Şarkiyat Mecmuası
Journal of Oriental Studies

İSTANBUL UNIVERSITY PRESS

Şarkiyat Mecmuası - Journal of Oriental Studies 37, (2020): 51-64
DOI: 10.26650/jos.2020.007

Araştırma Makalesi / *Research Article*

Spaces between Words as a Visual Cue when Reading Chinese: An Eye-Tracking Study*

Çince Okumada Görsel Bir İpucu Olarak Sözcük Sınırı Kullanımı: Bir Göz İzleme Çalışması

Sıla AY[1], İsmigül CANTÜRK[2], Tuğba AKGÜR[2]

ABSTRACT

In texts written in languages that use the Latin alphabet, the spaces left between words serve as visual cues to understand the text. In the written Chinese language, there are no spaces between words. Chinese differs from alphabetic languages in many respects, including its synonymous and multi-meaning symbolic language elements. The absence of a visual clue to indicate word boundaries creates ambiguity in Chinese sentences or may lead to the emergence of different meanings. A large number of Chinese characters, spelling features, and lack of boundaries between words cause various difficulties for foreign students. Turkish students who study Chinese as a foreign language may find its unfamiliar orthographic features, such as a spelling system without spaces between words, difficult to understand as Turkish is written using an alphabet. According to some studies on the early stages of Chinese language learning, artificially-triggered spaces in writing can have a positive effect on the reading process. Yet, this finding that the spaces between Chinese words facilitate reading comprehension is controversial. In this study, an eye movement tracking technique is used to investigate whether orthographically-triggered spelling differences or adding spaces between words affect the reading process of foreign language students. The results are discussed through Chinese grammatical features.

Keywords: Chinese reading, Chinese orthography, word boundary, eye tracking

ÖZ

Latin alfabesi kullanan dillerde yazılı metinlerde sözcük aralarında boşluk bırakılmakta ve bu durum metnin anlaşılmasında görsel bir ipucu görevi görmektedir. Ancak bu çalışmaya konu olan Çincede, yazım sırasında sözcükler arasında boşluk yoktur. Çince eş biçimli ve çok anlamlı sembolik dil ögeleri ile alfabetik dillerden birçok açıdan ayrılmaktadır. Sözcük sınırlarını gösteren bir ipucunun olmayışı, Çince tümcelerde anlam belirsizliği yaratmakta veya farklı anlamların ortaya çıkmasına yol açabilir. Çince imlerin sayıca çok olması, yazım özellikleri ve sözcük sınırlarının olmaması yabancı öğrenciler için çeşitli zorluklara neden olmaktadır. Çinceyi yabancı dil olarak öğrenen Türk öğrenciler, Türkçenin alfabetik yazım sistemine sahip olması sebebiyle, sözcükler arasında boşluk bırakmayan ve aşina olmadıkları Çince yazım sistemini zor bulabilirler. Çince öğreniminin ilk aşamalarında yapılan bazı araştırmalara göre sözcükler

* This study was presented as a paper at the 32nd National Linguistics Congress organized by İzmir Dokuz Eylül University on 3-4 May 2018 with the title of "Çince Okumada Görsel Bir İpucu Olarak Sözcük Sınırı Kullanımı: Bir Göz İzleme Çalışması".

[1]Ankara Üniversity, Faculty of Languages and History-Geography, Ankara, Turkey

[2]Düzce University, School of Foreign Languages, Düzce, Turkey

ORCID: S.A. 0000-0002-8773-5522;
İ.C. 0000-0001-8827-3449;
T.A. 0000-0003-3593-8181

Corresponding author/Sorumlu yazar:
Sıla AY (Assoc. Prof.),
Ankara Üniversity, Faculty of Languages, History and Geography, Ankara, Turkey
E-mail: ay@ankara.edu.tr

Başvuru/Submitted: 27.10.2020
Revizyon Talebi/Revision Requested: 17.11.2020
Son Revizyon/Last Revision Received: 30.11.2020
Kabul/Accepted: 01.12.2020

Atıf/Citation: Ay, Sila, Canturk, Ismıgul, Akgur, Tugba. "Spaces between Words as a Visual Cue when Reading Chinese: An Eye-Tracking Study". *Şarkiyat Mecmuası - Journal of Oriental Studies* *37* (2020), 51-64. https://doi.org/10.26650/jos.2020.007

arasında boşluk bırakarak yazmak, okuma süreci üzerinde olumlu bir etkiye sahip olabilir. Çince sözcükler arasındaki boşlukların, Çince okumayı kolaylaştırıp kolaylaştırmadığı oldukça tartışmalı bir konudur. Göz hareketlerini izleme tekniğinin kullanıldığı bu çalışmada, ortografik olarak tetiklenen yazım farklılıklarının veya sözcükler arasına boşluk eklenmesinin Çince yabancı dil öğrencilerinin okuma sürecine etkisi olup olmadığı araştırılmıştır. Çalışma sonuçları, Çince dilbilgisi özellikleriyle birlikte ele alınmıştır.

Anahtar kelimeler: Çince okuma, Çince yazım, sözcük sınırı, göz izleme

1. Literature Review

Reading process in Chinese is different from reading alphabetic languages. Compared to alphabetic languages such as English, the notion of word in Chinese is not clear enough, and this makes it difficult to define Chinese word boundaries. The abundance of two-syllable characters and the fact that the syllables and semantic values of the lines and parts that make up these characters do not follow a certain rule, are amongst reasons that renders Chinese writing system a difficult one (Yen, Tsai, Tzheng and Huang, 2008). In teaching Chinese as a foreign language, the Chinese characters' structure and the lack of spaces between words in the sentence are seen as obstacles in developing the learners' reading comprehension skills, as these features are not compatible with the characteristics of their mother tongues. The ability to recognize characters, to distinguish them from other characters in the sentence and to understand the whole sentence involves a much more complex process. Therefore, the absence of space between words in Chinese sentences can cause ambiguity. The ambiguity of meaning in the written language is divided into different categories at word, phrase and sentence levels (see examples 1, 2 and 3 accordingly).

Example 1: Word level ambiguity:
长春 - (Chángchūn) - name of a city
长 - 春 - (cháng - chūn) - long spring

Example 2: Phrase level ambiguity:
学校医院 - (xuéxiàoyīyuàn) - hospital of school
学校 - 医院 - (xuéxiào yīyuàn) - school and hospital

Example 3: Sentence level ambiguity:
学习汉语 - 语法很重要。 - (Xuéxí Hànyǔ - yǔfǎ hěn zhòngyào) – Grammar is very important in learning Chinese.
学习汉语语法 - 很重要。- (Xuéxí Hànyǔ yǔfǎ - hěn zhòngyào) - Studying Chinese grammar is very important.

As it can be seen from the examples, the lack of clear word boundaries in Chinese and the absence of spaces between words appear as significant difficulties in teaching Chinese both as a native language and as a foreign language. In order to understand the learning process of learners and to examine the use of word boundaries as a visual cue for reading Chinese texts, studies that measure eye movements and comprehension were conducted and study results showed differences. For example, Zang et al. (2013) examined sixteen Chinese native children's and sixteen adults' eye movement when reading word spaced and unspaced Chinese text. Child participants were in the third grade of primary school and adults were

undergraduate students at university. On the basis of the early local measures, they found that the word spacing manipulation had a greater beneficial effect for children than adults. They computed the total sentence reading times, and there was no reliable difference in term of reading time for all participants.

On the other hand, Bai et al. (2008), in their study with Chinese native speakers, found that the presence of space between the words in the sentence or marking as a clue of separating the words neither prevented nor facilitated the reading process.

In the study of Hsu and Huang (2000), Chinese native speakers' reading time and comprehension rate of Chinese sentences with and without spaces between words were measured. It was understood that spaces between words accelerated reading but did not have any effect on reading accuracy.

On the other hand, in another study conducted on Chinese native speakers, it was understood that adding spaces between words in Chinese texts had no facilitating effect on native speakers (Li et. al., 2010, 1381). Researchers think that this may be due to the fact that native Chinese readers are familiar with the absence of spaces between the words or with the graded reading method restraining their reading process.

Although the studies on this topic are highly diversified, the ones related to teaching Chinese as a foreign language are rather inconclusive. Bai et al. (2010) in his study consisting of two different eye-tracking tests conducted with Chinese language learners, who are native English speakers, found out that adding spaces between words had a positive function in reading comprehension rate.

Similarly, Shen et al. (2012) in his study conducted with students from four different countries concluded that the presence of spaces in Chinese texts facilitated the reading of second language learners, regardless of alphabetic status and word spacing in their native language.

2. Method

2.1. Participants

Twenty-seven undergraduate students participated in the experiment. They were all learning Chinese as a foreign language who were grouped in two levels (elementary and advanced) by the researchers. Ten students were at elementary level and seventeen students were classified as advanced level students. Elementary level students were in their second year at the Sinology department of a university and had never been to China. Advanced level students were seniors at the same department and had studied in China for at least 1 year. Participants were told that they were going to read and understand sentences presented under two different spacing conditions. They were assured that there was not any reading aloud task (as most of them were reluctant to participate if there were any).

2.2. Materials and Design

A total of 30 sentences were constructed. All the sentences consisted of 7 areas of interest (AOI) [subject +adverb of time + function word (preposition of place/在) + place name + verb + function word (tense suffix/了) + object] and had five free morphemes and two bound morphemes (as shown in figure 1). Each sentence was presented to the participants twice, once with spaces between the words and once in the conventional form of Chinese, in a random order. The sentences were ranged randomly by a computer so the participant could not have an educated guess of what kind of stimuli he/she should see next. Two practice sentences, one for each spacing condition, were included at the beginning of the first session. After the first 30 sentences there was a 5-minute break. In total each participant read 60 sentences. After each of these sentences, a *wh-question* (who, where, when and what) was presented to test the comprehension. Participants were asked to choose the right answer from the two options by using the mouse as a pointer.

Table 1. An example of one of the sentences used in the experiment.

(1) Word spaced condition						
经理	上午	在	展会	介绍	了	新品。
Jīnglǐ	shàngwǔ	zài	zhǎnhuì	jièshào	le	xīnpǐn.
AOI-1	AOI-2	AOI-3	AOI-4	AOI-5	AOI-6	AOI-7
Manager	morning	at	fair	introduce	pst	new product
‘At the fair, the manager introduced a new product in the morning.’						
(2) Unspaced condition						
经理上午在展会介绍了新品。						

2.3. Apparatus

Participants’ eye movements were recorded using the SMI RED 500 eye tracking system. The stimuli were presented on a 22-in. (55.8-cm) DELL monitor with a 1689 × 1050 pixel resolution. Participants were seated at a distance of 70 centimetres from the computer screen where the stimuli were presented, and their eyes were fixed with the jaw stabilizer. The stimuli were presented with Song font type in 36 font size. Prior to the start of each session, a five points calibration was completed.

2.4. Procedure

The eye tracking procedure was explained to the participants and it was emphasized that they should keep their heads still. The participants were tested individually. Participants were informed that they would read sentences under different spacing conditions. They were told to read the sentences silently and press the button to see the following comprehension question. In total the experiment took approximately 30 minutes. The eye tracker was placed in a sound proof room of the linguistics laboratory which was dimly lit.

3. Results

The comprehension rate was 93% making it possible to conclude that participants read and understood the sentences. The obtained data were analysed both globally and locally by the Linear Mixed Effect Model (LME). We used *lme4* package in R (R Core Team, 2013) by using *lmer* function for all fixation duration measures and *glmer* (Bates, Maechler &Bolker, 2013) function for number of fixation measure to fit generalized linear mixed-effects regression models (Baayen, 2008), with Condition factor (spaced, unspaced) and Language Level (elementary, advanced as fixed factors. In addition to fixed factors considered in simple linear regressions, LME models account for random variation induced by items and participants.

All data points above or below twice the standard deviations from the mean were excluded from the fixation duration data.

For global eye tracking measurements, total reading time, average fixation duration and total number of fixations were computed.

Total reading time: Reading times were shorter for spaced writing (M=3760 ms, SD=560) and reliably longer for unspaced writing condition (M=3833 ms, SD=463). But this difference was not reliably different (ps >.05).

Average fixation duration: There was no reliable effect of neither the writing condition (spaced/unspaced) nor the language level (elementary/advanced) (see table 1). That is to say, participants found neither of the conditions easier to read.

Table 2. Average fixation duration ANOVA Analysis for Two Conditions and Two Levels

Avarage fixation duration			
	β(SD)	**t**[1]	**p**
(Intercept)	5.579 (0.016)	344.25	<0.001
Condition	-0.031 (0.017)	-1.87	0.06
Level	-0.01 (0.032)	-0.30	0.76

Total number of fixation: Although it is anticipated that the total reading time and the number of fixations are highly correlated, in this case number of fixation showed reliable difference both in condition and language level variables (see table 2) whereas there was none in total reading time. Number of fixation in spaced writing and elementary language level were higher.

Table 3. Total number of fixation Non- Parametric Analysis for Two Conditions and Two Levels

	Total number of fixation		
	β(SD)	**z**	**p**
(Intercept)	2.169 (0.077)	28.14	<0.001
Condition	0.131 (0.028)	4.63	<0.001
Level	-0.756 (0.152)	-4.97	<0.001

1 |z| and |t| > 2.0, statistically significant (Gelman & Hill, 2007).

In addition to the global analyses, local analyses were also done. For these analyses first-pass duration, first fixation duration, second-pass duration and dwell time for each area of interest were computed.

First-pass Duration: First-pass duration time was statistically different only in AOI-3 and AOI-6 which are both bound morphemes. Figure 1 shows the first-pass duration and language level relation of spaced and unspaced sentences. In AOI-3 there is a statistically reliable difference concerning the level variable. Elementary level participants tend to have a longer first-pass duration.

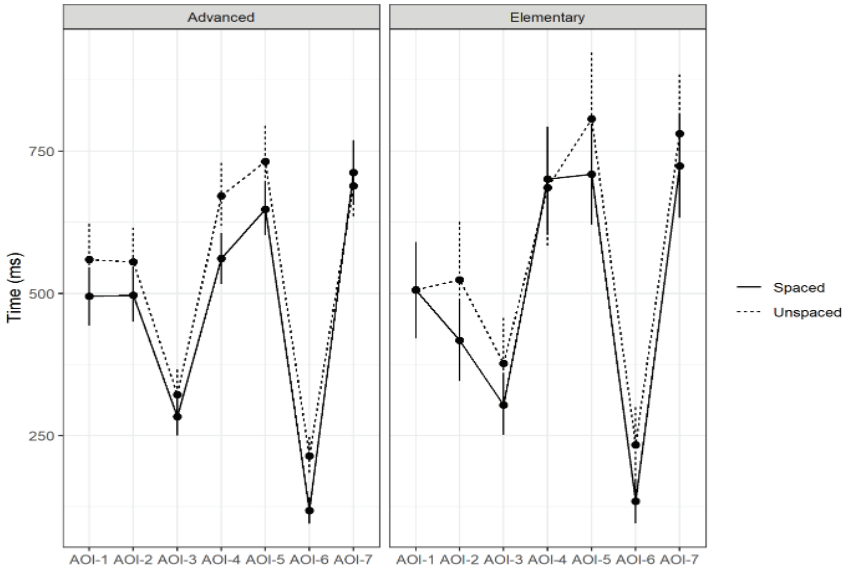

Figure 1. First-pass duration of participants (advanced and elementary language levels)

Concerning the first-pass duration, in AOI-6 there is a statistically reliable difference in the condition (spaced/unspaced) variable. In unspaced sentences this area of interest tends to have a longer first-pass duration (see figure 2).

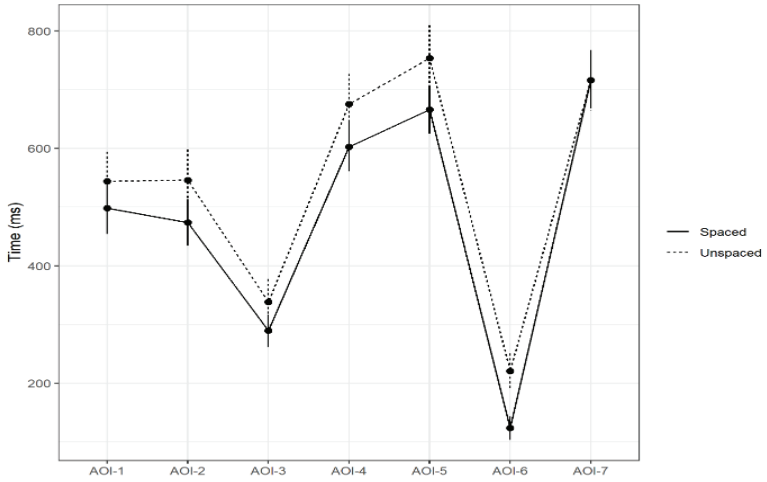

Figure 2. First-pass duration on two conditions (spaced and unspaced writing)

First Fixation: There was a statistically significant difference in AOI-6 in both condition (spaced/unspaced) and language level (elementary/advanced) variables concerning the first fixation measures (see figure 3 and figure 4).

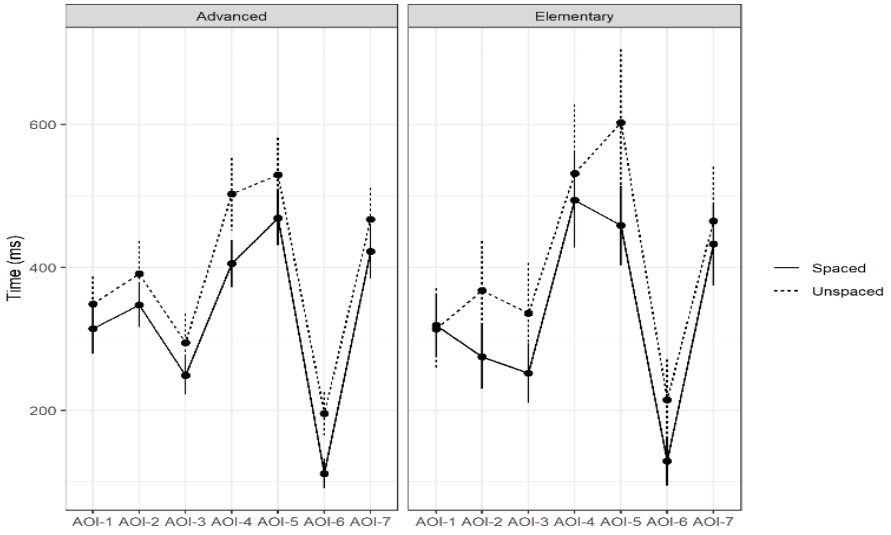

Figure 3. First fixation duration of participants (advanced and elementary language levels)

Also in AOI-3 there is a significant difference in condition (spaced/unspaced writing) where unspaced writing has a longer first fixation duration (see figure 4).

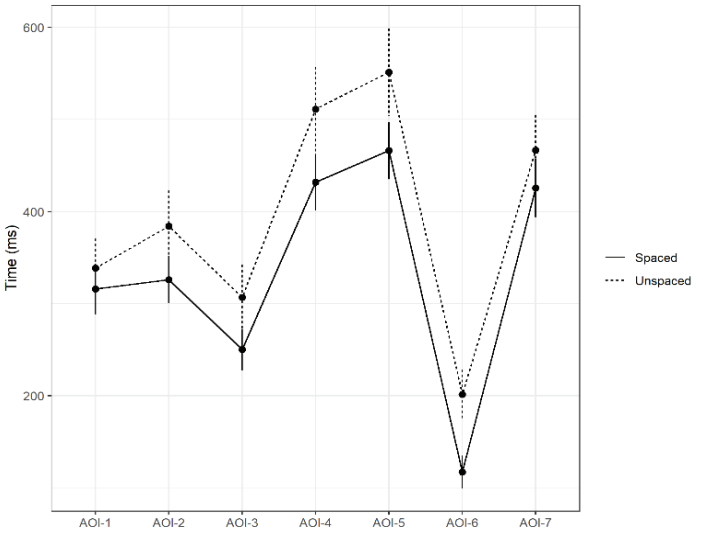

Figure 4. First fixation duration on two conditions (spaced and unspaced writing)

Second-pass Duration: In late processing (second-pass duration), like it was observed in the first-pass duration, again in AOI-3 there is a statistically reliable difference concerning the level variable. Elementary level participants tend to have a longer second-pass duration too.

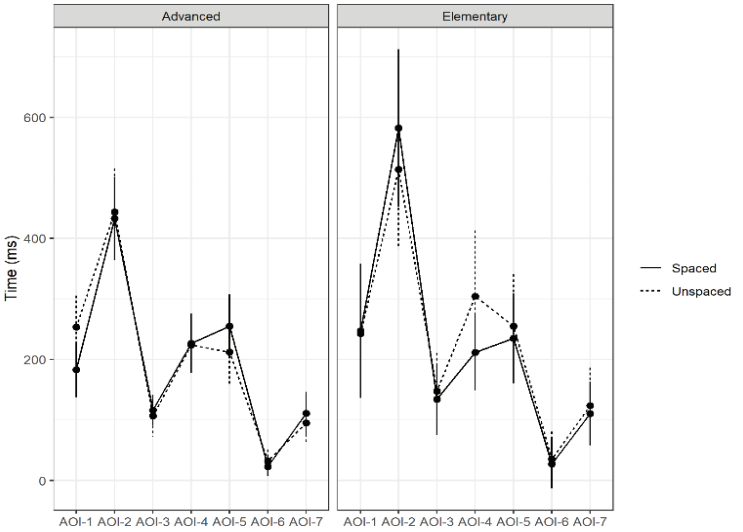

Figure 5. Second-pass duration of participants (advanced and elementary language levels)

Concerning the second-pass duration, in AOI-4 there is a statistically reliable difference in the condition (spaced/unspaced) variable. In unspaced sentences this area of interest tends to have a longer second-pass duration (see figure 6).

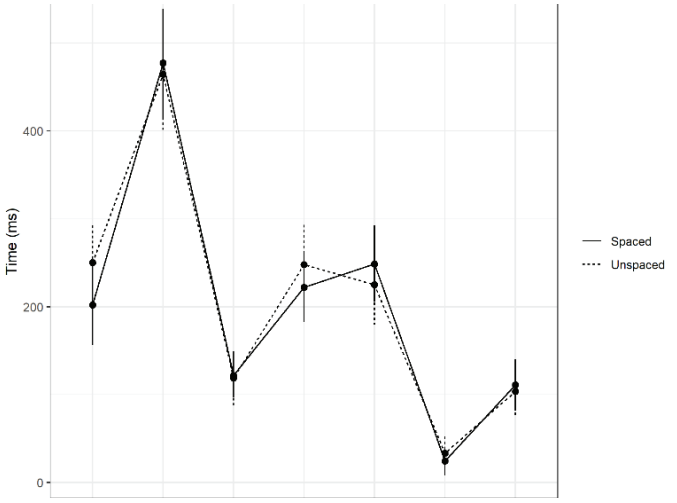

Figure 6. Second-pass duration of participants (advanced and elementary language levels)

Dwell Time. There was a statistically significant difference in AOI-3, AOI-4 and AOI-6 in condition (spaced/unspaced) and a statistically significant difference in AOI-6 in language level (elementary/advanced) variables concerning the dwell time (see figure 7 and figure 8).

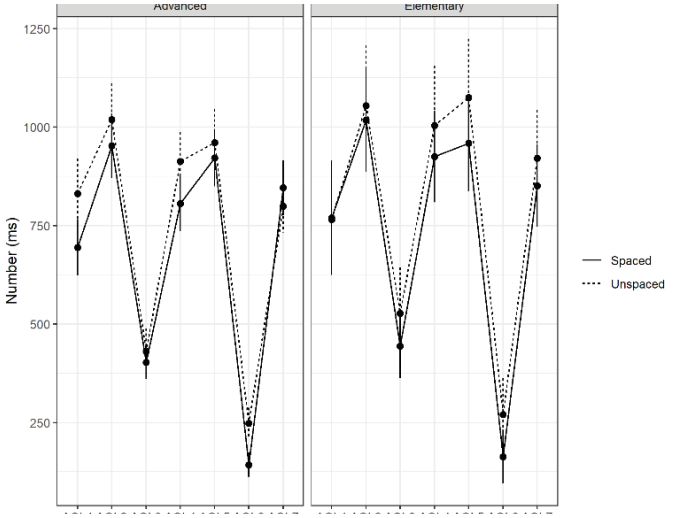

Figure 7. Dwell time of participants (advanced and elementary language levels)

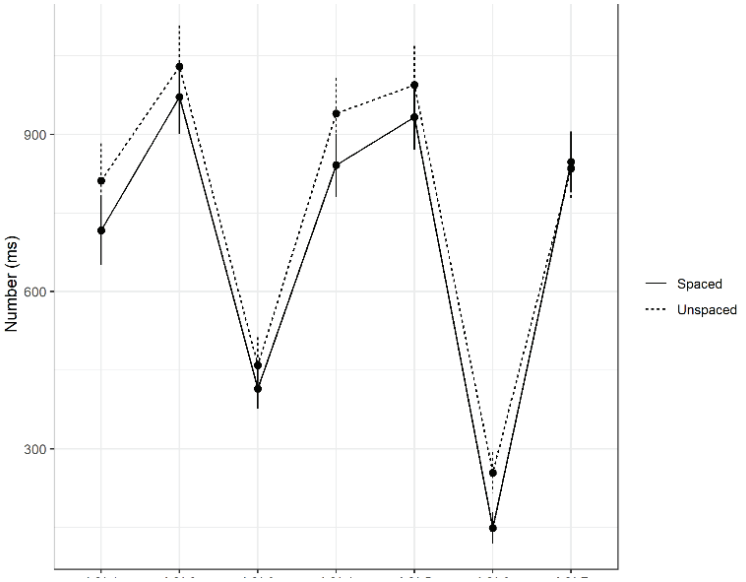

Figure 8. Dwell time of participants (advanced and elementary language levels)

4. Conclusion and Discussion

Our findings show that the total reading time for spaced writing and unspaced writing condition was not reliably different. This result is similar to Inhoff et al.'s (1997) study where they examined how inserting spaces between words in Chinese influenced reading. Their study showed no reliable differences in total reading times for any of the presentation conditions. In another study, Bai et al. (2008) investigated the influence of spacing information on eye movement behaviour during Chinese reading where total reading times for text presented under nonword and single character spacing conditions were not reliably different too. But in contrast, in Shen et al.'s (2012) study, reading times were shortest for word-spaced text and reliably longer for normal unspaced and character-spaced text and reliably longer again for nonword-spaced text. Their study indicated that word-spaced text was easiest for non-native Chinese readers to process and even easier to process than normal unspaced text.

As for the average fixation duration there was also no reliable effect of neither the writing condition (spaced/unspaced) nor the language level (elementary/advanced). In Bai et al.'s (2008, 6) study, for average fixation duration there was a signification effect of presentation condition. Average fixation durations were longer under normal spacing conditions than under single character, word spacing, and nonword spacing conditions. Also, average fixation durations were longer under word and non-word spacing conditions than under single spacing conditions. Finally, average fixation durations did not differ between word and nonword spacing conditions.

As mentioned in the results, the total number of fixation showed reliable difference both in condition and language level variables. Number of fixation in spaced writing and elementary

language level were higher. Presumably, it is because the participants need more time to process what they read, during the early stages of learning Chinese.

Concerning the local analyses there is a statistically significant difference between spaced and unspaced words in the early processing (first-pass duration and first fixation duration) of AOI-3 (在zài - function word) and AOI-6 (了le - function word). These two area of interests are both functional words (dependent form units). On the other hand in late processing (second-pass duration), only in AOI-4 (展会zhǎnhuì - place word) there is a significant difference.

The difference in the early processing of AOI-3 (在) and late processing for the area of interest 4 (展会), can be evaluated together. In Chinese, the function word '在' can add different meanings to sentence according to the word it is used together with. The reason why the participants focus on this preposition in the first fixation duration may be that they are trying to understand the correct function of '在' in the sentence. Because spaced or unspaced writing of the '在' can change the meaning of the sentence the place name and '在' may have attracted the attention of the students in early and late processing. If '在' has location meaning, it is read with the place word (see 1).

(1) 她在学校。
Tā zài xuéxiào.
'She is at school.'

'在' can also be used as auxiliary verb to express that an action is ongoing or in progress (see 2). This is the equivalent of present continuous in English.

(2) 我在看书。
Wǒ zài kànshū.
'I am reading book.'

Both early processing (first-pass duration and first fixation duration) and dwell time calculations showed that, there is a significant difference between the conditions of spaced and unspaced writing for the function word '了'. The meaning of '了', depending on whether it is spaced or unspaced in a sentence, causes uncertainty in meaning. As shown in figure 2 the function word '了' has several different meanings and usage.

Table 4. Examples of the function word '了'in different meanings and usage.

Past Tense Meaning: 衣服已经干了。(Yīfú yǐjīng gànle.) Dress is already dried.
Condition Change Meaning: 我有女朋友了。(Wǒ yǒu nǚ péngyǒule.) I have a girlfriend. (did not have before) 下雨了。(Xià yǔle.) It's raining. (it suddenly started)
'Almost' Meaning: 快到了。(Kuài dàole.) We're almost there. 我们就要毕业了。(Wǒmen jiù yào bìyèle.) We're about to graduate.
Continuity Meaning: 我在这儿住了三年了。(Wǒ zài zhè'er zhùle sān niánle.) I've been living here for three years. (I still live here)
Exaggeration: 太难了。(Tài nánle.) Very hard! 太可爱了。 (Tài kě'àile.) Very sweet!

As it can be seen from the examples above, functional words in Chinese can change the meaning of the sentences. For this reason, students, whose native language is an alphabetic language, can have difficulty in understanding and using the functional words '了' and '在' because of their complex usage and meaning features. According to our study results, we can say that the students focused on functional words in order to understand the meaning of the sentence. Because spaced or unspaced writing of functional words is very important to determine the meaning of the sentence. Participants may find it difficult to understand the function of the particle '了', they try to understand sentences correctly. The particle '了'is often followed by a verb to indicate various additional meanings. In Chinese, the aspect and the time of an action are not entirely expressed in one grammatical form, the aspectual particle '了' is used to indicate the completion of an action but it doesn't necessary show that the action took place in the past (Li, 2009).

Concerning the second-pass duration, in unspaced sentences AOI-4 tends to have a longer duration. As mentioned before participants were instructed to answer a wh- question for testing the comprehension and one of these questions was "where" so it may be assumed that they try to find answer the question to this question correctly by having a longer duration in this area of interest which represents the place where the action takes place.

Statistically significant difference of dwell time in AOI-3, AOI-4 and AOI-6 in condition (spaced/unspaced) and a statistically significant difference in AOI-6 in language level (elementary/advanced) variables concerning the dwell time is foreseen as there were differences in all other measurements in these area of interests.

To conclude, this study shows a minor difference in processing the spaced and unspaced sentences, by participants who are learning Chinese as a foreign language. The only difference occurs when it comes to processing the bound morphemes which tend to take more time and afford.

References/Kaynakça

Bai, Xuejun, Yan Guoli, Zang Chuanli, Liversedge Simon P. and Rayner Keith, "Reading spaced and unspaced Chinese text: Evidence from eye movement", *Journal of Experimential Psychology: Human Perception and Performance,* 34/5 (2008), p. 1277-1287.

Bai, Xuejun 白学军, Zhang Tao张涛, Tian Lijuan 田丽娟, Liang Feifei 梁菲菲 and Wang Tianlin 王天林, "Ci qie fen dui meiguo liuxuesheng hanyu yuedu yingxiang de yan dong yanjiu词切分对美国留学生汉语阅读影响的眼动研究 (Eye Movement Research on the influence of word segmentation on American Students' Chinese Reading)", *Psychological Research*, 3/5 (2010), p. 25-30.

Baayen, R.H. Analyzing Linguistic Data: A Practical Introduction to Statistics. Cambridge: Cambridge University Press. (2008), doi: 10.1017/CBO9780511801686.

Bates, D., Maechler, M., and Bolker, B. "lme4: Linear-mixed Effects Models Using S4 Classes", (2013), http://CRAN.R-project.org/package=lme4 (R package version 0.999999-2).

Blythe, I. Hazel, Liang Feifei, Zang Chuanli, Wang Jingxin, Yan Guoli and Bai Xuejun, "Inserting spaces into Chinese text helps readers to learn new words: An eye movement study", *Journal of Memory and Language*, 67 (2012), p. 241-254.

Gan, Kok Wee, "Integrating word boundary identification with sentences understanding", *31st Annual Meeting of the Association for Computational Linguistics,* (1993), p. 301-303.

Gelman, A. & Hill, J. Data analysis using regression and hierarchical/multilevel models. New York, NY: Cambridge, 2007.

Hsu, Sheng-Hsiung and Huang Kuo-Chen, "Interword spacing in Chinese text layout" *Perceptual and Motor Skills*, 91/2 (2000), p. 355-365.

Inhoff, A. W., Liu, W., Wang, J. and Fu, D. J. Use of Spatial Information During the Reading of Chinese Text", *In Cognitive research on Chinese Language*. Jinan, China: Shan Dong Education Publishing, 1997.

Li, Dejin 李德津 and Cheng Meizhen 程美珍*A Practical Chinese Grammar for Foreigners*外国人实用汉语语法Waiguo ren shiyong hanyu yufa. Beijing: Beijing Language and Culture University Press, 2009.

Li, Xingshan, Liu Pingping and Keith Rayner, "Eye Movement guidance in Chinese reading: Is there a preferred viewing location? *Vision Research*, (2011), p. 1146.

Li, Xin 李馨, Bai Xuejun 白学军, Yan Guoli 闫国利, Zang Chuanli 藏传丽 and Liang Feifei梁菲菲, "Kongge zai wenben yuedu zhong zuoyong 空格在文本阅读中作用 (Function of the space form on text reading)", *Advances in Pyschological Science*, 18/9 (2010), p. 1377-1385.

R Core Team "R: A language and environment for statistical computing. R Foundation for Statistical Computing", Vienna, Austria, (2012), ISBN 3-900051-07-0, URL http://www.R-project.org/.

Shen, Deli, Liversedge Simon P., Tian Jin., Zang Cuanli., Cui Lei., Bai Xuejun., Yan Guoli and Rayner Keith, "Eye Movement of Second Language Learners When Reading Spaced and Unspaced Chinese Text", *Journal of Experimental Psychology*, 18/2 (2012), p. 192-202.

Yen, Miao-Hsuan, Tsai Jie-Li, Tzheng Ovid J.l. a Huang Daisy L., "Eye movements and patafoveal word processing in reading Chinese, *Memory & Cognition*, (2008), p. 1033-1045.

Zang, Chuanli, Liang Feifei, Yan Guoli, Liversedge Simon, "Interword Spacing and Landing Position Effects During Chinese Reading in Children and Adults", *Journal of Experimental Psychology: Human Perception and Performance.* 39/3 (2013), p. 720–734.