

《心理学报》审稿意见与作者回应

题目：并列和偏正结构双字合成词的注视位置效应

作者：孟红霞 白学军 臧传丽 闫国利

第一轮修改说明

(修改日期：2012-12-31)

非常感谢《心理学报》两位专家的审稿意见。下面根据专家建议，逐一说明修改之处。

根据第一位审稿专家建议修改：

1. 引言部分

1) 引言提及本研究将解决“**什么因素**决定读者阅读过程的注视位置效应”的问题，在本研究中的结论部分并未给出回答。

谢谢专家的建议和提醒。

本研究主要解决双字合成词结构这一因素是否影响注视位置效应，根据本研究两个实验的结果，我们认为双字合成词结构不影响阅读过程中的注视位置效应。

2) 文章叙述了“中文阅读过程中“where”的决定是否以词为基础”，共包括三种不同的观点：①中文阅读过程中“where”的决定不是以词为基础；②中文阅读过程中眼跳目标的选择是以词为基础，而非汉字；③中文阅读过程中眼睛移向何处（where）的决定以“字”“词”相结合的方式为基础。建议作者在此论述基础上给出合理客观的评论，以及观点不相一致的原因所在，以帮助读者更加详细地了解。

非常感谢专家的建议和意见。具体修改内容如下：

上述三种观点不同的原因可能是：(1) 实验材料词边界的划分。Tsai 和 McConkie (2003) 以及 Yang 和 McConkie (1999) 的研究，没有严格控制实验材料词边界划分的一致性；而其他研究都对此变量进行了严格控制。(2) 阅读材料的呈现方式。Yang 和 McConkie (1999) 的实验材料中每个汉字的像素大小为 16×16 像素，词与词之间和词内汉字之间的空格大小为 8×8 像素，即空格为半个汉字的大小，所有实验句都采用上述格式，与中文读者正常阅读的格式完全不同。其他研究几乎都采用了正常的呈现方式。(3) 目标区域的分析。Li 等人 (2011) 研究中，分析四字词条件下的实验结果时，将四字词作为目标区域；但是分析双字词条件时，除了将双字词作为目标区域之外，还将双字词之后的两个汉字也作为分析的对象，使得双字词条件下可能包含 2 个或 3 个词，而四字词条件下只有四字词一个词。因此，中文阅读过程中，眼睛移向何处是以“词”还是“字”，亦或“字”“词”相结合的方式为基础，还需要更多的实验证据来验证。

3) “词的哪些特性影响中文阅读过程中的注视位置效应”，这一问题与本研究的实验目的、实验设计和实验材料的选择有何关系？建议补充说明，使研究逻辑更为清晰。

非常感谢专家的建议和意见。具体修改内容如下：

目前，较多研究支持中文阅读过程中眼跳目标的选择是以词为基础。那么词的哪些特性

影响中文阅读过程中的注视位置效应？汉语中词的特性包括低水平视觉因素（如笔画数和词长等）和高水平语言（认知）因素（如词频、预测性等）。是否与拼音文字相似，高水平语言（或认知）因素不影响注视位置效应？吴捷，刘志方和刘妮娜（2011）探讨了词频、语境的预测性及合理性对目标词首次注视位置的影响。结果发现，词频、语境的预测性和合理性均不影响读者对目标词的首次注视位置。但是，合理性影响读者对目标词的再注视概率，词频和语境的预测性不影响对目标词的跳读率和再注视率。郭晓峰（2012）考察了词频和语境的预测性对注视位置效应的影响，与吴捷等人的研究结果一致，发现词频和语境的预测性均不影响读者对目标词的首次注视的落点位置。但是，上述两项研究并不能全面回答“词的哪些特性影响中文阅读过程中的注视位置效应”这一问题。

中文书写文本系统中，大多数词是双字词（现代汉语频率词典，1986），构成双字词的两个汉字的字义与整个词的词义之间主要存在以下两种关系：一种是整个词的词义由两个汉字的字义共同决定，如“报刊”，另一种是整个词的词义主要由第二个汉字的字义决定，如“民警”（符淮青，2004；张良斌，2008）。并列结构和偏正结构的双字合成词正好符合上述两种关系。整词的词义与构成整词的汉字的字义之间的关系是否影响阅读中的注视位置效应。为了更好地回答上述两个问题，本研究选取了两种结构的双字合成词，即并列结构和偏正结构，进一步考察词的另一个高水平语言（或认知）因素——双字合成词结构，是否影响中文阅读过程中的注视位置效应。

4) “我们预测，中文阅读过程中存在偏向注视位置，具体为在单次注视中，读者的首次注视更多地落在词的中心位置；当首次注视落在词首位置时，再注视概率最高。”预测的理据何在？对研究结论进行逻辑假设将更为合理，不建议采用主观预测。

非常感谢专家的建议和提醒。具体修改内容如下：

根据前人的研究结果（白学军等，2012；白学军等，2011；Shu et al., 2011；Zang et al., 2012），在本研究条件下，中文阅读过程中可能存在偏向注视位置，具体为在单次注视中，读者的首次注视更多地落在词的中心位置；当首次注视落在词首位置时，再注视概率最高。

5) 为什么要采用双字合成词结构作为实验材料？“双字合成词结构是汉语词的高水平语言（认知）因素之一。”这一类型材料的特性是否会影响到实验结论？

非常感谢专家的建议。

首先，在汉语中，据统计双字词占 71%（现代汉语频率词典，1986）。双字词中绝大多数是双字合成词（钱玉莲，2008），因此，双字合成词是非常具有代表性的。

其次，汉语双字合成词最能够验证本研究所要探讨的问题，即中文阅读过程中的注视位置效应是否受整词的词义与构成整词的汉字的字义之间的关系影响。构成双字词的两个汉字的字义与整个词的词义之间主要存在以下两种关系：一种是整个词的词义由两个汉字的字义共同决定，另一种是整个词的词义主要由第二个汉字的字义决定。并列结构和偏正结构的双字合成词正好符合上述两种关系，因此为了考察这两种关系是否影响注视位置效应，本研究选择了并列和偏正结构的目标词作为实验材料。

2. 研究设计部分

1) 词切分一致性的控制与否十分重要，这是与前人研究结论相违背的关键，建议详细介绍词切分一致性的操作定义。

非常感谢审稿专家的提醒。

参考白学军等人(2012)和Yan等人(2009)的评定方法,对本实验材料的词切分一致性进行评定。第一步,根据《现代汉语词典》(2005)进行词切分,并编码;第二步,请未参加正式实验的20名大学生判断是否认可此种切分方式,若不认可,则标记出自己认为合理的切分方式,并进行编码;最后,计算两次编码的一致性百分比为95.8%,一致性范围在81%~100%之间。

2) 并列与偏正结构材料的尾字字频有无差异,并列与偏正结构的句子长度有无差异,需要提供检验结果。

非常感谢专家的建议。具体修改内容如下:

尾字字频结果:

对实验1中并列结构和偏正结构目标词的尾字字频进行统计分析发现,并列结构目标词的尾字字频, $M = 1.32$, $SD = 256.99$;偏正结构目标词的尾字字频, $M = 3.73$, $SD = 796.64$ 。对并列结构和偏正结构目标词的尾字字频进行 t 检验发现, $t(138) = -2.42$, $p = 0.017$,表明并列结构和偏正结构的尾字字频存在显著差异。

对实验2中并列结构和偏正结构目标词的尾字字频进行统计分析发现,并列结构目标词的尾字字频, $M = 1.14$, $SD = 250.98$;偏正结构目标词的尾字字频, $M = 3.25$, $SD = 755.79$ 。对并列结构和偏正结构目标词的尾字字频进行 t 检验发现, $t(138) = -2.22$, $p = 0.028$,表明并列结构和偏正结构的尾字字频存在显著差异。

但是,根据吴捷,刘志方和刘妮娜(2011)以及郭晓峰(2012)的结果,词频不影响读者对目标词的首次注视的落点位置,但是两项研究在词频影响读者对目标词的再注视概率方面存在差异,因此本研究中严格控制了词频这个无关变量。

根据Yan, Tian, Bai和Rayner(2006)的研究,在双字词识别过程中,当双字词为低频词时,首字字频影响读者对双字词的识别,具体表现为首字字频高的双字词的识别时间短于首字字频低的双字词。因此为了控制首字字频对实验结果的干扰,本研究严格控制了首字字频。但是根据Yan, Tian, Bai和Rayner(2006)的实验结果,尾字字频在任何情况下都不影响读者对双字词的识别过程,因此在本研究中并未控制尾字字频。

句子长度结果:

实验1中并列结构的句子长度, $M = 20.64$, $SD = 2.24$,偏正结构的句子长度, $M = 21.03$, $SD = 2.40$,对并列结构和偏正结构的句子长度进行了 t 检验,结果发现, $t(138) = -0.98$, $p = 0.327$;实验2中,并列结构的句子长度, $M = 20.96$, $SD = 2.13$,偏正结构的句子长度, $M = 21.26$, $SD = 2.42$,对并列结构和偏正结构的句子长度进行 t 检验发现, $t(138) = -0.78$, $p = 0.437$,表明两个实验中并列结构和偏正结构的句子长度之间均无显著差异。

3) 请在附录中呈现所有实验材料。

我们根据专家的建议在附录中提供了两个实验的所有实验材料。

4) 选取眼动指标的具体依据,每一眼动指标所代表的意义。

在选取眼动指标时,研究者主要参考的是前人的研究(白学军等,2011;Joseph et al., 2009; Li et al., 2010; Tsai & McConkie, 2003; Winkler et al., 2009; Yan et al., 2009; 臧传丽, 2010)。在前人研究的基础上,结合本研究的研究目的,选取了研究中提到的六个指标。

平均首次注视位置及首次注视位置分布,指的是读者首次加工目标词时,可能只需要一次注视就能识别该词,也可能需要两次和两次以上的注视才能识别该词。无论读者对目标词有一次注视还是两次和两次以上的注视,所有首次注视落点的平均位置,以及读者所有首次落在目标词四个区域的百分比的分布图。

单次注视中的平均首次注视位置及首次注视位置分布，指的是读者首次加工目标词时，只需要一次注视就能识别该词，在此种情况下所有注视落点的平均位置，以及所有注视落在四个区域中的百分比分布图。此指标表明当只有一次注视来识别目标词时，注视点落在目标词的哪个位置，表明此位置就是识别该词的最佳注视位置。

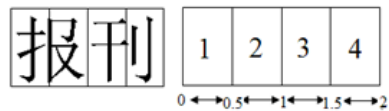
多次注视中的平均首次注视位置及首次注视位置分布，指的是读者首次加工目标词时，需要多次注视才能识别该词，在此种情况下所有落在目标词上的第一个注视点的平均位置，以及所有首次注视落在四个区域中的百分比分布图。此指标表明当需要多次注视识别目标词时，第一个注视点的落点位置。

向前注视的平均注视位置及注视位置分布，指的是首次加工目标词时，不管是单次注视还是多次注视，所有落在目标词上的所有注视点的位置的平均值。

首次注视位置上的再注视概率，指的是当首次注视落在词的某一个位置上时，读者对该词产生再注视的概率。此指标表明读者的首次注视落在词的哪个位置上时，只需要一次注视就能识别该词。

首次注视位置上的再注视位置，当首次注视落在词的某一个位置上时，如果发生了再注视，那么再注视落在词的哪个位置上。此指标表明读者的再注视模式是否有规律。

5) 对数据的收集介绍不清楚，注视位置等数值表示什么？有无单位？



如上图所示，本研究将一个双字词分为4个区域，一个区域代表半个汉字。数据收集过程中，眼动仪主要记录读者加工目标词时，每一个注视点的具体落点位置，是落在区域1、区域2、区域3还是区域4。注视位置等数值表示的是所有被试对每一种条件下所有目标词的注视的落点位置的平均值，单位为整字。

3. 讨论部分

1) 单次注视与多次注视的本质区别是什么？阅读者选择进行单词注视与多次注视到底基于什么？这是否也包含着情境、文本在加工过程中的作用体现。这与本研究所采用的实验材料有无关系？

谢谢专家的建议。

单次注视与多次注视的本质区别是，阅读过程中读者对词汇的识别过程客观上存在两种现象：一是读者在第一遍阅读中需要一次注视就能识别该词，即单次注视；二是读者需要两次及两次以上的注视才能识别该词，即多次注视。单次注视表明读者只需通过一次注视就能完成对该词的识别过程；多次注视表明读者在目标词的首次注视上并未提取到该词汇的全部信息，也就是并未完成词汇识别过程，因此需要再注视来完成对该词的识别过程。

阅读者选择进行单次注视与多次注视基于其首次注视的落点位置，如果首次注视落在最佳注视位置上，就很可能是单次注视；如果首次注视未落在最佳注视位置，则很可能是多次注视。

为了排除情境、文本在加工过程中的作用，在造句时我们保证了一组目标句中，目标词之前的内容完全相同，详见附录。

2) 研究发现双字合成词结构不影响中文阅读过程的注视位置效应，但前人研究者认为双字合成词结构影响读者的词汇识别，作者认为，这是两个不同的研究内容。并采用孟红霞

和白学军(准备中)的研究予以说明,解释略显牵强。这是否与本研究所采用的双字合成词材料均在句中呈现有关?同时,两种研究内容的不同具体体现在什么地方?请说明。

非常感谢专家的建议和意见。具体修改内容如下:

前人的研究考察的是双字合成词结构对词汇识别过程的影响,而本文重点考察的是双字合成词结构是否影响读者对两种结构双字合成词的首次注视位置和再注视概率,这是两个不同的研究内容,前人的研究考察的是何时移动眼睛,而本研究考察的是眼睛移向何处。而且有研究表明有两个独立的系统分别决定着何时移动眼睛和眼睛移向何处(Rayner & Pollatsek, 1981)。Rayner 和 Pollatsek (1981) 在实验一中采用移动窗口范式,随机改变窗口大小。结果发现眼跳距离随窗口大小的变化而变化;但注视时间不受影响。实验二,研究者操纵当前注视词的呈现时间。当一次注视开始的时候,屏幕出现掩蔽刺激使中央凹视觉中文本信息出现时间延迟。结果发现注视时间的随延迟时间的变化而变化;但眼跳距离不受影响。这表明眼睛移向何处和眼睛何时移动是两个独立的过程。

3) 讨论中的部分文字与引言部分重复,如 4.1 和 4.2 的第一段,建议进行调整,使文章的论述精炼,重点突出。

非常谢谢专家的提醒和建议。具体修改内容如下:

4.1 中文阅读过程中读者的眼跳策略

以拼音文字为实验材料的大量眼动研究发现:拼音文字阅读过程中存在偏向注视位置和最佳注视位置,表明拼音文字阅读过程中读者眼跳选择的目标是“词”,而不是“字母”(Rayner, 2009; Rayner & Liversedge, 2011)。关于中文阅读过程中是否存在偏向注视位置和最佳注视位置,一直存在争论和分歧(Li et al., 2011; Tsai & McConkie, 2003; Yan et al., 2010)。本文在前人研究的基础上,进一步考察中文阅读过程中是否存在偏向注视位置和最佳注视位置,以及眼跳选择的目标是什么,为中文阅读过程中眼动控制的第二个问题提供更多实验证据。

4.2 双字合成词结构不影响中文阅读过程中的注视位置效应

如上所述,本研究发现中文阅读过程中眼跳目标的选择是基于词。那么词的哪些特性影响阅读中的注视位置效应呢?已有研究发现,词的一些高水平语言(或认知)因素(如词频、语境的预测性和合理性等)不影响中文阅读过程中的首次注视位置(郭晓峰, 2012; 吴捷等, 2011)。中文文本书写系统中大多数词是双字词,构成双字词的两个汉字的字义与整个词的词义之间主要存在两种关系。整词的词义与构成整词的汉字的字义之间的关系是否影响阅读中的注视位置效应。中文文本书写系统中的并列结构和偏正结构的双字合成词正好符合上述两种关系,因此本文的第二个研究目的是考察并列结构和偏正结构双字合成词上的注视位置效应是否存在差异。

4) “倒 U 型线性趋势”和“负向线性趋势”的表述欠妥,建议作者将语句表述清晰。

非常感谢专家的建议。具体修改内容如下:

在单次注视条件下,读者对目标词的首次注视更多地落在词的中心位置,首次注视位置在四个区域中的分布呈倒“U”形分布。多次注视条件下首次注视更多地落在词的开始部分,首次注视位置在区域 1、2、3、4 中的分布呈逐步下降的趋势。

5) 研究与 Li 等人(2011)的结果不一致,与白学军等人(2011)的结果一致。建议作者详细阐述加以说明,尤其是对于 Li 等人(2011)研究与本研究不相一致的可能原因。建议采取更确凿的证据证实本文作者研究的可靠性和合理性。

非常感谢专家的建议。具体修改内容如下:

这可能是由于 Li 等人 (2011) 研究中, 在分析四字词条件下的实验结果时, 将四字词作为目标区域; 但是在分析双字词条件时, 除了将双字词作为目标区域之外, 还将双字词之后的两个汉字也作为分析的对象, 使得双字词条件下可能包含 2 个或 3 个词, 而四字词条件下只有四字词一个词。而白学军等人 (2011) 以及本研究在分析结果时, 无论哪种条件下都只包含一个词。对于此问题还需要更多的实验证据来进一步验证。

6) 引言中的研究假设部分提及“重点考察并列结构和偏正结构双字合成词上的注视位置效应是否有差异”, 但两个实验的研究结论均未出现阳性结果, 这说明对于双字合成词意义的理解与其结构方式无关, 换句话说, 对双字合成词的加工是一个十分稳定的过程, 这是否由于双字合成词结构材料在句中出现的的原因? 若如此, 则研究结论的适用范围是否需要进行调整? 作者还需进行深入探讨。

非常感谢专家的建议和意见。具体修改内容如下:

首先, 专家解释得非常合理, 为了排除本研究结果是由目标词在句中出现的的原因造成, 研究者在造句的过程中, 对每一组目标句, 做到了目标词之前的内容完全相同, 排除了情境、文本在加工过程中对实验结果的影响。

其次, 中文阅读过程中的眼动控制包含两个基本问题: 一是阅读材料的什么因素决定读者何时 (when) 移动眼睛, 从时间的维度考察哪些因素影响词汇识别过程; 二是什么因素决定读者的眼睛移向何处 (where) (Rayner, 2009), 从空间的维度考察哪些因素影响读者的注视位置。本研究探讨的是第二个问题, 即什么因素决定读者阅读过程中的注视位置效应。前人研究通过传统的反应时实验发现双字合成词结构影响词汇识别过程, 是从时间的维度来考察; 本研究发现并列结构和偏正结构双字合成词上的注视位置效应没有差异, 是从空间的维度来考察。而且, 有研究发现, 何时移动眼睛和眼睛移向何处是两个相对独立的过程 (Rayner & Pollatsek, 1981)。

4. 其他细节问题

1) 文中交叉使用“偏向注视位置”、“最佳注视位置”和“PVL”、“OVP”的用法, 建议统一。

非常感谢专家的建议, 我们统一了用词, 为“偏向注视位置”和“最佳注视位置”。

2) 建议作者认真核对英文摘要, In Experiment 1, 33 undergraduates were asked to read the experimental sentences which included target words, and another 29 undergraduates took part in Experiment 2. 被试人数与原文不符。

非常感谢专家的提醒, 我们对英文摘要中的被试人数进行了修改, 谢谢。

根据第二位审稿专家建议修改:

本文以眼动技术考察了结构双字合成词的注视位置效应, 资料翔实, 选题和方法有意义。请作者回答以下的一些问题:

1. 实验中考察并列和偏正结构双字合成词是否存在偏向位置上的差异, 但是实验中基于阴性结果认为并列和偏正结构双字合成词不存在偏向位置上的差异, 且实验中仅仅是这两种双字合成词进行对比, 没有其他的参照, 欠缺说服力。

非常感谢专家的建议。具体修改内容如下:

与拼音文字相似, 以中文为材料的眼动研究发现, 词的一些高水平语言 (认知) 因素, 如词频和预测性等不影响注视位置效应。但是, 中文书写文本系统中, 大多数词是双字词 (现代汉语频率词典, 1986), 构成双字词的两个汉字的字义与整个词的词义之间主要存在以下

两种关系：一种是整个词的词义由两个汉字的字义共同决定，如“报刊”，另一种是整个词的词义主要由第二个汉字的字义决定，如“民警”（符淮青，2004；张良斌，2008）。并列结构和偏正结构的双字合成词正好符合上述两种关系。整词的词义与构成整词的汉字的字义之间的关系是否影响阅读中的注视位置效应。为了更好地回答上述两个问题，本研究选取了两种结构的双字合成词，即并列结构和偏正结构，进一步考察词的另一个高水平语言（或认知）因素——双字合成词结构，是否影响中文阅读过程中的注视位置效应。

2. 本文得出结论“当首次注视落在词的开头部分时，再注视该词的概率增加”，这样的结果是基于对多次注视下数据的分析，然而文中并没有给出单次注视和多次注视的具体比例，仅在 2.7.3 中表明“多次注视中的首次注视在所有首次注视中所占的比例为 16.2%”，那么可以推测被试对目标词进行多次注视的比例显著少于单次注视，那么基于小部分比例的数据得出的结论，是否具有代表性，尚有待商榷，请对此问题进行进一步的说明。

谢谢专家的意见和建议。

“当首次注视落在词的开头部分时，再注视该词的概率增加”，此结果是基于“首次注视位置上的再注视概率”指标下的数据分析，即首次注视落在词的不同位置上发生再注视的概率有多大。当统计首次注视落在词的开头部分（区域 1）的再注视概率时，分子为首次注视落在区域 1 上发生再注视的个数，分母为首次注视落在区域 1 上单次注视和发生再注视的个数。

3. 论文中的一个结论（无论并列结构还是偏正结构的双字合成词，首次注视的位置都在词的中心），请作者在讨论中增加对这一重要结论的原因解释。

非常感谢专家的建议。具体修改内容如下：

本研究的两个实验都比较一致地发现，中文阅读过程中首次注视位置的分布存在分离的现象，即在单次注视条件下，读者对目标词的首次注视更多地落在词的中心位置，首次注视位置在四个区域中的分布呈倒“U”形分布；多次注视条件下首次注视更多地落在词的开始部分，首次注视位置在区域 1、2、3、4 中的分布呈逐步下降的趋势。这与前人的研究结果比较一致（白学军等，2012；白学军等，2011；Li et al., 2011；Shu et al., 2011；Zang et al., 2012），与拼音文字的结果亦比较一致（Rayner & Liversedge, 2011）。结果表明，在中文和拼音文字阅读过程中，读者在一个词上注视一次还是注视多次是预先计划好的。当读者的首次注视落在词的中心位置时，即最佳注视位置时，往往只需要一次注视就可以识别该词，不需要进行再注视。但是由于眼跳存在一定的误差，读者的首次注视可能偏离了最佳注视位置，即落在词的开始部分时，在这种情况下读者就需要再进行一次再注视，进而完成对该词的识别，增加识别该词的成功率。

4. 对于偏正结构的词来说，首个语素能非常好的预测第二个语素（歌舞、芹菜）和首个语素很难预测第二个语素（白菜），这两种情况，偏向注视位置是否不同？

非常感谢专家的意见和建议。

本研究主要考察并列结构和偏正结构双字合成词上的注视位置效应是否存在差异，因此并未操纵偏正结构中第一语素对第二语素的预测性，这为我们指明了进一步研究的方向，我们将对此问题做进一步探讨。

5. 词频的数据是基于 1986 年的《现代汉语频率词典》，距今将近 20 年时间。最好能提供被试对于词汇的熟悉程度的数据。

非常感谢专家的建议。修改内容具体如下：

分别邀请 40 名未参加正式实验的大学生对实验一和实验二的目标词进行熟悉度的五点等级评定，其中“1”代表非常熟悉，“5”代表非常不熟悉。结果发现：实验一，并列结构目标词 $M = 1.28$, $SD = 0.27$ ，偏正结构目标词 $M = 1.32$, $SD = 0.27$ ，表明目标词的熟悉程度较高。进一步对并列结构和偏正结构目标词的熟悉度进行 t 检验发现， $t(38) = 0.45$, $p = 0.658$ ，这表明并列结构和偏正结构目标词的熟悉程度之间不存在显著差异；实验二，并列结构目标词 $M = 1.20$, $SD = 0.10$ ，偏正结构目标词 $M = 1.26$, $SD = 0.12$ 。对并列结构和偏正结构目标词的熟悉度进行 t 检验发现， $t(38) = -1.62$, $p = 0.113$ ，表明并列结构和偏正结构目标词的熟悉程度之间不存在显著差异。

6. 请作者提供实验材料。

根据专家的建议，我们将实验材料添加到附录中，实验 1 和实验 2 的材料详见附录。

7. 当 $p > 0.05$ 时，请提供 P 的具体数据。

根据专家的建议，我们提供了所有当 $p > 0.05$ 时的 p 的具体数值。具体详见文中黄色背景的内容。

第二轮修改说明

(修改日期：2013-1-29)

非常感谢《心理学报》两位专家的审稿意见。下面根据专家建议，逐一说明修改之处。

根据第一位审稿专家建议修改：

1. 引言部分和讨论部分写作方式过于分散（引言分了 13 个段落），不够系统概括，不易于读者的理解。要围绕研究问题系统综合地展开论述。

非常感谢专家的建议和提醒。作者对本文的引言部分和讨论部分进行了总结和概括，具体修改内容请见文中黄色阴影的部分。

2. 拼音文字的研究发现，词的高水平语言(或认知)因素(linguistic /cognitive processing)（如词频和语境的预测性等）不影响偏向注视位置和最佳注视位置，中文作为一种表意文字，与拼音文字存在很大差异。到底基于中文的什么特性，使研究者要对中文展开研究？如此引出本研究的内容，略显牵强。

非常感谢专家的建议和提醒。具体修改内容如下：

中文作为一种表意文字，与拼音文字存在很大差异(Zang, Liversedge, Bai, & Yan, 2011)。例如，中文的基本书写单位是汉字，每个汉字占有相同的空间，但在视觉复杂性上存在很大差异，这种差异主要体现在笔画数的不同。另外，与拼音文字不同，中文书写文本中没有明显的词边界信息，语言学家关于中文中什么是词仍然存在争论。因此，以拼音文字为实验材

料得到的眼动研究结果，不能直接推论到中文。

3. 这一问题仍是作者需要继续回答的问题：“什么因素决定读者阅读过程的注视位置效应？”作者的回答是：双字合成词结构不影响阅读过程中的注视位置效应。仅仅从排除的角度去回答似乎不够令人满意。研究者多次提及要考察“词的哪些特性影响中文阅读过程中的注视位置效应”，但是最终的研究结论也仅是验证了别人的研究结果[如，吴捷，刘志方和刘妮娜（2011）和郭晓峰（2012）]。这是否能够回答作者自己所设定的研究目标呢？写作时是否可以将研究的目标定位得更为切合本研究的内容，否则会因为研究目标而误导读者。

非常感谢专家的意见和建议。具体修改内容如下：

目前，较多研究支持中文阅读过程中眼跳目标的选择是以词为基础。那么词的哪些特性影响注视位置效应，哪些特性对注视位置效应不产生影响？汉语中词的特性包括低水平视觉因素（如笔画数和词长等）和高水平语言（认知）因素（如词频、预测性等）。高水平语言（或认知）因素是否影响注视位置效应？前人的研究发现，词频和语境的预测性均不影响中文阅读过程中读者的首次注视位置（吴捷，刘志方，刘妮娜，2011；郭晓峰，2012）。但是，上述两项研究并不能全面回答上述问题。

4. 引言部分，需要详细检查表述。随后，一些研究得到了与 Yan 等人一致的结果（白学军，梁菲菲，闫国利，田瑾，臧传丽，孟红霞，2012；白学军，孟红霞，王敬欣，田静，臧传丽，闫国利，2011；陈燕丽，史瑞萍，田宏杰，2004；臧传丽，2010；Shu, Zhou, Yan, & Kliegl, 2011），如此表述欠妥。

非常感谢专家的建议和提醒。非常抱歉，作者在写作过程中出现了书写错误，具体修改内容如下：

随后，一些研究得到了与 Yan 等人一致的结果（白学军，梁菲菲，闫国利，田瑾，臧传丽，孟红霞，2012；白学军，孟红霞，王敬欣，田静，臧传丽，闫国利，2011；臧传丽，2010；Shu, Zhou, Yan, & Kliegl, 2011）。

作者对本文的表述和用语进行了详细地检查。

5. 该研究结果支持“战略-战术”模型，这是本研究的一个重要结论，但是结合这一模型进行的论述还不透彻，除了战略战术模型，还是无其他模型？在有比较的前提下进行论述会更有说服力，使结论的得出更为可信。建议作者进行深入论述。

非常感谢专家的建议和提醒。具体修改内容如下：

除了“战略-战术”模型之外，还有两个比较有影响力的模型：E-Z 读者模型和 SWIFT 模型（Liversedge, Gilchrist, & Everling, 2011）。两个模型都假设，所有的眼跳均计划落在目标词的中间位置（即最佳注视位置），这与“战略-战术”模型的词间眼跳策略比较一致。但并不是所有眼跳都是完美的，在实际执行眼跳计划的过程中经常出现误差。首先，两个模型都假设存在随机误差，它是一种正态分布，其标准差随实际眼跳距离的变化而变化。其次，眼跳过程中存在一种偏好，即偏好大约 7 个字母的眼跳距离。这两种倾向导致实际眼跳的落点位置分布偏向词的开头和结尾部分。关于第二种偏好眼跳距离，此结果是在以拼音文字为材料的眼动研究的结果基础上提出，此结果是否适用于中文阅读，还有待于进一步研究。因

为, 中文的平均眼跳距离为 2-3 个汉字, 英文的平均眼跳距离为 7-9 个字母 (Rayner, 2009)。中文中的 2-3 个汉字和英文中的 7-9 个字母是否具有相同的心理意义, 还需要进一步的探讨。

6. 从本研究的主要结论看, 多数主要观点均与前人的研究 (白学军等, 2012; 白学军等, 2011; Zang et al., 2012) 相一致, 本研究的创新点体现在什么地方? 与前人研究相比, 有哪些实质性的进展?

非常感谢专家的意见和建议。具体修改内容如下:

本研究的创新点: 根据现代汉语频率词典 (1986), 双字词占中文中所有词的大约 70%, 双字词在中文中占有一个很重要的地位。在双字词中, 构成双字词的两个汉字的字义与整个词的词义之间存在两种关系: 一是词义由两个汉字的字义共同决定, 如“报刊”, 二是词义主要由第二个汉字的字义决定, 如“民警”(符淮青, 2004; 张良斌, 2008)。双字词中的并列结构和偏正结构的双字合成词正好符合上述两种关系。字义与词义之间的关系是否影响读者对目标词的注视位置? 因此, 本文选择了并列结构和偏正结构的双字合成词作为目标词, 考察两种结构双字合成词上的注视位置效应是否存在差异。

与前人研究相比, 本文发现并列结构和偏正结构上的注视位置效应不存在差异, 即字义与词义之间的关系不影响读者的注视位置, 进一步扩展了关于中文阅读过程中注视位置效应的研究结果, 为建立中文的眼动控制模型提供了更多的实验证据。

第三轮修改说明

(修改日期: 2013-4-18)

非常感谢《心理学报》两位专家的审稿意见。下面根据专家建议, 逐一说明修改之处。根据第一位审稿专家建议修改:

1. 问题提出部分: 缺乏双字合成词的结构影响注视位置效应的理论假设。

非常感谢专家的建议。具体修改内容如下:

目前, 研究者运用计算机模拟, 提出了各种模型, 以解释阅读过程中出现的注视位置效应。其中, O'Regan 和 Jacobs (1992) 提出了战略—战术模型 (strategy-tactics theory)。该模型认为读者阅读过程中采用了一种词间眼跳战略 (strategy), 该策略引导读者注视每个词的最佳注视位置。但是由于眼跳误差的存在, 眼跳经常未落在计划注视的目标上。读者此时可以使用词内战术 (tactic): 如果眼睛没有跳到最佳注视位置附近, 那它就会立即移向单词的另一端, 对该单词进行再注视。除了“战略-战术”模型之外, 还有两个比较有影响的模型: E-Z 读者模型和 SWIFT 模型 (Liversedge, Gilchrist, & Everling, 2011)。首先, 三个模型的观点并不是对立的, 三者有一个相同的假设, 即所有的眼跳均计划落在目标词的中间位置 (即最佳注视位置)。战略-战术模型与 E-Z 读者和 SWIFT 模型的不同体现在误差方面: E-Z 读者和 SWIFT 模型将战略-战术模型中的误差分为了两部分: 随机误差 (它是一种正态分布, 其标准差随实际眼跳距离的变化而变化) 和偏好眼跳距离 (眼跳过程中存在一种偏好, 即偏好大约 7 个字母的眼跳距离)。三个模型在结果的预期方面并不存在差异, 只是在结果的解释方面对误差有不同的观点。

2. 实验部分:

(1) 实验 1: 缺乏词频、字频信息来源的介绍。

非常感谢专家的建议和提醒。

两个实验中的词频和字频信息均来源于 1986 年北京语言学院语言教学研究所编著的《现代汉语频率词典》。

(2) 实验 1 中“目标词的预测率为 0”。不知目标词的预测率是如何计算出现? 是评定的, 还是被试填写出来的概率? 为什么会是 0?

非常感谢专家的意见和建议。

目标词的预测率是被试填写出来的概率。请 20 名本科生对目标词的预测性进行评定。评定时给 20 名本科生呈现目标词之前的实验句子, 如“市委书记表示政府将鼓励各地_____”, 要求被试在该未完成的句子后面填上他们能想到词。结果发现目标词的填写率为 0%。

(3) 实验 1 中作者写到“为了考查双字合成词结构是否影响读者阅读过程中的再注视模式, 对再注视位置与首次注视位置之间的关系进行分析, 见图 7。需要指出的是, 由于没有足够的证据, 因此没有进行进一步的统计分析”。这里“没有足够的证据”是什么意思?

谢谢专家的建议和意见。

没有足够的证据指的是, 实验 1 中并列结构和偏正结构目标词上的再注视概率分别为 16% 和 15%, 这 16% 和 15% 的再注视又分别分布在目标词的四个位置上, 因此再注视落在每个位置上的比率就更低, 导致没有足够的数据进行统计分析。

(4) 实验 2 为什么和实验 1 统计指标不完全相同?

非常感谢专家的建议和提醒。

实验 2 的目的是为了进一步验证实验 1 的结果, 因此在写文章时, 为了使重点突出, 节省篇幅, 只报告了其中三个指标的结果。

(5) 中心位置的操作性定义是什么? 是否注视位置 2 和 3 就是中心位置? 为什么实验 1 中首次注视位置的偏好注视位置 2 和 3 没有区别, 但实验中注视位置 2 的比率大于位置 3 的比率?

非常感谢专家的意见和建议。

根据前人的研究(白学军, 梁菲菲, 闫国利, 田瑾, 臧传丽, 孟红霞, 2012), 称区域 1 为目标词的开端部分, 区域 2 和 3 为目标词的中心位置, 区域 4 为目标词的结尾部分。虽然实验 1 中首次注视落在区域 2 (29%) 和 3 (31%) 的比例没有区别, 实验 2 中首次注视落在区域 2 (35%) 的比例大于区域 3 (25%), 但是由于区域 2 和 3 共同组成了目标词的中心位置, 研究者在分析原因时, 以中心位置作为一个整体, 实验 1 和实验 2 中首次注视落在中心位置的比例均为 60%。

三、讨论部分:

(1) 作者认为“在中文和拼音文字阅读过程中, 读者在一个词上注视一次还是注视多次是预先计划好的”。这句话是什么意思? 读者在阅读之前会事先计划好注视一次还是多次吗?

非常感谢专家的意见和建议。

本研究发现, 如果首次注视落在词的中心位置, 那么只需要一次注视就能识别该词; 如果首次注视没有落在词的中心位置, 将计划再注视, 从而增加该词识别的成功率。这些结果表明, 落在词上的首次注视位置不是随机的, 而是有计划、有目的的, 即在执行实际的眼

跳运动之前,运球运动系统已经完成了对该词的注视计划,因此读者在一个词上决定注视一次还是再注视是预先计划好的。

(2) 作者写道“本研究的两个实验分别统计了所有向前的注视位置及其分布”,但实验 2 并没有给出向前的注视位置及其分布的数据。

非常感谢专家的建议和提醒。

非常抱歉,由于作者的疏忽,导致行文中出现了上述疏漏。具体修改内容如下:

本研究的实验 1 统计了所有向前的注视位置及其分布。结果发现,读者对目标词所有向前的注视位置分布并非是一条平行于 x 轴的直线,所有向前的注视并非均匀分布在双字词的四个区域,而是在最后一个区域有明显的下降趋势,与 Li 等人(2011)的结果不一致,与白学军等人(2011)的结果一致。

(3) 不清楚战略-战术模型和 E-Z 模型和 SWIFT 模型的区别,也不清楚这些模型的预期,因此从本文看不出为什么支持战略-战术模型。

非常感谢专家的建议和提醒。

关于阅读过程中读者的眼跳策略,如引言所述,主要存在三种模型。这三种模型都假设读者的眼跳计划落在词的中间位置,但是三个模型对误差的解释存在差异。E-Z 读者模型和 SWIFT 模型将误差分为了随机误差和偏好眼跳距离。偏好眼跳距离是基于拼音文字的研究结果得出,此结果是否适用于中文眼动研究,还有待于进一步的研究。因为,中文的平均眼跳距离为 2-3 个汉字,英文的平均眼跳距离为 7-9 个字母(Rayner, 2009)。中文中的 2-3 个汉字和英文中的 7-9 个字母是否具有相同的心理意义,还需要进一步的探讨。本研究两个实验的结果均发现,中文阅读过程中注视位置分布存在分离的现象,单次注视事件中,读者的首次注视往往落在词的中间位置,只需要一次注视就可以完成对目标词的识别,较少需要再注视;当首次注视落在词首位置时,会计划一次再注视,并且再注视往往落在词的结尾部分。因此,“战略-战术”模型能够更好地解释本文的研究结果。

根据第二位审稿专家建议修改:

《并列和偏正结构双字合成词的注视位置效应》有一定的理论意义,但存在以下问题:

1. 该研究结果支持“战略-战术”模型,但在引言部分未提该模型;它和 E-Z 读者模型和 SWIFT 模型的区别也没有表述清楚;没有比较如果根据三种不同的模型,实验应该会分别得到什么样的结果。

谢谢专家的意见和建议。

目前,研究者运用计算机模拟,提出了各种模型,以解释阅读过程中出现的注视位置效应。其中,O'Regan 和 Jacobs(1992)提出了战略—战术模型(strategy-tactics theory)。该模型认为读者阅读过程中采用了一种词间眼跳策略(strategy),该策略引导读者注视每个词的最佳注视位置。但是由于眼跳误差的存在,眼跳经常未落在计划注视的目标上。读者此时可以使用词内战术(tactic):如果眼睛没有跳到最佳注视位置附近,那它就会立即移向单词的另一端,对该单词进行再注视。除了“战略-战术”模型之外,还有两个比较有影响的模型:E-Z 读者模型和 SWIFT 模型(Liversedge, Gilchrist, & Everling, 2011)。首先,三个模型的观点并不是对立的,三者有一个相同的假设,即所有的眼跳均计划落在目标词的中间位置(即最佳注视位置)。战略-战术模型与 E-Z 读者和 SWIFT 模型的不同体现在误差方面:E-Z 读者和 SWIFT 模型将战略-战术模型中的误差分为了两部分:随机误差(它是一种正态

分布,其标准差随实际眼跳距离的变化而变化)和偏好眼跳距离(眼跳过程中存在一种偏好,即偏好大约7个字母的眼跳距离)。三个模型在结果的预期方面并不存在差异,只是在结果的解释方面对误差有不同的观点。

2. 作者对评审专家的意见“实验中考察并列和偏正结构双字合成词是否存在偏向位置上的差异,但是实验中基于阴性结果认为并列和偏正结构双字合成词不存在偏向位置上的差异,且实验中仅仅是这两种双字合成词进行对比,没有其他的参照,欠缺说服力”解释得不到位,研究的意义显示不出来。

非常感谢专家的意见和建议。

首先,中文书写系统中,大多数词是双字词(现代汉语频率词典,1986),大约占71%,双字词中大多数是双字合成词(钱玉莲,2008)。根据构成双字合成词的语素之间的不同关系,可以将双字合成词分为几种不同的结构类型,其中并列结构和偏正结构占绝大多数(冯丽萍,2003)。其次,本文的研究目的就是考察并列结构和偏正结构双字合成词上的注视位置效应是否存在差异,因此只选取了这两种类型的双字合成词,没有选择其他类型的双字合成词。

同时,审稿人提出一个很好的思路,我们将在后面的研究中继续进行此项研究。

3.词汇切分由专家来完成还是由普通人完成,多少人完成,文中没有交代,也没有给出解释。这是个关键问题,因为数据是基于切分好的词收集的。有些双字词和紧邻的后两个词又可以组合成新词,有些不组合,且往往是成对的词,这方面没有控制,如实验1中的智谋过人、寝食不安、文武才能、因果复句、功过是非、功利主义、水力资源、短跑成绩、手艺高低、书报价格、专车数量、利弊条件、危难处境、手工产品、功绩卓著、水分循环、盈亏相抵、产销渠道、身心健康、水源短缺、水稻种植、温饱问题、物资供应等,实验2在此不一一列举。

非常感谢专家的意见和建议。

参考白学军等人(2012)和Yan等人(2009)的评定方法,对本实验材料的词切分一致性进行评定。第一步,根据《现代汉语词典》(2005)进行词切分,即在《现代汉语词典》中能查到的词将其切分为一个词,并编码;第二步,请未参加正式实验的20名大学生判断是否认可此种切分方式,若不认可,则标记出自己认为合理的切分方式,并进行编码;最后,计算两次编码的一致性百分比为95.8%,一致性范围在81%~100%之间。

4.材料熟悉度评定“并列结构目标词 $M = 1.28$, $SD = 0.27$,偏正结构目标词 $M = 1.32$, $SD = 0.27$,表明目标词的熟悉程度较高”,但并不是意味着每一个的熟悉度都高,应该选定一个标准,如将5点量表中低于2或者3的项目保留,再进行平均数、标准差的计算和t检验,这就要求编写材料前就要考虑可能有不符合标准的材料,在编写数量上要大于实验用材料。

非常感谢专家的意见和建议。

研究者对两个实验的熟悉度评定数据进行了仔细检查和核对,实验1中140个目标词的平均熟悉度在1.00-2.70之间,熟悉度的最大平均值即2.70;实验2中140个目标词的平均熟悉度在1.00-2.60之间,熟悉度的最大平均值即2.60。而且实验1和实验2中熟悉度评定数据的标准差都较小,表明数据的离散程度较小。

5.句子通顺性评定“请不参加正式实验的38名本科生采用七点量表对70对实验句进行通顺性评定,“1”代表非常不通顺,“7”代表非常通顺,结果显示 $M = 5.87$, $SD = 0.60$ ”,并不能“表明所有实验句都比较通顺”,只是通顺性程度较高,如实验一的16、25、27、34、

37、49、53 等句可能通顺性程度就不高，22 句中的目标词用了多音字，27 和 70 句理解起来可能有些困难等等，接下来应该删除通顺性低于 3 或者 4 的项目，再补充新的句子，进行二次评定，选出 70 对符合标准的句子，进行通顺性平均程度的 t 检验。

非常感谢专家的意见和建议。

研究者对两个实验的句子通顺性评定数据进行了仔细检查和核对，实验 1 中 70 对实验句的通顺性评定结果在 3.35-6.80 之间，只有一个句子的通顺性评定在 4 以下，即“著名舞蹈家杨丽萍凭借艺术意境极强的舞蹈抒发其内心的情感。”；实验 2 中 70 对实验句的通顺性评定结果在 4.06-6.88 之间。

6.实验二“要求 20 名大学生对目标词在句子中的预测性进行评定，结果发现目标词的预测性仅为 1.09%”，应该将能够预测出目标词的句子替换掉。

非常感谢专家的意见和建议。

本研究中目标词的预测率是被试填写出来的概率。请 20 名不参加正式实验的本科生对目标词的预测性进行评定。评定时给 20 名本科生呈现目标词之前的实验句子，如“市委书记表示政府将鼓励各地_____”，要求被试在该未完成的句子后面填上他们能想到词。结果发现目标词的填写率为 1.09%，目标词的填写率是非常低的。而且有研究发现，预测性不影响读者对目标词的注视位置效应（吴捷，刘志方，刘妮娜，2011；郭晓峰，2012）。

7.因变量指标上没有考虑多次注视情况下的回视来源，也就是说，多次注视次数中有可能混淆了由于语境不同（特别是实验一）或者目标词之后句子阅读产生的困难导致的回视，如果是这样，则区分不了词类不产生效应是否是语境效应相互抵消的结果。

非常谢谢专家的意见和建议。

本研究中的单次注视即在首次阅读（first pass）目标词时，对目标词只有一个注视点或只有一次注视的情况；多次注视即在首次阅读目标词时，对目标词有两个或两个以上向前的注视点的情况，并不包含词内回视和词间回视的数据。

8.关键区只设定了目标词所在的 4 个区，一般的做法还需要设定紧邻非目标词前后所在的一、两个区作为参考。

非常感谢专家的意见和建议。

对关键区的设定，本研究参考的前人的设定方法（白学军，孟红霞，王敬欣，田静，臧传丽，闫国利，2011；白学军，梁菲菲，闫国利，田瑾，臧传丽，孟红霞，2012；Shu, Zhou, Yan, & Kliegl, 2011; Yan, Kliegl, Richter, Nuthmann, & Shu, 2010; Zang, Liang, Bai, Yan, & Liversedge, 2012），前人在设定关键区时，都只是设定了目标词所在的 4 个区。

第四轮修改说明

（修改日期：2013-5-28）

非常感谢《心理学报》编委专家的审稿意见。下面根据专家建议，逐一说明修改之处。本文有一定发现，作者也对审稿人的意见做出了回应。但还存在下述问题：

1.行文口语化，句子经常不完整，标点符号使用不规范，参见稿件的批注。建议作者让

文字水平好的人通读一下。

非常感谢专家提出的宝贵建议。

作者邀请了文字水平较高的人对行文进行了一定的修改。

2.统计指标的使用不统一， F_2 时时有时无。 p 值的使用不规范，多数给出具体值，但也使用0.05等。建议全部改为0.05、0.01、0.001。

谢谢专家的建议。

关于 F_2 的统计，参考前人的研究（Zang, Meng, Liang, Bai, & Yan, 2013），本文对所有指标下的平均注视位置进行了被试分析（ F_1 ）和项目分析（ F_2 ），而在对每个指标下注视点落在不同区域的百分比进行分析时，只进行了被试分析，因为落在不同区域的数据分布并不相同，不适合做项目分析。

p 值的使用，按照专家的建议，统一改为了0.05、0.01和0.001。

3.讨论部分对双字合成词的结构无影响不令人信服。所以不影响可能与双字合成词的大小有关。如果双字合成词小，就可能不影响。

非常感谢专家的意见和建议。

按照三个眼跳策略模型的观点，阅读过程中读者在一个词（词 n ）上注视一次还是注视多次是预先计划好的。预先计划是读者在加工词 $n-1$ 时，通过副中央凹区域对词 n 进行了预加工。根据本文的研究结果，读者对目标词的双字合成词结构类型的预加工并没有达到可影响随后的眼跳选择目标环节的水平。或者还有另外一个可能，即由于双字合成词结构加工水平较高，读者对词 n 的预加工还没有达到此水平，而可能只是一些低水平的加工，如对词 n 的笔画数的加工，因此才会出现双字合成词结构不影响中文阅读过程中的注视位置效应。

专家提出，双字合成词结构之所以不影响注视位置效应，可能与双字合成词的大小有关。语言学家一致认为，由两个或两个以上的语素（亦称作词素）组成的词叫双字合成词。现代汉语词汇中，双字合成词占了绝大多数。双字合成词中，多数由两个语素构成，两个以上的语素构成的是少数（张良斌，2008）。双字合成词的大小，即一个双字合成词由几个语素构成。汉语词汇中，多数双字合成词由两个语素构成，因此本文只选择了由两个语素构成的双字合成词作为实验材料。注视位置效应是否受双字合成词的大小的影响，将是我们今后的研究方向。

第五轮修改说明

（修改日期：2013-6-7）

非常感谢《心理学报》编委专家的审稿意见。下面根据专家建议，逐一说明修改之处。

作者对论文进行了修改，仍存在下述问题：

1. 前言部分的介绍缺乏规范的语言学术语。文章通篇以词、字来论述，而在规范术语中，对组成词的汉字应采用“语素”（以往文献也叫做“词素”）的术语。目前，关于词与语素

的关系的研究多的是,但作者一点也没有提及,是不合适的。作者在行文中用“词首”、“词的开头部分”之类的概念(应为“首语素”)是也不合适的。“词的中心位置”又是指什么?是两个字的中间还是其他地方?这样不规范的术语有不少。

非常谢谢编委专家提出的修改意见!

语素是语言中最小的音义结合体,是最小的语法单位。词是语言中最小的能独立运用的音义结合体,词由语素构成(陆俭明,2005)。但是,为了与已有的文献叙述保持一致(白学军等,2012;白学军等,2011;臧传丽,2010;Shu et al.,2011;Tsai & McConkie,2003;Yan et al.,2010;Yang & McConkie,1999;Zang et al.,2013),本文在介绍中文的注视位置效应的相关研究时,仍然采用“字”和“词”的表述。

关于语素在双字合成词识别过程中的作用,有研究发现合成词在心理词典中是以分解的语素方式表征的(张必隐,1993),有研究者认为是以整词形式表征的(zhou & Marslen-Wilson,1994,1995),还有研究者认为,既存在语素表征,也存在整词表征,词汇识别是语素与整词激活相互作用的结果(彭聃龄,丁国盛,王春茂, Taft, 朱晓平,1999;王文斌,2001)。

本文参照前人的研究,提出了“词的中心位置”和“词的开端部分”等概念(白学军,梁菲菲,闫国利,田瑾,臧传丽,孟红霞,2012;白学军,孟红霞,王敬欣,田静,臧传丽,闫国利,2011;臧传丽,2010;Shu, Zhou, Yan, & Kliegl, 2011; Tsai & McConkie, 2003; Yan, Kliegl, Richter, Nuthmann, & Shu, 2010; Yang & McConkie, 1999; Zang, Meng, Liang, Bai, & Yan, 2013)。根据前人的研究,在分析眼动数据时,将每个目标词划分为四个区域(见图1)。对于分别落在区域1,2,3,4中的注视点,注视位置分别编码为0-0.5,0.5-1,1-1.5,1.5-2。区域1为目标词的开端部分,即“词首”,区域2和3为目标词的中心位置,区域4为目标词的结尾部分。

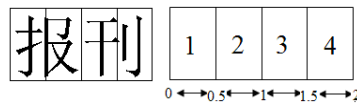


图1 目标词兴趣区划分方式

2. 作者虽然说明了文章中的“合成词”是双字词,但作为主要的研究发现,还是应该进行必要的限定。语言中非双字的合成词并不少。“双字合成词”比“合成词”准确。

谢谢编委专家提出的修改意见!

按照专家的建议,我们将文中的“合成词”严格限定为“双字合成词”。

3. 如果采用语素的概念,由两个汉字组成的词有的是双语素,如作者采用的,有的就只有一个语素,如“玳瑁”、“逍遥”,这样的合成词就更有代表性,更有研究价值。但作者并未采用。并列结构与偏正结构虽然不同,但均由两个语素构成。作者发现它们的结构无影响也是可能的。总之,在缺乏语素概念的情况下讨论合成词的结构的影响,是本研究的一大硬伤。修改时应尽力弥补。

非常感谢编委专家提出的宝贵修改建议。

由两个或两个以上的语素(亦称作词素)组成的词叫合成词(陆俭明,2005)。专家提到的“逍遥”和“玳瑁”两个词,均由一个语素构成,不是合成词,而是单纯词,即由一个语素构成的词(陆俭明,2005)。本研究主要考察的是并列结构和偏正结构双字合成词上的注视位置效应是否有差异,并未考察单纯词和合成词的注视位置效应是否存在差异。编委专

家为我们提供了今后的研究方向,即由两个汉字构成的单纯词和合成词上的注视位置效应是否存在差异,我们将做进一步的研究。谢谢编委专家!

4. 文字仍然有不规范的地方,如第一句话“阅读过程中,”应为“在阅读过程中,”。非常感谢专家的意见和建议!

非常抱歉,让编委专家两次提出同样的意见。作者再次对文章进行了认真细致的推敲,并邀请语言学教师对文章做了最后修改。