a group of native Chinese participants with no knowledge of the Korean script showed a statistically significant mean accuracy rate of .65, indicating a successful statistical learning of Korean "words".

Experiment 1

Our attempts at investigating statistical learning in reading scripts found that the Fiser and Aslin paradigm was more effective than the visually adapted version of the Saffran paradigm. But the success of statistical learning was demonstrated with the Chinese speakers reading the Korean script. The success could be due to the paradigm. It could also be due to a differential transfer effect. The Chinese speakers exposed to the Fiser and Aslin paradigm were experienced readers of Chinese characters and they might have benefited from a transfer of such an experience to the reading of the Korean characters. The foreign speakers exposed to the Saffran paradigm, on the other hand, were experienced readers of alphabet, and that experience could not easily transfer to the processing of Chinese characters. To determine if the native language experience contributes to the differential success of the Saffran paradigm and the Fiser and Aslin paradigm, Experiment 1 was conducted using the Fiser and Aslin paradigm but applied to a group of foreign learners of Chinese (with an alphabetic language background) using Chinese characters as the visual linguistic stimuli.

Method

Participants

Thirty non-Chinese native speakers, aged from 18-66 years old (*M* =27.3, *SD* = 8.83). Participants were all beginners of Chinese language, their learning hour was from 8-240 hours (*M* =104.1, *SD* = 90.7). Their 1st language was English, Russian, Italian, Spanish, French, German, Wolof, Pilipino, Bahasa, and Thai language.

Materials

Six experimenter defined Chinese words(三上、久也、女上、三也、大女、也大) were made from 6 Chinese characters, none of these words can be found in real usage. Each Chinese word was repeated for 300 times then composed a continuous pseudorandom Chinese word stream which contained 3600 Chinese characters. In this stream, same Chinese words were not displayed subsequently. The TPs of adjacent characters between words were from 0 - .29, the TPs of adjacent characters within words were from .46 – 1. The TPs among the adjacent Chinese characters were the only clue for participants determining word boundaries. The 2AFC test contained 36 questions, each question was made of one defined word and one reversed word (non-word). Test questions were randomly presented. The sequences of the choice were counter-balanced.

Procedure

The general procedure was adopted from Fiser and Aslin's visual statistical learning paradigm (Fiser & Aslin, 2002). Participants were asked to watch an animation of the stream of 3600 Chinese characters in subsequent days. For each day, participants watched 1800 Chinese characters of the stream, in which characters were displayed separately in 9 sessions. When the animation started, a black rectangle was displayed at the center of the screen as the occluder. One Chinese character was moving at a constant speed from the back of the occluder to the edge of the screen, then moving to the back of the occluder. The next Chinese character showed from the opposite side of the occluder then moved to the edge of the other side of the screen and moving to the back of the occluder. The animation was kept playing till it ran out of 3600

Chinese characters of the stream. Each Chinese character was fully displayed on the screen for 2 seconds. After finishing the passive viewing stage of the 2nd day, participants were asked to complete a 2AFC test. In the test, the participants' task was to decide which combination or sequence of Chinese characters is more like the one they had seen in the learning stage, then pressed the keyboard to indicate their answer. The method of displaying the test choices was exactly the same as the method of displaying the Chinese character in the learning stage. There was no time limitation of the test, participants were encouraged to follow their instincts if they were not sure about the answer. At the end of the experiment, the participants completed a questionnaire about the Chinese characters used in the experiment. We investigated participants' understanding of the Chinese characters in phonetic, semantic and syntactic levels through the following questions: (1) What is the pinyin of the Chinese character? (2) What is the meaning of the Chinese character? (3) Please raise some examples of words that contain the experiment Chinese characters.

Results

30 participants' average accuracy of the test was .67(SD = .16, Max =1, Min = .44), which was higher than the random guessing rate (.5) significantly: $t_1(29) = 6.04$, $p < .001$. Cohen's $d = 1.10$. 25 participants' accuracy were above .5, sign test, $p < .001$.

The responses of the questionnaire were collected and analyzed. For each Chinese character, participants got 2 points for correctly answered both the pinyin and the correspondent meaning of the Chinese character, 1 point for answered correctly either the pinyin or the meaning of the Chinese character, 0 points for no correct answer neither in the pinyin nor in the meaning of the character. The average grade of the questionnaire of 30 participants was 8.43. We divided participants into 2 groups by 6 points of their questionnaire score. 20 participants got above 6 points of the questionnaire, 10 participants got 6 or under 6 points of the questionnaire. There was no significant difference in the average accuracy of the test between two groups: $F(1,28) = 0.01$, $p > .05$, $MSe = 0.03$.

Discussion

The results of Experiment 1 demonstrated that non-natives of Chinese could extract statistical information from a continuous Chinese character stream. Participants were not told about the statistical regularities of the Chinese character stream, however, they could find word boundaries and segment character combinations based on the statistical regularities from it. The different degrees of participants' previous knowledge of the experimental characters did not affect their statistical learning outcomes of the Chinese character stream. To confirm further that the success of the Fiser and Aslin paradigm was not due to the orthographic similarities between the participants' native language and the language of the experimental stimuli, Experiment 2 was conducted by applying the paradigm to native Chinese speakers but with Hebrew letters. Hebrew letters are alphabetic, bearing no similarity to Chinese characters.

Experiment 2

Method

Participants

Fifteen native Chinese speakers, aged from 20-25 years old (M = 22.3, SD = 1.67). None of the participants had experience of the Hebrew language.

Materials

Twelve Hebrew letters were selected and divided into 4 different letter combinations: שקט, בדם, למג, אףצ (hereafter called “word”). Each word was repeated 24 times before being connected as a continuous pseudo-random Hebrew letter stream. The TP of adjacent Hebrew letters between words was .33, and the TP of adjacent Hebrew letters within words was 1. The TPs among the adjacent Hebrew letters were the only clue for participants determining word boundaries. The test contained 32 questions. Each question was made of one defined Hebrew word (e.g. ףצא) and a non-word (e.g. קמא).

Procedure

The procedure was the same as the procedure used in our previous experiment with the Korean letters.

Results

The average accuracy of the 15 participants on the test was .71 (SD = .25, Max = 1, Min = .28), which was significantly greater than the .5 random guessing rate: $t(14) = 3.17$, $p < .05$, Cohen’s $d = 1.19$. Ten out of 15 participants scored above .5 (sign test: $z = 1.94$, $p < .05$).

Discussion

Experiment 3 showed that Chinese participants could segment Hebrew words from a continuous Hebrew letter stream. The participants did not know Hebrew letters, nor the regularities of the letter stream they were exposed to. However, they could distinguish the difference between defined words and non-words based on the embedded statistical information from the adjacent Hebrew letters. The results of Experiment 2 and Experiment 3, along with our previous experiment using the Korean letters, indicated that statistical learning could be induced and observed with the Fiser and Aslin paradigm, and that this has little to do with the type of orthography or the linguistic similarity between orthographies.

Experiment 3

Although the Fiser and Aslin paradigm is useful for investigating statistical learning of visual linguistic stimuli, the task does not resemble the normal way of reading linguistic scripts. To understand how readers of Chinese LEARN to segment words in a Chinese sentence, we need to find a task as similar to normal reading as possible and to demonstrate a statistical learning effect with Chinese characters in foreign learners of Chinese. A good candidate is the Rapid Serial Visual Presentation (RSVP) of characters which has been used in past investigations of visual word recognition (Potter, 1983; Yen & Chien, 2011; Öquist & Goldstein, 2002; Cao, Yang, & Yan, 2017). The purpose of Experiment 3 was to investigate visual statistical learning of

Chinese characters using the RSVP paradigm. In order to ensure that the participants kept their eyes on viewing the fast flashing characters shown in the RSVP task, three different kinds of primary task was attempted. One required the participants to detect English letters that sporadically appeared among the characters. In Experiment 3-1, the participants needed to press a button whenever they detected an English letter. In Experiment 3-2, they were not required to respond. Experiment 3-3 employed a different primary task as a cover task. The participants were simply told to pay attention to the flashing characters because they would be asked to write down the characters at the end of the experiment.

Method

Participants

Experiment 3-1 recruited 15 foreign beginning learners of Chinese (4 males, 11 females; 18-49 years of age (M = 24.2, SD = 7.57)。Hours of learning Chinese ranged from 30 to 180 (M = 98.1, SD = 70.1). Their mother languages included Polish, Kazakh, Russian, French, English, German, and Spanish. They were all at the beginners' level of learning Chinese, as inferred from the fact that they were all learning from the first book of various Chinese language textbooks (當代中文 1、精通中文 1、視聽華語 1).

Experiment 3-2 recruited another 15 foreign beginning learners of Chinese (9 males, 6 females; 18-46 years of age (M = 22.5, SD = 7.1)。Hours of learning Chinese ranged from 75 to 225 (M = 166.7, SD = 54). Their mother languages included French, English, German, Spanish, Tagalog, Portuguese, and Indonesian. They were all at the beginners' level of learning Chinese, as inferred from the fact that they were all learning from the first book of various Chinese language textbooks (當代中文 1、菁華中文 1、視聽華語 1).

Experiment 3-3 recruited yet another 15 foreign beginning learners of Chinese (8 males, 7 females; 21-65 years of age (M = 31.7, SD = 11.2)。Hours of learning Chinese ranged from 45 to 150 (M = 72, SD = 31.2). Their mother languages included French, English, German, Swedish, Tagalog, and Thai. They were all at the beginners' level of learning Chinese, as inferred from the fact that they were all learning from the first book of various Chinese language textbooks (當代中文 1，精通中文 1).

Materials

The stimuli used for learning and testing were the same as those in Experiment 1. The 3600 characters of the stimulus string for learning were divided into nine sets. For Experiments 3-1 and 3-2, each set of 400 characters were mixed with 20 English letters (randomly chosen for each set). The English letters were randomly inserted into the 400 Chinese characters, with the restrictions that they did not appear at the beginning or end of the string and that they appeared between "words" and not within a "word".

Procedure

For Experiment 3-1 and 3-2, before the experimental task was given, the participants were shown the 6 stimulus characters, told the pronunciation of each, and checked to ensure they could tell apart the six characters. At the end of the experiment, the participants were given a memory test, in which they were shown the 6 characters and asked to write down the pinyins and the meanings. For Experiment 3-3, the participants were not shown and familiarized with the stimulus characters before the experimental task was given. Instead, they were told that a memory test of the characters would be given at the end of the experiment, in which they had to write down the six characters that appeared in the subsequent viewing task.

For the experiment task, the nine stimulus sets were administered in nine separate sessions, with a break in between. A session started with a fixation cross displayed at the center of the screen. Pressing the spacebar removed the cross and brought in the stimulus string. The string of Chinese characters (and English letters in Experiment 3-1 and 3-2) were shown one at a time at the center of the screen. Each character or letter appeared for 350 ms, followed by a 150-ms blank screen. The participants were asked to view each character and to press ‘1’ on the numeric keypad (Experiment 3-1) or to give no response (Experiment 3-2) whenever and as soon as an English letter appeared. The next character came on regardless of whether they responded in time. For Experiment 3-3, there were no English letters in the stimulus string. The participants were simply reminded that they would be asked to write down the Chinese characters at the end of the experiment. At the end of the learning, the participants were given a surprise 2AFC test, in which they were shown two character strings, one above the other, and they had to indicate which string was the one they had seen in the sequence during learning. Following the 2AFC test, the memory test was given.

Results

Experiment 3-1

The mean accuracy of the 15 participants was .53 (*SD*= .09, *Max* = .64, *Min* = .39), which was not significantly greater than the .5 guessing rate: *t*(14) = 1.29, *p* = .11, Cohen’s *d* = .47. Nine out of the 15 participants scored above .5. The mean signal detection *d’ was* .11, (*SD* = .33, *Max* = .52, *Min* = -.40). Table 1 presents the proportion of responses cross-tabulated according to the type of stimulus and the type of responses.

Table 1. Crosstabulation of the proportion of responses in Experiment 3-1.

		Stimulus	
		Word	Nonword
Participants’ response	Word (seen)	0.54	0.48
	Nonword (not seen)	0.46	0.52

The result of the memory test showed a mean score of 5.8 (*SD*= .77, *Max* = 6, *Min* = 3) for pinyin and 4.93 (*SD*= .96, *Max* = 6, *Min* = 2) for meaning. Because there was no significant effect of statistical learning and the result of the memory test approach ceiling, there was no need to run a correlation test between the two.

Experiment 3-2

The mean accuracy of the 15 participants was .53 (*SD*= .10, *Max* = .67, *Min* = .36), which was not significantly greater than the .5 guessing rate: *t*(14) = 1.02, *p* = .16, Cohen’s *d* = .42. Eight out of the 15 participants scored above .5. The mean signal detection *d’ was* .10, (*SD* = .39, *Max* = .68, *Min* = -.50). Table 1 presents the proportion of responses cross-tabulated according to the type of stimulus and the type of responses.

Table 2. Crosstabulation of the proportion of responses in Experiment 3-2.

		Stimulus	
		Word	Nonword
Participants’ response	Word (seen)	0.57	0.51

	Nonword (not seen)	0.43	0.49

The result of the memory test showed a mean score of 5.67 (*SD* = .49, *Max* = 6, *Min* = 5) for pinyin and 5.53 (*SD*= .64, *Max* = 6, *Min* = 4) for meaning. Because there was no significant effect of statistical learning and the result of the memory test approach ceiling, there was no need to run a correlation test between the two.

Experiment 3-3

The mean accuracy of the 15 participants was .57 (*SD* = .14, *Max* = .78, *Min* = .25), which was significantly greater than the .5 guessing rate: $t(14) = 3.06$, $p < .01$, Cohen's $d = 1.16$. Eleven out of the 15 participants scored above .5. The mean signal detection *d' was* .26, (*SD* = .55, *Max* = 1.10, *Min* = -.96). Table 1 presents the proportion of responses cross-tabulated according to the type of stimulus and the type of responses.

Table 3. Crosstabulation of the proportion of responses in Experiment 3-3.

		Stimulus	
		Word	Nonword
Participants' response	Word (seen)	0.56	0.41
	Nonword (not seen)	0.44	0.59

The result of the memory test showed a mean score of 3.95 (*SD* = 1.82, *Max* = 6, *Min* = 1) for pinyin and 4.45 (*SD* = 1.36, *Max* = 6, *Min* = 2) for meaning. The accuracy of the remembered characters was fairly high, with a mean of 5.85 (*SD* = .37, *Max* = 6, *Min* = 5).

Discussion

Employing the RSVP paradigm, which was a better simulation of normal reading than the Fiser and Aslin paradigm, Experiment 3-3 observed a significant effect of statistical learning with Chinese characters, but Experiment 3-1 and 3-2 did not. Requiring an explicit motoric response does not convincingly explain the failure of statistical learning in the latter experiments because Experiment 3-2 did not require it. The presence of the sporadically shown English letters obviously had something to do with the failure of statistical learning in these experiments, but it is not clear in what way. Attention might be an explanation. The sporadic appearance of an English letter among a sequence of Chinese characters might induce an attentional capture effect (Yantis, 1996), disrupt the implicit computation of the transitional probabilities between characters in the stimulus string, and result in the failure of statistical learning. The cover task adopted in Experiment 3-3 served to direct the participants' attention to the characters themselves (and quite effectively, based on the high accuracy of the participants' memory of the characters) and there were no extraneous stimuli or tasks to distract their attention away from the character string. With this kind of cover task, the sequential nature of the RSVP paradigm was effective to induce computation of transitional probabilities between characters and statistical learning.

General Discussion

The present study applied the Fisher and Aslin paradigm to investigating statistical learning in reading a string of linguistic symbols (Korean letter, Hebrew letters, and Chinese characters). Statistical learning was observed regardless of the type of orthography and orthographic similarity between the participants' native language and the experimental stimuli. The results indicate that statistical learning can be induced and observed with visual linguistic stimuli. More interestingly, the present study also succeeded in inducing and observing statistical learning with the RSVP paradigm, which is more similar to normal reading than the Fiser and Aslin paradigm. The statistical learning effect (.57) observed in the RSVP of Chinese characters was comparable to the effect (.57) observed with the auditory version of the Saffran et al. paradigm using Chinese characters in our previous study. It was also similar to the effect (.58) originally reported by Saffran et al. with auditory English syllables.

Two conclusions can be drawn from the current and our previous investigations of statistical learning in reading Chinese. First, in reading unspaced Chinese sentences, Chinese readers compute statistical information contained in the sentence and use that information to find word boundaries and identify words. Second, statistical learning requires sequential processing of input, and can be observed only if a task can effectively induce sequential processing. However, we quickly note that statistical learning may not depend on sequential processing entirely. Transitional probabilities between characters can be computed and accumulated when a string of characters appears in isolation as in a typical and common word learning experience. Nonetheless, sequential processing may still be an essential characteristic of statistical learning.

References

Armstrong, B. C., Frost, R., & Christiansen, M. H. (2017). The long road of statistical learning research: past, present and future. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *372*(1711), 20160047.

Aslin, R. N., Saffran, J. R., & Newport, E. L. (1999). Statistical learning in linguistic and nonlinguistic domains. In B. MacWhinney (Ed.), *The emergence of language* (pp. 359-380). Mahwah, NJ: Erlbaum.

Bai, X.-J., Liang, F.-F., Blythe, H. I., Zang, C.-L., Yan, G.-L., & Liversedge, S. P. (2013). Interword spacing effects on the acquisition of new vocabulary for readers of Chinese as a second language. *Journal of Research in Reading*, *36*(S1), S4-S17.

Bai, X.-J., Yan, G.-L., Liversedge, S. P., Zang, C., & Rayner, K. (2008). Reading spaced and unspaced Chinese text: Evidence from eye movements. *Journal of Experimental Psychology: Human Perception and Performance, 34*(5), 1277-1287.

Baker, C. I., Olson, C. R., & Behrmann, M. (2004). Role of attention and perceptual grouping in visual statistical learning. *Psychological Science*, *15*(7), 460–466.

Baldwin, D., Anderson, A., Saffran, J., & Meyer, M. (2008). Segmenting dynamic human action via statistical structure. *Cognition*, *106*(3), 1382–1407.

Cao, H.-W., Yang, K.-Y., & Yan, H.-M. (2017). Character Decomposition and Transposition Processes of Chinese Compound Words in Rapid Serial Visual Presentation. *Frontiers in Psychology, 8*, 483. DOI=10.3389/fpsyg.2017.00483

Chen, J.-Y. (1998). Word recognition during reading of Chinese sentences: Evidence from studying the word superiority effect. In A. Inhoff, J. Wang, & H. C. Chen (Eds.), *Reading Chinese script: A cognitive analysis* (pp. 239-256). N.J.: Erlbaum.

Chen, J.-Y., & Li, C.-Y. (2016, Aug). *Technological shaping of working memory: A difference between Chinese phonology-based and orthography-based typing*. The 38th Annual Meeting of the Cognitive Science Society, Philadelphia, Pennsylvania, USA.

Cheng, C.-M. (1981). The cognitive processes of Chinese word recognition. *Chinese Journal of Psychology*, 2*3*, 137-153. (in Chinese: 鄭昭明，漢字的認知歷程)

Chuang, Y.-L. (2018). *The facilitatory effect of word-unit presentation on reading Chinese with an application in teaching Chinese as a second language*. Unpublished doctoral thesis, National Taiwan Normal University.

Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review, 12*(5), 769-786. doi: 10.3758/bf03196772

Conway, C. M., & Christiansen, M. H. (2005). Modality-constrained statistical learning of tactile, visual, and auditory sequences. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *31*(1), 24-39.

Cornoldi, C., & Vecchi, T. (2003). *Visuo‐spatial working memory and individual differences*. Hove, UK: Psychology Press.

Cunillera, T., & Guilera, G. (2018). Twenty years of statistical learning: from language, back to machine learning. *Scientometrics*, *117*, 1–8.

Fiser, J. Z., & Aslin, R. N. (2002). Statistical learning of higher-order temporal structure from visual shape sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(3), 458–467.

Frost, R., Armstrong, B. C., Siegelman, N., & Christiansen, M. H. (2015). Domain generality versus modality specificity: The paradox of statistical learning. *Trends in Cognitive Sciences*, *19*(3), 117–125.

Hamrick, P., & Rebuschat, P. (2012). How implicit is statistical learning. In J. N., Williams and P. Rebuschat (Eds.), *Statistical learning and language acquisition* (pp. 365–382). Boston/Berlin: Walter de Gruyter.

Hay, J., Pelucchi, B., Graf Estes, K., & Saffran, J. R. (2011). Linking sounds to meanings: Infant statistical learning in a natural language. *Cognitive Psychology*, *63*, 93–106.

Hoosain, R. (1992). Psychological reality of the word in Chinese. In H.-C. Chen, & O. J. L. Tzeng (Eds.), *Language processing in Chinese* (pp. 111-130). Amsterdam: North-Holland.

Huang, C.-R., Chen, K.-J., Chen, F.-Y., & Chang, L.-L. (1997). Segmentation standard for Chinese natural language processing. *Computational Linguistics and Chinese Language Processing, 2*(2), 47-62.

Hue, C.-W. (1989). Word superiority effect and inferiority effect: A study of Chinese word recognition. *Chinese Journal of Psychology*, *31*, 33-39.

Kim, R., Seitz, A., Feenstra, H., & Shams, L. (2009). Testing assumptions of statistical learning: Is it long-term and implicit? *Neuroscience Letters*, *461*(2), 145–149.

Kittleson, M. M., Aguilar, J. M., Tokerud, G. L., Plante, E., & Asbøjrnsen, A. E. (2010). Implicit language learning: Adults' ability to segment words in Norwegian. *Bilingualism*, *13*(4), 513-523.

Li, X.-S., Rayner, K., & Cave, K. (2009). On the segmentation of Chinese words during reading. *Cognitive Psychology, 58*, 525-552.

Liu, I.-M., Yeh, C.-H., Wang, L.-H., & Chang, Y.-K. (1974). The effect of word spacing on reading efficiency in Chinese. *Chinese Journal of Psychology, 16*, 25-32.

Öquist, G.; Goldstein, M. (2002). Towards an improved readability on mobile devices: Evaluating adaptive Rapid Serial Visual Presentation". *Human Computer Interaction with Mobile Devices. Lecture Notes in Computer Science*, 2411, 225. doi:10.1007/3-540-45756-9_18.

Paribakht, T., & Wesche, M. (1993). The relationship between reading comprehension and second language development in a comprehension-based ESL program. *TESL Canada Journal, 11*(1), 9-29.

Peng, R.-Y., & Chen, J.-Y. (2004). Even words are right, odd ones are odd: Explaining the word segmentation inconsistency among Chinese readers. *Chinese Journal of Psychology, 46*, 049-055. (in Chinese: 彭瑞元、陳振宇，「偶語易安、奇字難適」：探討中文讀者斷詞不一致之原因)

Perruchet, P., & Pacton, S. (2006). Implicit learning and statistical learning: One phenomenon, two approaches. *Trends in Cognitive Sciences*, *10*(5), 233–238.

Potter, M. C. (1983). Rapid serial visual presentation (RSVP): A method for studying language processing. In D. Kieras and M. Lust (Eds.), *New methods in reading comprehension research* (pp. 91-118). Hillsdale, N.J.: LEA.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*(5294), 1926–1928.

Saffran, J. R., Newport, E. L., Aslin, R. N., Tunick, R. A., & Barrueco, S. (1997). Incidental language learning: Listening (and learning) out of the corner of your ear. *Psychological Science*, *8*, 101–105.

Shu, X., Wang, J., Shen, X., & Qu, P. (2017). Word segmentation in Chinese language processing. *Statistics*

and its Interface, *10*(2), 165-173. https://doi.org/10.4310/SII.2017.v10.n2.a1
Tang, L. X., Geva, S., Xu, Y., & Trotman, A. (2009). Word segmentation for Chinese Wikipedia using n-gram mutual information. *Proceedings of the 14th Australasian Document Computing Symposium*, Sydney, Australia, 4 December.
Toro, J. M., Sinnett, S., & Soto-Faraco, S. (2005). Speech segmentation by statistical learning depends on attention. *Cognition*, *97*(2), B25–B34.
Yan, M., Kliegl, R., Richter, E. M., Nuthmann, A., Shu, H. (2010). Flexible saccade-target selection in Chinese reading. *Quarterly Journal of Experimental Psychology, 63*, 705–725.
Yantis, S. (1996). *Attentional capture in vision.* In A. F. Kramer, M. G. H. Coles, & G. D. Logan (Eds.), *Converging operations in the study of visual selective attention* (p. 45–76). American Psychological Association. https://doi.org/10.1037/10187-002
Yen, C. C., & Chien, Y. H. (2011). Rapid serial visual presentation display on a small screen: reading Chinese while walking. *Perceptual and Motor Skills, 112*(3), 991-1003.

108年度專題研究計畫成果彙整表

<table>
<tr><td colspan="3">計畫主持人：陳振宇</td><td colspan="3">計畫編號：108-2410-H-003-043-</td></tr>
<tr><td colspan="6">計畫名稱：再探閱讀中文文句時的斷詞是否為一種基於統計學習的結果</td></tr>
<tr><td colspan="3">成果項目</td><td>量化</td><td>單位</td><td>質化
（說明：各成果項目請附佐證資料或細項說明，如期刊名稱、年份、卷期、起訖頁數、證號...等）</td></tr>
<tr><td rowspan="6">國內</td><td rowspan="6">學術性論文</td><td>期刊論文</td><td>0</td><td rowspan="2">篇</td><td></td></tr>
<tr><td>研討會論文</td><td>0</td><td></td></tr>
<tr><td>專書</td><td>0</td><td>本</td><td></td></tr>
<tr><td>專書論文</td><td>0</td><td>章</td><td></td></tr>
<tr><td>技術報告</td><td>0</td><td>篇</td><td></td></tr>
<tr><td>其他</td><td>0</td><td>篇</td><td></td></tr>
<tr><td rowspan="6">國外</td><td rowspan="6">學術性論文</td><td>期刊論文</td><td>0</td><td rowspan="2">篇</td><td></td></tr>
<tr><td>研討會論文</td><td>0</td><td></td></tr>
<tr><td>專書</td><td>0</td><td>本</td><td></td></tr>
<tr><td>專書論文</td><td>0</td><td>章</td><td></td></tr>
<tr><td>技術報告</td><td>0</td><td>篇</td><td></td></tr>
<tr><td>其他</td><td>0</td><td>篇</td><td></td></tr>
<tr><td rowspan="10">參與計畫人力</td><td rowspan="5">本國籍</td><td>大專生</td><td>0</td><td rowspan="10">人次</td><td></td></tr>
<tr><td>碩士生</td><td>0</td><td></td></tr>
<tr><td>博士生</td><td>1</td><td>本研究主題論文產出。</td></tr>
<tr><td>博士級研究人員</td><td>0</td><td></td></tr>
<tr><td>專任人員</td><td>1</td><td>協助處理研究計畫相關事宜。</td></tr>
<tr><td rowspan="5">非本國籍</td><td>大專生</td><td>0</td><td></td></tr>
<tr><td>碩士生</td><td>0</td><td></td></tr>
<tr><td>博士生</td><td>0</td><td></td></tr>
<tr><td>博士級研究人員</td><td>0</td><td></td></tr>
<tr><td>專任人員</td><td>0</td><td></td></tr>
<tr><td colspan="3">其他成果
（無法以量化表達之成果如辦理學術活動、獲得獎項、重要國際合作、研究成果國際影響力及其他協助產業技術發展之具體效益事項等，請以文字敘述填列。）</td><td colspan="3"></td></tr>
</table>