Word knowledge influences character perception

Xingshan Li

Key Laboratory of Behavioral Science,

Institute of Psychology, Chinese Academy of Sciences

Alexander Pollatsek

Department of Psychology, University of Massachusetts Amherst

Correspondence to:

Xingshan Li

Institute of Psychology, Chinese Academy of Sciences, China

Email: lixs@psych.ac.cn

Abstract

In two experiments, we examined whether context information can affect the activity of the nodes at the character level. Chinese readers viewed two Chinese characters; one was intact, but the other (the target) was embedded in a rectangle of visual noise and increased in visibility over time. The two characters constituted a word in one condition but did not in the other condition. The task was to press a button to indicate whether the character in the noise was at the top or bottom of the rectangle. (They did not have to identify the character.) Response times were faster in the word condition than in the nonword condition. As the "wordness" of the stimulus was logically irrelevant to judging the location of the target character, the data indicate that processing at the word level can feed back to fairly low level judgments such as where a character is.

Models of word processing in reading usually assume that the processing has multiple levels: a feature level, a character (or letter) level, and a word level (McClelland & Rumelhart, 1981; Massaro, 1998). In all such models, the feature level encodes visual feature information which feeds into character (letter) identification, which in turn, feeds into word identification. However, word superiority effect (WSE) studies demonstrated that character identification is facilitated (even when guessing is controlled for) when the character is part of a word (Reicher, 1969; Wheeler, 1970) compared to when it is embedded in a series of nonword letters or when it is shown in isolation. Such a phenomenon clearly casts doubt on whether word recognition is merely the result of the feed-forward mechanism sketched above.

Perhaps the most influential current model of word recognition is the Interactive Activation Model (IAM, McClelland & Rumelhart, 1981) which assumes that word processing is an interactive process. (The IAM model was largely proposed to explain phenomena such as the WSE.) Through the interconnections between the nodes in a three-level network (feature level, character level and word level), the activity of a node can affect the activity of the nodes in the same level and the other levels including "lower" levels. Most notably, to explain the word superiority effect, IAM assumes that the activity of the nodes at the character level is affected by activity of nodes at the word level. The activity of the character nodes increases faster when the character is part of a word (or even part of a pseudoword) than when it is not. Hence, a character belonging to a word is identified faster than when it is not part of word or when it is shown in isolation.

However, not all of the models of word processing that can explain the WSE assume that context affects the activity in the character level. For example, the Fuzzy Logic Model of Perception (FLMP, Massaro, 1998) assumes that context information in the word level does not

feedback to the character level and affects the activity of related nodes. Instead, FLMP assumes that sensory information (feature level) and context information (word level) are integrated at the decision making stage using Fuzzy Logic (Zadeh, 1965). Since the decision making stage receives two sources of information (sensory information and context information) when the letter is shown in a word, but only one source of information (sensory information) when it is not, decisions about what letter was present in a given location are better when the letter is shown in the context of a word than when it is not. That is, contrary to IAM, FLMP predicts that context information does not affect the activity of the nodes at character level.

These two models share a multiple-level processing view but differ in whether context information feeds back to the character level and affects the activity at that level. Since both models can explain the word superiority effect (McClelland & Rumelhart, 1981; Massaro, 1998), telling the models apart using quantitative modeling is likely to be difficult, as the goodness of fit may depend on auxiliary assumptions rather than the basic structure of the model. Standard letter identification paradigms are also not well suited for testing whether context information affects activity at the character level because the context manipulation is not logically irrelevant to the task. That is, as indicated above, the improved performance in the word context condition could be due either to increased activity at the character level or to the context improving the chances of selecting the correct identity of the character from partial information.

There are two recent studies that could be viewed as evidence against Massaro's model (Barnhart & Goldinger, 2010; Norris, Butterfield, McQueen & Cutler, 2006). In Norris, et al., the key manipulation was employing ambiguous stimuli that were halfway between 'H' and 'N' in a word context that supported one or the other interpretation (e.g., *REIG?*, where '?' represents the ambiguous letter). Their key finding was that people were more likely to classify the ambiguous

letter consistent with the word context. However, this finding could simply reflect that the decision about the letter identity reflected a mixture of the evidence from the word level and the evidence from lower levels using Fuzzy Logic. Barnhart and Goldinger investigated the effects of various variables affecting word perception (e.g., spelling-to-sound consistency) on naming handwritten words and found that these effects were bigger for handwritten words than for printed words. Although their conclusions were not stated in terms of a refutation of the Fuzzy Logic model, they do suggest that word-level information was feeding back to letter-level identification. However, it is not clear whether some version of the Fuzzy logic model could explain these results as well: if the computation of the name of the word comes from both information at the word and letter level, and the information at the letter level is more ambiguous for handwriting, one might expect that effects such as spelling-to-sound consistency (which are presumably occurring beyond the letter-identity level) will be magnified for handwritten words.

The present experiments take a different approach to testing whether word-level information feeds back to letter-level information. They employed an object localization task in which participants did not need to identify the character to do the task. Thus, sensory information is necessary for the task, but the context information is quite removed from the response decision. Participants saw two Chinese characters, which constituted a word in the word condition, but did not constitute a word in the non-word condition. The two characters were to the left and right of a fixation cross. One, the *context character,* was fully visible. The other, the *target character*, was presented in a rectangle of visual noise and it appeared either at the top or bottom of this rectangle. The participants' task was to decide which location ('top' or 'bottom') the target character was in. Therefore, knowing whether the two characters formed a word was irrelevant to making the judgment of 'top' or 'bottom'. Thus, if feedback from the word level to the character

level is occurring which actually improves the perceptibility of the target character (as a model such as the IAM predicts), the location of the target should be detected better when the two characters form a word. In contrast, if word context only increases the probability that the participant will be able to make the correct decision about which character was present (as a model such as the FLMP predicts), it should have no effect on performance in the location task.

To make a difference in performance between word and nonword conditions more likely to be detected, the task was made difficult by putting the target character in a rectangular background of dynamic noise (see Figure 1). The average intensity of the noise was constant throughout a trial, but the contrast of the character increased gradually. Increasing the contrast of the character gradually made sure that the task was not easy at the beginning of a trial but it was simple enough at the end of a trial to guarantee that participants could do the task without error. Another reason that we embedded the target character in noise is that it made the task difficult to do if it was solely based on intensity information.

Insert Figure 1 about here

Another difference between the paradigm used in the current experiments and the traditional word superiority effect experiments was that the stimuli were not shown briefly. Instead, participants could control the duration of the stimulus presentation so that the accuracy was very high. By measuring the time it took to make the responses, we could measure the efficiency of character detection. Hence, participants could make their response as soon as they perceived some features of a character (such as some radicals of the characters) instead of recognizing the character.

In this study, we used Chinese characters for the following reasons. First, there are many two-character Chinese words. Among the 56,008 words that are included in one published source

(*Lexicon of common words in contemporary Chinese*, 2009), 72% are 2-character words. Second, the direction of reading is relatively flexible in Chinese. Traditionally, Chinese text was written from top–to-bottom, and the columns went from right-to-left. This tradition is still kept in calligraphy. Modern Chinese text is usually written from left-to-right, and with rows going from top–to-bottom, but some printed materials still use the top to bottom order for characters. In the current study, the two characters were arranged left-to-right, but there was a horizontal displacement between the characters.

Experiment 1

Method

  Participants. Forty undergraduate students (26 women and 14 men) from China Agricultural University with normal or corrected-to-normal vision were paid to participate in the experiment. The participants' ages ranged from 19 to 25 years (Mean= 21.4 years).

  Apparatus. Stimulus presentation and response registration were controlled by a personal computer. Stimuli were presented on a 15-inch LCD monitor with a resolution of 1280 x 800 pixels and a refresh rate of 60 Hz. Participants viewed the stimuli about 70cm from the monitor.

  Stimuli. Two characters were presented in each trial: a context character and a target character (see Figure 1). The context character was shown in black (RGB:0,0,0) on a gray background (RGB:128,128,128). The target character was shown in a background of dynamic noise. Dynamic Gaussian white noise of mean 0 and variance 1[1], which was updated each frame, was added to a rectangular window (3.3ºX6.5º) that covered the target character (see Figure 1). The contrast of the target character increased gradually from 0 to as large as 20%, which took 200 frames (about 3300ms) to reach the maximal contrast. A red cross was shown at the center of the screen throughout the trial (RGB:255,0,0; size 0.4º x 0.4º, line width 0.1º). The size of the

target character was 2.2º x 2.2º, which was large enough so that participants could detect the location of the target when they fixated on the cross. On half of the trials, the context character was on the left of the fixation cross, and the target character was on the right of the fixation cross, whereas on the other half of trials, the positioning was the opposite (see Figure 1). The target character was equally likely to be located in one of two locations, either above or below the fixation cross.

There were 16 practice trials and 320 experimental trials. On half of trials (both the practice trials and experimental trials) the context character and the target character constituted a word (the word condition), whereas on the other half of the trials, the two characters did not constitute a word (the nonword condition). The words were medium frequency, ranging from 30 to 100 per 1,116,417 words[2] (mean=53). The target character frequencies and character complexities did not differ significantly across conditions (Fs<1). The average target character frequencies were 950 (SD=669）and 923 (SD=653) occurrences per million and the average target character complexities were 8.19 (SD=2.00) and 8.24 (SD=1.97) strokes for the nonword and word conditions, respectively. The cue character frequencies and complexities did not differ significantly across conditions either (Fs<1). The average cue character frequencies were 939 (SD=634) and 886 (SD=616) occurrences per million and the average cue character complexities were 8.15 (SD=1.94) and 8.03 (SD=1.81) strokes for the nonword and word conditions, respectively. Participants were not told that half the stimuli were words and half were nonwords.

Task and responses. Participants were asked to press the "↑" key on a standard keyboard if the character that appeared in the noise rectangle was at the top half of the rectangle and press the "↓" key if the target was at the bottom half. (They did not need to identify the target character.) Participants were instructed to respond as quickly and as accurately as possible.

Procedure. At the beginning of a trial, a red cross was shown at the center of the screen. Participants were instructed to fixate on the fixation cross throughout the trial. They pressed the space bar on the keyboard to initiate a trial, at which point the two characters and the noise rectangle appeared. As indicated above, at the beginning of the trial, the contrast for the target character was 0. Stimuli disappeared as soon as participants made the response. There were 16 practice trials at the beginning of the experiment, followed by 320 experimental trials. There was a break after the practice trials and after every 80 experimental trials. For the practice trials, the feedback "incorrect" was presented for 500ms if a participant made an error in that trial. Participants were asked to make a response as quickly and as accurately as possible.

Results and Discussion

Accuracy. Accuracy was 99.5%, indicating that participants could do the task very well.

Reaction time. Incorrect trials and any trials with a reaction time (RT) shorter than 100ms or when the RT was over three SDs from the mean RT on that condition were excluded from the RT analysis. In total, 1.6% of trials were excluded from the RT analysis.

If word knowledge could affect the activity in the character level or lower level, we would expect that the responses would be faster in the word condition than in the non-word condition. The results were consistent with this prediction (see Table 1). The mean RT was 11.5ms (SE=5.7ms) longer in the non-word condition than in the word condition. A repeated-measures 2 (target location: left or right) x 2 (wordness: word or nonword) Analysis of Variance (ANOVA) indicated that there was a main effect of wordness, $F(1, 39)=4.02$, $MSE=.001$, $p=.05$, $\eta_p^2 =.09$, but perhaps more importantly, a significant interaction between target location and wordness, $F(1, 39)=5.34$ $MSE=.001$, $p<.05$, $\eta_p^2=.12$. When the target character was on the right, RT was 23.1ms longer (SE=8.6ms) in the nonword condition than in the word condition,

$F(1,39)=10.58$, $MSE=0.001$, $p<.01$, $\eta_p^2=.21$, whereas the RTs were identical when the target was on the left (difference=0.0ms, SE= 6.5ms; $F<1$). The main effect of left vs. right (Left: 2191ms, Right 2194ms) was also not significant ($F<1$). The shorter mean RT in the word condition than in the nonword condition when the target was on the right indicated that word knowledge could affect processing at the character level or possibly even at a lower (featural) level.

Insert Table 1 about here

Experiment 2

One possibility for the interaction is that the location of the two characters was the cause, most plausibly which visual field the characters fell in. Some studies have shown that characters falling on the left visual field and the right visual field are processed by different hemisphere of the brain. Moreover, other studies suggest that the two hemispheres use different strategies in lexical processing (Zaidel, Clarke, & Suyenobu, 1990). (However, it is not clear how these differences would produce our observed interaction.) The other possibility is that there is greater facilitation from the first character of a 2-character Chinese word to the second character than from the second character to the first. Experiment 2 was designed to distinguish between these two possibilities by presenting the two characters in the word condition in the opposite order from which they standardly appear in print (i.e., the first character on the right and the second character on the left). If the effect observed in Experiment 1 was merely due to visual field differences, the pattern of results should be identical to that of Experiment 1. On the other hand, if the effects were due to greater facilitation from the first character to the second, the pattern should be opposite to that in Experiment 1. Another plausible pattern is that facilitation from the first character of a word to the second requires the two characters to be in the order usually seen in print. In this case, there should be no context effect observed in Experiment 2.

Method

 Participants. Thirty-nine undergraduate students (22 women and 19 men) from the same participant pool as that in Experiment 1 participated in Experiment 2. None of these participants had participated in Experiment 1. One participant was excluded from the RT analysis since his/her accuracy was low (85%).

 Apparatus. The apparatus was identical to that used in Experiment 1.

 Stimuli. The stimuli were identical to those in Experiment 1 except that the position of the characters was switched. That is to say, the second character in a word was shown on the left, and the first character was shown on the right. The positioning of the characters in the nonword condition was also reversed from that in Experiment 1.

 Tasks and procedure. The tasks and procedures were identical to that in Experiment 1. Again, participants were told nothing about the word-nonword manipulation.

Results and Discussion

 Accuracy. Accuracy was again very high (99.3%) except for the excluded participant.

 Reaction time. Incorrect trials and any trials with RTs shorter than 100ms or when the RT was greater than 3 SDs from the mean RT on that condition were excluded. In total, 1.3% of trials were excluded from the RT analysis.

 Consistent with Experiment 1, the mean RT was 13.7ms (SE=7.0ms) longer in the non-word condition than in the word condition (see Table 1). As in Experiment 1, there was a significant main effect of wordness in a repeated-measures 2 (target location: left or right) x 2 (wordness: word or nonword) ANOVA, $F(1, 37)=4.02$, $MSE =.001$, $p =.05$, and a significant interaction between target location and wordness, $F(1, 37)= 5.34$, $MSE=.001$, $p<.05$. The main effect of target location was again not significant (Left: 2384ms, Right: 2395ms; $F(1, 37)= 2.16$,

MSE=.002, *p*>.10). The pattern was the opposite of that in Experiment 1: the mean RTs were 25.0ms (SE=9.9ms) longer in the nonword condition than in the word condition (when the target character was on the left, $F(1,37)=11.87$, *MSE*=.001, *p*<.01, whereas there was only a 2.2ms (SE=7.5ms) difference between the mean RTs when the target character was on the right (F <1).

The results of Experiment 2 replicated the major findings from Experiment 1 and indicated that the word benefit, although not large, is quite reliable. The fact that the pattern of the effect was opposite from that of Experiment 1 indicated that the interaction was not due to the spatial location of the characters, but instead likely due to a "deeper" explanation relating to how the two characters of a Chinese word interact with each other.

We thought it was of interest to do a meta-analysis of the two experiments: the factors were 2 (wordness) x 2(target location: left or right) x 2 (Experiment). There was a main effect of wordness, $F(1, 76)=7.65$, *MSE*=.002, *p*<.001, η_p^2 =.09. More importantly, the three way interaction between experiment, target location and wordness, $F(1, 76)=9.65$, *MSE*=.001, *p*<.01, η_p^2 =.11, indicates that there was a highly significant difference in facilitation across the two experiments between when the context character was the first character of the word and when it was second. There was also a main effect of experiment, $F(1, 76)=13.92$, *MSE*=.22, *p*<.001, η_p^2 =.16. RTs were shorter in Experiment 1 (2193ms, SE=37ms) than in Experiment 2 (2390ms, SE=38ms). However, it is not clear whether this is merely due to differences in the participants sampled or to differences between the configurations of stimuli in the two experiments.

Although the effect size of the word benefit was small compared to the overall RTs, it was statistically significant in both experiments. Moreover, 28 out of 40 participants in Experiment 1, and 23 out of the 38 participants in Experiment 2 showed better word than nonword performance in the conditions where there was a difference in the means. A sign test

showed that the number of participants who showed a word benefit was significantly more than half in Experiment 1, $t(39)=2.76$, $p<0.01$, although this effect was not significant in Experiment 2, $t(37)=1.32$, $p=0.19$. We also computed the SD in RTs for each participant separately (pooling all the conditions). In Experiment 1, the mean over participants of these SDs was 308ms, and in Experiment 2, it was 334ms. This indicates that the RTs were reasonably closely clustered around a value slightly over 2 seconds.

Additional analysis. We were interested in examining whether the size of the word-nonword difference changed over the course of the experiment to determine whether using the context character to determine the target character's location benefited from practice. Accordingly, we analyzed performance as a function of trial block for both Experiments 1 and 2 (see Table 1). The complete data are presented, but our analyses will be only on conditions where there was a word-nonword difference in the main analyses. For both experiments, it is clear that the word-nonword difference varied across trial blocks with no pattern; moreover, the pattern was different in the two experiments. For Experiment 1, the main effect of word-nonword, and the linear trend of block, were both significant, $F(1,39)=7.10$, $p<.05$, $F(1,39)=5.38$, $p<.05$, respectively, but the interaction was not significant, $F(3,117)=1.75$, $p>.10$. For Experiment 2, the pattern was similar: a main effect of word-nonword, $F(1,37)=6.34$, $p<.05$, an almost significant linear trend of block, $F(1,37)=4.05$, $p=.051$, and a non-significant interaction, $F(3,111)<1$. Thus, the context effect was not something learned over the course of the experiment.

Discussion

In the location detection task we employed, there was an advantage of the word condition over the nonword condition when the target was the second character of a word, but not when the target was the first character of a word. This advantage was unaffected by whether the target was

presented in the LVF or the RVF and also unaffected by the spatial order in which the two characters of a word were presented. This indicates, contrary to a feed-forward model such as FLMP (Massaro, 1998), that word-level information is feeding back to enhance processing of the visual information of the target character. Our experiment can not discriminate between whether this feedback only goes back to the character level or actually feeds back through the character level to the level of visual features.

Both Experiments 1 and 2 indicated that this word-level information only had a facilitative effect if the context character was the first character of a two-character word and the target character was the second. However, what is a bit puzzling is why there was about as large an effect when the relative spatial locations of the characters were opposite to the usual (but not universal) left-to-right order of the two characters. We think the most plausible hypothesis is that the "correct" temporal order of the characters when the context character was the first character of the word overrode the "incorrect" spatial order in Experiment 2, especially because the context character was plausibly encoded over a second before the information from the target character was visible enough to allow a response. This, of course, does not necessarily mean that the characters of two-character Chinese words are processed in series in normal reading. However, the fact that there was no word advantage when the second character was the context character is consistent with a serial processing hypothesis.

On the other hand, this asymmetry between the first and second character of a word to produce the observed word advantage in the present experiments is also consistent with a parallel encoding of the two characters in normal reading. Although characters are a higher-order unit than letters in an alphabetic script, the data from English would be consistent with such a parallel model. First, there is abundant data indicating that, for shorter words at least, the letters in a word

are processed in parallel. For example, the data from the WSE experiments indicate that identification accuracy is as good for a letter in a 4-letter nonword is as for a letter in isolation (e.g., Reicher, 1969). Moreover, in spite of this parallel processing, there is an abundance of research indicating that the beginnings of English words play an especially important role in the early encoding processes of these words (Rayner, 1998; White, Johnson, Liversedge, & Rayner, 2008). This advantage makes sense as one would expect that, in all spoken languages, one would want as much information to be packed into the beginning of the utterance as possible. Obviously, however, future research is needed to further elucidate the details of the operation and limitations of this feedback mechanism.

In closing, we should note that prior studies have shown that words play an important role in the reading of Chinese sentences (Rayner, Li, Juhasz, & Yan, 2005) and that the context can also affect character identification (Li, Rayner, & Cave, 2009; Cheng, 1981). Thus, it is plausible that the findings from the current study could be extended to other languages such as English.

References

Barnhart, A. S., & Goldinger, S. D. (2010). Interpreting Chicken-Scratch: Lexical Access for Handwritten Words. *Journal of Experimental Psychology-Human Perception and Performance, 36*(4), 906-923.

Cheng, C. M. (1981). Perception of Chinese-characters. *Acta Psychologica Taiwanica, 23*(2), 137-153.

Li, X. S., Rayner, K., & Cave, K. R. (2009). On the segmentation of Chinese words during reading. *Cognitive Psychology, 58*(4), 525-552.

Massaro, D. W. (1998). *Perceiving talking faces: From speech perception to a behavioral principle*: MIT Press.

Massaro, D. W., & Cohen, M. M. (1991). Integration versus interactive activation: The joint influence of stimulus and context in perception. *Cognitive Psychology, 23*, 558-614.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception 1: An account of basic findings. *Psychological Review, 88*(5), 375-407.

Norris, D., Butterfield, S., McQueen, J. M., & Cutler, A. (2006). Lexically guided retuning of letter perception. *Quarterly Journal of Experimental Psychology, 59*(9), 1505-1515.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*, 372-422.

Rayner, K., Li, X. S., Juhasz, B. J., & Yan, G. L. (2005). The effect of word predictability on the eye movements of Chinese readers. *Psychonomic Bulletin & Review, 12*(6), 1089-1093.

Reicher, G. M. (1969). Perceptual recognition as a function of meaningfulness of stimulus material. *Journal of Experimental Psychology, 81*(2), 275-&.

Wheeler, D. D. (1970). Processes in word recognition. *Cognitive Psychology, 1*(1), 59-85.

White, S. J., Johnson, R. L., Liversedge, S. P., & Rayner, K. (2008). Eye movements when reading transposed text: The importance of word-beginning letters. *Journal of Experimental Psychology: Human Perception and Performance, 34*(5), 1261-1276.

Zadeh, L.A. (1965). Fuzzy sets, *Information and Control,* 8 (3), 338–353.

Zaidel, E., Clarke, J.M., & Suyenobu, B. (1990). Hemispheric independence: A paradigm case for cognitive neuroscience. In Scheibel, Arnold B.; Wechsler, Adam F. (1990). *Neurobiology of higher cognitive function. UCLA forum in medical sciences*, No. 29. (pp. 297-355). New York, NY, US: Guilford Press.

Footnotes

1. We used the 'imnoise' function in Matlab to add noise.

2. The source for word frequency was the Chinese Dictionary (National Languages Committee, 1997) with a corpus of 1,116,417 words.

Acknowledgments

Table 1

Examination of Context Effects Overall and as a Function of Trial Block

	Target Left			Target Right		
		Experiment 1				
	Nonword	Word	Difference	Nonword	Word	Difference
Overall	2.191	2.191	0.000	2.206	2.183	0.023
	(.035)	(.032)		(.034)	(.032)	
Block1	2.219	2.236	-0.016	2.251	2.229	0.023
	(.038)	(.033)		(.036)	(.031)	
Block2	2.212	2.196	0.031	2.196	2.183	0.013
	(.037)	(.035)		(.035)	(.033)	
Block3	2.186	2.214	-0.011	2.214	2.162	0.052
	(.039)	(.036)		(.040)	(.035)	
Block4	2.148	2.162	-0.003	2.162	2.159	0.003
	(.038)	(.040)		(.041)	(.040)	
		Experiment 2				
	Nonword	Word	Difference	Nonword	Word	Difference
Overall	2.397	2.372	0.025	2.396	2.394	0.002
	(.043)	(.040)		(.044)	(.041)	
Block1	2.415	2.393	0.023	2.411	2.444	-0.033
	(.043)	(.043)		(.045)	(.041)	
Block2	2.423	2.407	0.016	2.423	2.390	0.033
	(.047)	(.044)		(.044)	(.044)	
Block3	2.391	2.365	0.026	2.384	2.399	-0.015
	(.044)	(.042)		(.043)	(.043)	
Block4	2.366	2.324	0.042	2.371	2.345	0.026
	(.059)	(.047)		(.058)	(.050)	

(The entries in the cells are the response times in seconds. The numbers in parenthesis are the standard errors)

Figure Caption

Figure 1. Two examples of stimuli: Left panel (target on left); Right Panel (target on right). The

contrast shown in this figure is the maximum contrast in the experimental display.

Figure 1