

目录

中文摘要	I
ABSTRACT	III
第一部分 文献综述	1
1 概念概述	1
1.1 预期性效应	1
1.2 整合效应	3
1.3 预期性效应与整合效应的混淆	3
2 字母语言中预期性效应在目标前词上的体现	5
2.1 冠词位置上的预期性效应	5
2.2 形容词位置上的预期性效应	7
3 中文研究中预期性效应在目标前词上的体现	7
3.1 量词	7
3.2 量词位置上的预期性效应	8
4 预视加工对预期性的影响	9
4.1 预视加工	9
4.2 预期性效应与预视加工的相互影响	10
第二部分 研究思路	11
1 问题提出	11
2 研究思路及设计	12
3 研究意义	13
3.1 理论意义	13
3.2 实践意义	13
第三部分 实证研究	14
实验 1: 中文阅读理解中, 目标前词量词位置上的预期性效应	14
1 实验目的与假设	14
2 研究方法	14
3 实验结果	17

4 讨论.....	20
实验 2: 量词和目标词间添加一个汉字, 目标前词量词位置上的预期性效应.....	21
1 实验目的与假设.....	21
2 研究方法.....	21
3 实验结果.....	22
4 讨论.....	26
实验 3: 量词和目标词间添加两个汉字, 目标前词量词位置上的预期性效应.....	27
1 实验目的与假设.....	27
2 研究方法.....	28
3 实验结果.....	29
4 讨论.....	32
第四部分 总讨论.....	32
1 目标词上的预期性效应.....	33
2 目标前词位置上的预期性效应.....	34
3 预期与文本不一致时, 修改错误预期.....	35
4 对教育教学的启示.....	36
5 研究不足与展望.....	36
第五部分 结论.....	37
参考文献.....	38
攻读学位期间发表的学术论著.....	50
致谢.....	51

中文摘要

在阅读理解过程中，预期性是减轻认知负荷，促进高效阅读的重要途径。因此，探究阅读理解过程中的预期性效应，有利于帮助学生掌握高效阅读的方法，进行快速阅读，为教师进行阅读指导提供依据。跨语言的研究表明，阅读理解中的预期性效应非常稳定，相比于低预期词，高预期词更易于加工，诱发的 N400 波幅更低，注视时间更短，跳读率更高。但有研究者提出，在目标词出现时，难以区分预期性效应和整合效应。高预期词更容易加工可能是自上而下的预激活，也可能由自下而上的语义整合导致。目标词的整合效应是在目标词出现时，将当前目标词输入的视觉信息与前文语境进行整合，而预期性效应是在目标词出现前，根据前文语境和大脑中储存的知识经验对即将出现目标词的预激活。因此，需要在目标词整合效应出现前，探究目标词(N)的预期性效应是否能够在目标词出现前，即目标前词(N-1)位置上体现，从而印证预期性效应对学生阅读理解的促进作用，为促进学生高效阅读提供方法论。

实验 1 采用自然句子阅读任务，结合眼动追踪技术，利用量词与名词语义一致的匹配性，操纵了目标词的高、低预期性水平。结果表明，目标词(N)上存在预期性效应，在所有注视时间指标上，高预期词的注视时间显著短于低预期词。此外，预期性效应不仅在目标前词量词(N-1)位置上存在，在数量词(Nc)上也存在。在首次注视时间、单一注视时间、凝视时间上均发现高预期词的注视时间更短，与前人研究结果一致。但读者在注视量词时，也会获取注视点右侧目标词的预视信息，对注视点右侧目标词进行预视加工。读者获取的目标词的预视信息会影响量词的加工，量词的结果可能受到目标词预视信息的影响。因此，需延长量词和目标词名词之间的距离，再次探究目标前词量词位置上是否存在预期性效应。

实验 2 采用和实验 1 相同的实验设计，依托实验 1 的材料，在量词和目标词之间添加一个汉字，延长量词和目标词之间的距离。结果发现，目标词(N)上存在稳定的预期性效应，高预期词在所有注视时间指标上的注视时间均更短。在目标前词量词(N-2)和数量词(Nc)位置上仍然存在预期性效应。在量词(N-2)上，所有注视时间指标上均表现出高预期词的注视

时间比低预期词更短。在数量词(Nc)上,高预期词的凝视时间更短,跳读率更高。虽然大多数研究者认为读者只能预视加工到注视点右侧一个字 N+1 的预视信息,但小部分研究者认为读者能预视加工到注视点右侧两个字 N+1 和 N+2 的词汇信息,而对 N+2 的预视加工仅停留在词汇层面,较难达到高水平的语义层面。在量词和目标词间添加一个字,读者在注视量词时,量词上的结果仍然可能受到目标词预视加工的影响。为避免预视加工到目标词的可能性,需要进一步延长量词和名词间的距离,再次探究目标前词量词位置上是否存在预期性效应。

实验 3 继续操纵目标词的高、低预期性水平,依托实验 1 的材料,在量词和目标词之间添加两个汉字,进一步延长量词和目标词之间的距离,探究目标前词量词(N-1)位置是否存在预期性效应。结果发现,在目标词(N)上,高预期词的首次注视时间,单一注视时间和回视路径时间显著短于低预期词,存在预期性效应。目标前词量词(N-3)和数量词(Nc)位置均存在稳定的预期性效应。在单一注视时间和凝视时间上,均发现与低预期词相比,高预期词的注视时间更短。由此证明,在目标词整合效应出现前,目标前词量词位置上存在预期性效应,且十分稳定。

本研究采用三个实验,结合中文特有的量词系统,发现目标词整合效应出现前,目标词的预期性效应能够在目标前词量词(N-1)位置上体现。延长量词和目标词的距离,避免注视量词时的预视加工对目标词的影响,目标前词量词依然存在预期性效应。研究为解决预期性效应与整合效应的争论和指导学生快速阅读提供支撑。

关键词 预期性效应,整合效应,量词,阅读理解

ABSTRACT

In the process of reading comprehension, predictability is an important way to reduce cognitive load and promote efficient reading. Therefore, exploring predictability effect in reading comprehension is conducive to helping students master efficient reading methods and speed reading, and providing teachers with a basis for reading instruction. Cross-linguistic studies shown that predictability effect in reading comprehension is very stable, with high predictability words being easier to process, inducing lower N400 amplitudes, shorter fixation time, and higher skipping rates than low predictability words. However, the researchers proposed that it was difficult to distinguish predictability effect with integration effect when target words (N) were present. Easier processing of high predictability words may be a result of top-down preactivation or by bottom-up semantic integration. The target integration effect is the integration of visual information from the current target word input with the preceding context when target word appears, whereas the predictability effect is the preactivation of the upcoming target word based on the preceding context and the stored knowledge experience in the brain before the target word appears. Therefore, it is necessary to explore whether the predictability effect of the target (N) can be reflected in the pre-target word (N-1) position before the target word integration effect appears, so as to corroborate the facilitating effect of predictability effect on the reading comprehension of students, and to provide a methodology for the promotion of students' efficient reading.

Experiment 1 used a normal sentence reading paradigm, combined with eye-tracking technology to manipulate the high and low predictability levels of the target words by utilizing the matching of classifiers and nouns for semantic consistency. The results showed that there was a predictability effect on target word (N), with high predictability words having significantly shorter fixations than low predictability words on all fixation measures. Furthermore, the predictability effect existed not only on the pre-target classifier (N-1) position, but also on the quantifier(Nc). Shorter fixation times were found for high predictability words on first fixation duration, single

fixation duration, and gaze duration, which was consistent with the results of previous studies. However, when readers fixated at classifier, they also acquired information about the target words to the right of the fixation point, and preview processing the target word on the right side of the fixation point. The preview information of the target word acquired by the reader affects the processing of the classifier, and the result of the classifier may be influenced by target preview information. Therefore, it is necessary to extend the distance between the classifier and target word to explore again whether there is an predictability effect on classifier position of the pre-target word.

Experiment 2 used the same experimental design as Experiment 1, relying on the material from Experiment 1, adding a Chinese character between the classifier and target to lengthen the distance between the classifier and target, and found that there was a stable expectancy effect on target (N), and that the high predictability word had a shorter fixation for all fixation measures. The predictability effect still existed on the pre-target word classifier (N-2) and quantifier (Nc) positions. On classifier (N-2), all fixation measures showed shorter fixation duration for high predictability words than for low predictability words. On quantifier (Nc), high predictability words had shorter gaze fixation duration and higher initial skipping rates.

While most researchers believed that readers can only preview the information of N+1 for one character to the right of fixation point, a small number of researchers believed that readers were able to preview the information for two words N+1 and N+2 to the right of fixation point, and that preview processing for N+2 was only at the lexical level, and it was more difficult to reach a higher level of the semantic level. By adding a word between classifier and target, the result on classifier may still be affected by the target word's preview processing when fixated at classifier. To avoid the possibility of preview processing to target, it is necessary to further extend the distance between the classifier and target to explore whether there is an predictability effect on classifier position of the pre-target word.

Experiment 3 continued to manipulate the high and low predictability levels of the target words, relying on the material from Experiment 1, and added two Chinese characters between the classifier and target to further lengthen the distance between the classifier and target, to explore whether there was a predictability effect for the position of the pre-target word classifier (N-1). It was found that there was a predictability effect on target word (N), where the first fixation duration, single fixation duration, and regression path duration were significantly shorter for high predictability words than for low predictability words. There were stable predictability effects for both pre-target word classifiers (N-3) and quantifier (Nc) position. Shorter fixation durations were found for both single fixation durations and gaze durations for high predictability words compared to low predictability words. This proves that there was a predictability effect for pre-target word classifiers that was stable before the target word integration effect emerged.

This study used three experiments, to explore the predictability effect of the target word can be reflected in the pre-target classifier (N-1) position before the target integration effect appeared in conjunction with the Chinese-specific classifier system. Predictability effects remain for pre-target word classifiers by lengthening the distance between the classifier and target and avoiding the effects of preview processing on target when fixating at classifier. The study provided support for resolving the predictability versus integration effects debate and instructing students to read quickly.

Key words predictability effect, integration effect, classifier, reading comprehension

第一部分 文献综述

阅读是人类独有的认知活动之一，是个体认识世界，丰富经验，获得体验感的重要途径。在阅读理解过程中，学生会利用多种方式提高阅读效率，常用的一种方式预期，根据当前语境提供的信息，利用头脑中积累的知识经验对即将呈现的文本内容进行预期，即对即将呈现内容的提前加工。学生能根据文本线索，核实和调整预期内容，从而减轻认知负荷，进行高效阅读。提前预期即将到来的信息的能力可能是大脑高效处理信息的关键(Bar, 2011; Chow et al., 2018)。更好地了解预期，对于提高学生阅读效率，指导快速阅读具有重要作用。

1 概念概述

1.1 预期性效应

预期性(Predictability)是指在特定语境中，读者提前预期和考虑到特定目标词出现的可能性(Miller & Selfridge, 1950)。在目标词出现前，读者根据上下文提供的语境信息和已往积累的知识经验，对即将出现的目标词产生预期，预先激活目标词的语言信息，属于自上而下的加工。词汇的预期性水平通常根据填充测验任务进行测量，即呈现目标词之前的文本片段，让读者据此填写接下来出现可能性最高的词，目标词的填充比率即目标词的填写次数与总次数的比值，也就是该目标词在当前文本中的预期性水平(Taylor, 1953)。当语境提供了很高的限制时，只有少数的词能构成句子的合理延续。相反，当句子语境提供了较低的限制时，许多潜在的词能够补全句子。相关研究表明，填充比率大于等于 0.7 的目标词为高预期词，填充比率小于等于 0.1 的为低预期词(Rayner & Well, 1996)。

研究者通常采用两种方式构建不同的预期性水平。最常用的方式是采用相同语境，选取不同填充比例的目标词，如例 1(目标词加粗呈现)。另一种方式是保持目标词不变，改变句子语境，进而改变目标词的填充比例，如例 2。

例 1：高预期：下雨天出门带**伞**是一个好习惯。

低预期：下雨天出门带**纸**是一个好习惯。

例 2：高预期：辽阔草原上奔驰的**马**十分健壮。

低预期：刘大爷家里散养的**马**十分健壮。

目前，阅读过程中的预期性效应(Predictability Effect)十分稳定，已有多方面证据支持，即高预期词比低预期词的加工速度更快(Ehrlich & Rayner, 1981)。在 ERP 技术中，主要通过 N400 波幅考察预期性效应。N400 波幅反映了从长时记忆中提取词汇语义的容易程度(Kutas & Federmeier, 2000; Lau et al., 2008)，与预期性水平呈反比关系(Dambacher et al., 2006; Kutas & Hillyard, 1984; Wlotko & Federmeier, 2012)。在字母语言中，研究者结合 ERP 技术发现高预期词诱发的 N400 波幅显著小于低预期词，表明高预期词更容易加工(Federmeier & Kutas, 1999; Grisoni et al., 2021; Kutas & Federmeier, 2011; Kutas & Hillyard, 1980; Lau et al., 2013; Lau et al., 2014; Stone et al., 2023)。另有研究者采用眼动追踪技术对预期性效应进行探究，发现读者对高预期词的跳读率比低预期词更高，注视时间更短，回视更少(Ashby et al., 2005; Drieghe et al., 2004; Hand et al., 2010; Kretzschmar et al., 2015; Rayner et al., 2004; Rayner et al., 2011; Rayner & Well, 1996; Staub, 2011; Reichle & Drieghe, 2013)。

预期性效应具有跨语言的稳定性，研究证实，中文研究同样存在显著的预期性效应。Zhang 等人(2019)利用 ERP 技术，操纵不同句子的高、低预期性水平，采用快速序列呈现范式(Rapid Serial Visual Presentation task, RSVP)系列呈现文本内容。结果发现，与低预期词相比，高预期词诱发的 N400 波幅更小。为考察预期性效应对眼球运动的影响，Rayner 等人(2011)对此进行探究。研究操纵了目标词的高、低预期性水平。结果表明，在首次注视时间和凝视时间上，高预期词的注视时间显著短于低预期词。其他研究也得到了一致的结果(Chang et al., 2023; Cui et al., 2022; Li et al., 2018; Liu et al., 2020; Rayner et al., 2005; Xu et al., 2022)。

眼动追踪技术具有更高的生态效度，并且能让读者在正常阅读过程中，实时观察读者的眼球运动。与 ERP 技术片段化系列呈现文本，无法进行回视相比，眼动追踪技术不对读者的阅读顺序设限，提高了生态效度，更具有优越性。因此，在本研究中，研究者利用眼动追踪技术对预期性效应进行探讨。

1.2 整合效应

整合(Integration)是指将当前输入的词汇与上下文相结合,在与前文语境的互动中,形成连贯语义表征,属于自下而上的加工(Hagoort et al., 2005)。整合效应(Integration Effect)指与语境不一致词相比,语境一致词更容易整合到前文语境中,更容易加工(Brown & Hagoort, 1993; Hagoort et al., 2009)。

Van Berkum 等人(1999)利用 ERP 技术,探究目标词的整合效应。实验呈现三句话的故事,前两句话构成句子语境,操纵最后一句话中目标词的语境连贯性。语境连贯水平下,目标词与前文语境语义一致。在语境异常水平下,目标词与前文语境语义不一致,但在最后一句话中语义合理,确保语境异常词诱发的 ERP 成分能够归因于语境异常词在语境中更难整合。结果表明,语境异常词诱发的 N400 波幅显著大于语境连贯词,语境异常词更难进行语义整合。Hagoort 等人(2004)结合 ERP 技术和 fMRI 技术,探究了语言理解中的语义整合和世界知识整合效应。实验材料包含三个水平:语义一致,世界知识违背(违背常识)和语义违背(语义不一致)。结果表明,语义违背条件和世界知识违背条件诱发的 N400 波幅均显著大于语义一致条件,并且两种违背条件在目标词出现时均激活了左下前额皮质(LIPC)。语义违背条件和世界知识违背两种条件下的 N400 波幅没有显著差异,表明词汇语义和世界知识的整合快速且同时发生。后续研究均证实了整合效应的存在(Chen et al., 2017; Hald et al., 2007; Helder et al., 2019; Zhu et al., 2019)。

1.3 预期性效应与整合效应的混淆

研究发现,高预期词的注视时间比低预期词更短,这种注视时间的差异属于预期性效应,高预期词比低预期词更容易加工(Rayner et al., 2005; Rayner et al., 2011; Li et al., 2018)。在目标词出现前,读者更多地预激活与高预期词有关的信息,因此高预期词呈现时,其激活程度更高,更容易加工,注视时间更短(Staub, 2015)。然而,其他研究者对此存在异议,认为这并非或未必全部属于预期性效应,而是整合效应(Huetting & Mani, 2016; Picking & Gambi, 2018)。如例 1 所示,相比于整合语境不一致词“纸”,语境一致词“伞”更容易整合。在目标词出现时,高预期词“伞”的语义与前文语境表述的语义雨天一致,比低预期词“纸”更容易整合到前文语境中,所以高预期词的注视时间更短。

语言理解包含两种加工方式,即自上而下的加工和自下而上的加工。读者在目标词呈

现前，根据前文提供的语境信息和长时记忆中积累的知识经验对即将出现的目标词产生预期，预先激活目标词的语言信息，属于自上而下的加工。读者在目标词呈现时，对目标词进行词汇识别，将目标词携带的语言信息与先前输入的语境信息相结合，进行语义整合，形成连贯的语义表征，属于自下而上的加工(Federmeier, 2007; Ferreira & Chantavarin, 2018; Kuperberg & Jaeger, 2016)。由此，目标词上存在预期性效应与整合效应的混淆(Picking & Gambi, 2018)。

Schwanenfligel 和 Shoben(1985)研究表明，在控制目标词词频和句子合理性的情况下，读者对高预期词进行词汇判断的速度比低预期词更快。Federmeier 和 Kutas(1999)操纵了目标词的类型为预期词，类别内违背和类别间违背三种水平。比如在实验材料“在晚宴上，我想知道为什么妈妈不喝汤。然后我注意到她没有勺子/刀/碗”中，“勺子”为预期词，“刀”为类别内违背，“刀”和“勺子”均为手持餐具，属于同一语义类别，共享许多语义特征，但与语境不一致，“碗”为类别间违背，“碗”属于容器餐具，与手持餐具不是同一语义类别，共享的共同特征较少，但共享更高级的语义类别餐具，同样与语境不匹配。结果发现，预期词诱发的 N400 波幅显著小于类别内违背和类别间违背，类别间违背诱发的 N400 波幅最大。该结果属于预期性效应，高预期词比违背词更容易加工，在预激活高预期词时，类别内违背词获得一定程度的激活，比类别间违背更容易加工。但该结果同样可以用整合效应来解释，预期词最容易整合到前文语境中，类别内违背与预期词共享较多语义特征，相对于类别间违背更容易整合到前文语境中。Rayner 等人(2005)操纵目标词的预期性水平，结果发现，与低预期词相比，高预期词的注视时间更短，跳读率更高。上述研究结果均存在难以区分目标词上的预期性效应和整合效应的现象(Picking & Gambi, 2018)，难以确定高预期词的促进效应是读者自上而下的预激活，还是自下而上的语义整合，需要进一步探究。

研究者普遍发现目标词上存在预期性效应(Cui et al., 2022; Li et al., 2018; Liu et al., 2020; Rayner et al., 2004; Rayner et al., 2011; Rayner & Well, 1996; Reichle & Drieghe, 2013; Staub, 2011; Xu et al., 2022)，但是该效应是在目标词呈现时才出现，还是在目标词呈现之前才出现呢？若预期性效应在目标词呈现时出现，则难以区分预期性效应与整合效应。但是，若在目标词出现前呈现预期性效应，则只能被解释为预期性效应。这是因为，目标词的整合效应只有在目标词出现时，才能识别目标词，将目标词的语言信息整合到前文语境，形成

连贯语义表征。预期性效应是对即将到来目标词的预激活，读者可能在目标词呈现前，提前预期到目标词的语言信息，并在目标前词上体现。因此，需要在目标词(N)整合效应发生前，探究目标词出现前，目标词的预期性效应是否能在目标前词(N-1)位置上体现。

2 字母语言中预期性效应在目标前词上的体现

2.1 冠词位置上的预期性效应

DeLong 等人(2005)在目标词整合效应出现前，探究目标词的预期性效应是否能在目标前词(N-1)上体现，采用快速序列呈现范式，利用英语中的语音规律，即以元音开头的单词前的冠词为“an”，以辅音开头的单词前的冠词是“a”，设置高预期(a kite)、低预期(an airplane)两种实验条件。此外，低预期条件下的冠词(an)与高预期目标词(kite)不匹配。如果读者能在目标前词的冠词上预期到目标词的信息，觉察到低预期条件的这种不匹配，冠词位置上就存在预期性效应。结果发现，在目标词名词(N)上，高预期条件下的 N400 波幅更小。在冠词(N-1)上，同样发现高预期条件的 N400 波幅显著小于低预期条件。由此表明读者在目标词出现前的冠词(N-1)上预期到目标词的相关信息，目标词的预期性效应能在目标前词冠词(N-1)位置上体现。

Martin 等人(2013)探究了双语学习者目标词的预期性效应是否能在目标前词(N-1)位置上体现。选取母语为英语和西班牙语-英语双语学习者两类被试群体，采用和 DeLong 等人同样的实验逻辑，结合英语冠词的语音规则，设置高、低预期两种条件。结果发现，在目标词(N)上，母语和双语者均发现高预期词诱发的 N400 波幅更低，母语者的预期性效应更大。在目标前词冠词(N-1)上，只有母语者存在预期性效应。这说明双语学习者不能像母语者那样预激活即将到来的单词，不存在目标前词位置上的预期性效应，只有在目标词出现时，才能进行语义整合。但 Martin 等人与 DeLong 等人的研究在实验材料及程序方面存在不同。DeLong 等人的实验程序以 500ms 的 SOA 逐字呈现每个单词，名词后面紧跟着其他词。在 Martin 等人的研究中，整体呈现冠词前的语境，冠词和名词以 700ms 的 SOA 逐字呈现。名词是句子的最后一个词。因此，读者有更多的时间阅读 Martin 等人材料的每个单词。Martin 等人使用的程序，将读者的注意力集中在最后一个名词上，更有可能诱导读者对名词进行预期。

Ito 等人(2017)再次探究双语学习者目标前词冠词(N-1)位置上是否存在预期性效应。继续沿用 DeLong 等人的实验逻辑和程序, 改编 Martin 等人的实验材料, 操纵目标词的预期性水平。结果表明, 在目标词(N)上, 只有母语者在高预期条件下诱发的 N400 波幅显著小于低预期词, 双语者的 N400 波幅没有显著差异。在目标前词冠词(N-1)上, 两者均不存在预期性效应。Ito 等人的结果与 Martin 等人不一致, 主要有以下原因。第一, Ito 等人采用逐词呈现的程序, 而不是像 Martin 等人将句子上下文作为整体呈现, 减少将读者的注意力集中到关键的冠词-名词组合上, 导致读者策略性地在冠词上预期名词的影响。第二, Ito 等人的研究中, 读者阅读了 160 个填充句和 64 个实验句, 而 Martin 等人没有使用填充句。当没有看到任何填充句时, Martin 等人的读者可能已经意识到, 在每个句子中出现较晚的冠词-名词组合总是高预期或相对不可预期的。这可能导致读者对 a/an 操作格外关注, 从而无意中鼓励参与者进行预期加工。而填充句形成的阅读情境会影响读者的预期加工 (Brothers et al., 2017; Lau et al., 2013), Ito 等人填充句的使用让实验操纵变得不明显, 从而减少了读者特别注意实验操纵的可能性。

Nieuwland 等人(2018)再次探究目标词的预期性效应是否能在目标前词(N-1)位置上体现。受心理学和神经科学研究中需要大样本研究的启发(Open Science Collaboration, 2015), 将被试量增加到 356 人, 采用了与 DeLong 等人相同的实验设计, 实验材料和实验程序。结果发现, 在目标词名词(N)上, 存在显著的 N400 效应, 高预期词诱发了更小的 N400 波幅。在目标前词冠词(N-1)上, 高预期条件诱发的 N400 波幅与低预期条件没有显著差异。Nieuwland 等人没有在目标前词(N-1)位置上发现预期性效应。这与 DeLong 等的研究结果不一致。分析两个研究出现不一致结果的原因, 研究者认为, 冠词只提供即将出现词汇的语音特征, 冠词后呈现的词汇不一定是名词。这一观点得到了语料库(Corpus of Contemporary American English and British National Corpus)的支持, 冠词 a/an 后紧跟名词的概率是仅为 33%, 实验材料中冠词名词连续出现的情况并不常见, 读者不会立即在冠词位置上预期名词的信息。此外, DeLong 等人(2017)提到 2005 年原始实验使用了大量包含冠词-名词组合的填充句, 即每句话都包含冠词-名词组合的不寻常的实验语境导致被试策略性地在冠词上预期即将到来的名词, 而 Nieuwland 等人未使用含冠词-名词组合的填充句, 所以两个研究得出了不一致的实验结果。

综上所述，英文研究中，在目标词整合效应出现前，目标前词冠词(N-1)位置上的预期性效应不一致。

2.2 形容词位置上的预期性效应

Van Berkum 等人(2005)研究了荷兰语中目标词的预期性效应是否能在目标前词(N-1)位置上体现。不定冠词后紧跟带有性别标记后缀的形容词，操纵不定冠词与形容词-名词后缀的性别一致性，高预期条件下冠词与形容词-名词后缀性别一致，低预期条件下冠词与形容词-名词后缀性别不一致。采用快速序列听觉呈现范式逐词呈现实验材料。研究结果发现，在名词(N)上，相比于高预期水平，低预期水平的 N400 波幅更大。在形容词(N-1)上，当读者遇到低预期条件的形容词时，觉察到低预期形容词与预期名词性别后缀不匹配，在 50-250ms 时间窗内产生更大的正波。证实目标前词形容词位置上存在预期性效应。

Otten 等人(2007)操纵了具有性别标记的形容词后缀与目标词名词的一致性，发现在名词(N)上，低预期词产生了更大的 N400 波幅。在形容词(N-1)开始后的 300-600ms，低预期词诱发了更大的负波。Otten 和 Van Berkum(2008)发现与带有性别标记的高预期词相比，低预期词在形容词开始后 900-1100ms 诱发了更大的负波，这是因为形容词和名词间隔至少三个词(即至少 1800ms)，所以在形容词开始后 900-1100ms 之间发现了较晚的负波。

总而言之，尽管不同研究在目标前词形容词上诱发了不同的脑电成分，上述研究均发现了目标前词(N-1)位置上的预期性效应。

3 中文研究中预期性效应在目标前词上的体现

3.1 量词

在中文中，当数词或者代词修饰名词时，名词前面必须有量词(Li & Thompson, 1989)，例如“一棵树、这条路”。量词及其相关名词的语义特征必须一致，如有生命性、形状和大小(Allan, 1977; Tai, 1994)。例如，有些量词只能搭配无生命名词(例如，朵，一朵花)，而一些量词可以修饰有生命名词(例如，名，一名教师，一名医生)(Huang et al., 2023)。

量词的一个重要属性是，它们能够限制紧随其后的名词(Qian & Garnsey, 2016)。量词和名词之间存在着某种特定的对应关系(Chou et al., 2014)。有些量词只能与极少数名词连

用, 对与其匹配的名词提供了较高的语义限制, 它们被称为特定量词, 如“盏”, 只能与灯连用, 构成“一盏灯”。少数量词如“个”和“些”与大量名词兼容, 对它们的配对名词的限制相对较低, 因此被称为一般量词(Gao & Malt, 2009)。无论是专用量词还是通用量词, 其搭配名词的合理性都可以通过填充概率测量。

量词与其相关联名词之间的密切关系已在多项研究中得到证实。Huetting 等人(2010)采用视觉情境范式, 考察量词是否影响名词的加工。视觉情境范式(Visual-world Paradigm)即向被试呈现视听双通道的信息, 在播放实验材料录音前, 给被试呈现图片材料, 要求被试根据录音对图片材料中的目标刺激进行注视, 计算目标物体的注视比例。研究给被试呈现四幅图片, 其中一幅是与量词匹配的目标物体, 另外三幅为干扰物。结果发现, 读者在听到含有特定量词的话语时, 更多注视与量词匹配的物体, 而不是干扰物。量词与名词不匹配会诱发更大的 N400 波幅(Chou et al., 2014; Qian & Garnsey, 2016; Zhang et al. 2012; Zhou et al. 2010)。

上述结果表明, 在中文中, 量词与其关联名词之间的语义特征必须匹配, 不匹配会导致加工难度增加。

3.2 量词位置上的预期性效应

Kwon 等人(2017)探究目标词整合效应出现前, 预期性效应是否能在目标前词量词(N-1)位置上体现。采用快速序列视觉呈现范式, 设置了预期词, 语义相关词, 语义无关词三种水平。操纵三种水平的目标前词量词与目标词语义一致, 分别是高预期词水平(点缀在夜空的几颗星在闪烁), 语义相关水平(点缀在夜空的几朵云在闪烁)和语义无关水平(点缀在夜空的几本书在闪烁), 其中目标词加粗呈现。语义相关词和语义无关词均为低预期水平, 其量词(“朵”和“本”)与高预期词(“星”)不匹配。结果发现, 在目标词名词(N)上, 高预期词引发的 N400 波幅显著小于语义相关词和语义无关词, 语义无关词诱发的 N400 波幅最大。在量词(N-1)上, 发现与目标词一致的结果, 高预期词诱发的 N400 波幅最小。分析研究结果, 语义相关词与语义无关词同为低预期词, 但语义相关词与高预期词共享语义特征, 所以语义相关词诱发的 N400 波幅比语义无关词更大。量词(N-1)上的结果表明, 读者在量词上预期到目标词名词的语言信息, 在目标词前的量词(N-1)位置上存在预期性效应。

然而, Kwon 等人未控制实验句的合理性, 量词上的结果可能是量词的整合效应, 即

高预期词水平的量词比语义无关水平更容易整合到前文语境中(Huang et al., 2022)。因此, 需要控制语义合理性, 进一步探究在目标词整合效应发生前, 目标词的预期性效应是否能在量词(N-1)位置上体现。

Huang 等人(2023)使用脑电(EEG)和表征相似性分析(RSA)技术, 结合量词和名词的生命性特征一致的特性, 考察了在低限制语境下, 是否能预期目标词的生命性语义特征。研究向被试呈现有生命限制或无生命限制的量词, 后面紧跟生命性一致或不一致的名词。例如, “位”属于有生命限制的量词, 可搭配“老人”, 构成生命性一致的量名词组合“一位老人”, “张”属于无生命限制的量词, 可构成生命性一致的量名词组合“一张桌子”。“一位桌子”和“一张老人”属于生命性不一致的量名词组合。结果发现, 在名词(N)上, 与生命性一致条件相比, 不一致条件诱发了更大的 N400 波幅, 反映了读者遇到生命性不一致名词时语义整合更困难。研究者量化了量词出现后神经活动模式之间的相似性。RSA 结果显示, 在名词呈现之前的 240ms, 受生命性限制的量词的神经活动模式之间的相似性大于无生命限制的量词, 表明读者在量词上预期到名词生命性的语义特征。

4 预视加工对预期性的影响

4.1 预视加工

在阅读理解过程中, 读者不仅能够获取当前注视点的信息, 也能从注视点右侧词获取预视信息, 即预视效应(Preview Effect, Engbert et al., 2005; Rayner, 1975)。读者获取的注视点右侧词的预视信息会影响当前注视词的加工, 即副中央凹-中央凹效应(Parafoveal-on-Foveal Effect, POF, Risse et al., 2008)。研究者通常采用边界范式对预视加工进行探究。边界范式(Rayner, 1975)即在目标词左侧设置一条看不见的边界, 读者的注视点在边界左侧时, 目标词处呈现预视词。当读者的注视点越过边界, 目标词取代预视词呈现。

拼音字母研究表明, 读者能获取多种水平的预视信息, 如正字法信息(Rayner, 1975; White et al., 2005; Williams et al., 2006), 单词的首、尾字母信息(Johnson & Eisler, 2012; Johnson et al., 2007; Johnson et al., 2018; Milledge et al., 2021), 语音信息(Ashby & Rayner, 2004; Chace et al., 2005; Slattery et al., 2011; Sparrow & Mielliet, 2002)。研究者对是否能预视到更高层次的语义信息, 仍然存在争议。有研究者认为难以获得语义预视信息(Altarriba, et

al., 2001; Hyönä & Häikiö, 2005; Rayner et al., 1986)。但也有研究者证实能够获取语义预视信息(Hohenstein et al., 2010; Schotter, 2013)。

中文研究发现, 读者不仅能获得字形、语音等预视信息(Liu et al., 2002; Pollatsek et al., 2000; Tsai et al., 2004), 还能获得稳定的语义预视信息(Yan et al., 2009; Pan et al., 2020; 王穗苹 等, 2009; Zhang et al., 2023; Zhu et al., 2021)。Yan 等人(2009)采用边界范式, 操纵了目标词的预视类型, 共五种预视类型, 包括一致预视(“户”), 正字法相关预视(“广”), 语音相关预视(“互”), 语义相关预视(“门”)和无关预视(“丹”)。结果表明, 读者在首次注视时间, 单一注视时间和凝视时间上, 均获得正字法和语义预视信息。语音预视效应只在凝视时间上获得。在目标前词区域, 读者还获得了语义 POF 效应。

4.2 预期性效应与预视加工的相互影响

在阅读理解过程中, 预视加工是注视点右侧对左侧信息的影响, 读者还会受到按照文本阅读方向, 注视点左侧对右侧的影响, 即预期性的影响。研究表明, 预视加工和预期性效应均会影响阅读过程, 二者之间存在交互作用。

White 等人(2005)操纵了目标词的预期性和词长预视类型。词长预视分为正确预视和无效预视两种水平。其中, 无效预视是指字母 s 被插入到目标词和词 N+1 之间, 从而使得预视词词长信息错误。结果发现, 预期性和词长预视在凝视时间和跳读率上存在交互作用。正确词长预视条件下, 存在预期性效应, 而在无效词长预视时, 预期性效应不存在。其他研究者也得出了一致结论(Juhaus et al., 2008; Milligan & Schotte, 2024; Schotter et al., 2015; Staub & Goddard, 2019)。

中文研究中, Chang 和 Zhang 等人(2020)探究了预期性是否影响预视加工。采用边界范式, 操纵目标词的预期性水平(高、低预期)和预视类型(一致预视和假词预视)。结果发现, 高、低预期条件的预视效应在首次注视时间、凝视时间上有显著差异, 高预期条件的预视效应更大, 高预期有效促进了读者的预视加工(Chang, Hao et al., 2020; Cui et al., 2021)。

Xu 等人(2022)探究了预期性是否能调节语义预视效果。采用高限制句, 操纵了目标词的预期性水平和预视类型。预视类型包含一致, 语义相关, 语义无关三种水平。结果表明, 高预期条件下, 语义相关水平的首次注视时间与语义无关没有显著差异, 而在低预期条件, 语义相关的注视时间比语义无关条件更短。

第二部分 研究思路

1 问题提出

研究发现,目标词上存在预期性效应与整合效应的混淆(Ehrlich & Rayner, 1981; Picking & Gambi, 2018; Rayner et al., 2011),难以确定高预期词的促进效应是读者自上而下的预激活,还是自下而上的语义整合。由于目标词的整合效应只有在目标词呈现时,才能将当前输入的目标词的语言信息整合到前文语境中,而预期性效应是根据上下文语境信息和知识经验对即将出现目标词的预激活。因此,为避免目标词上预期性效应与整合效应的混淆,需要在目标词整合效应发生前,在目标词出现前,探究读者在目标前词(N-1)上是否能预期到目标词的语言信息,即目标词的预期性效应是否能在目标前词(N-1)位置上体现,为帮助学生高效阅读提供依据和支撑。

在英文研究中,DeLong 等人(2005)结合冠词与预期名词的匹配性,探究在目标词整合效应出现前,目标词的预期性效应是否能在目标前词冠词(N-1)位置上体现。结果发现,在目标前词冠词(N-1)上,高预期词诱发的 N400 波幅比低预期词更小,表明冠词上存在预期性效应。然而后续的研究由于实验程序,材料设计及填充句的使用等原因,没有得出与 DeLong 等人一致的结论(Ito et al., 2017; Martin et al., 2013; Nieuwland et al., 2018)。这表明英文研究中的目标前词(N-1)位置上的预期性效应不一致。

在荷兰语研究中, Van Berkum 等人(2005)探究了荷兰语中目标词的预期性效应是否能在目标前词形容词(N-1)位置上体现,操纵不定冠词与形容词-名词后缀的性别一致性,发现在形容词(N-1)上,低预期词在 50-250ms 诱发了比高预期词更大的正波。Otten 等人(2007)及 Otten 和 Van Berkum(2008)分别发现在形容词(N-1)出现后的 300-600ms,形容词开始后 900-1100ms,低预期词诱发了更大的负波。

在拼音文字的研究中,在目标前词考察预期性效应,总是依据高度语法化的语言特征,如冠词和具有性别特征的形容词,并与所考察语言中的显性形态句法标记相关(Greenberg, 1978)。中文研究表明,量词能限制紧随其后的名词,量词与其相关的名词语义特征一致(Chou et al., 2014; Qian & Garnsey, 2016),在量词上预期即将出现的名词(Huetting et al., 2010; Klein et al., 2012)。读者能觉察到量词与其关联名词之间的语义特征的不匹配,语义不匹配

会导致加工难度增加。鉴于此，为提高研究结果的普遍性，本研究使用不局限于语法化特征的语义特征。使用汉语的量词系统，探究目标词的预期性效应是否能在目标前词量词(N-1)位置上体现(Kwon et al., 2017)。

Kwon 等人(2017)发现在量词(N-1)位置上，高预期词诱发的 N400 波幅显著小于低预期词。然而，Kwon 等人未控制实验句的合理性，量词上的结果可能是高预期水平的量词比低预期词更容易整合到前文语境中，是量词上的整合效应(Huang et al., 2022)。因此，需要在控制语义合理性的情况下，进一步探究目标词整合效应发生前，目标词的预期性效应是否能在量词(N-1)位置上体现。

2 研究思路及设计

本研究的具体思路如下：

实验 1 操纵目标词的预期性水平为高、低预期两种条件。考察在目标词整合效应出现前，目标词的预期性效应是否能在量词(N-1)位置上体现。

注视点右侧的预视信息可能会影响当前注视点的加工，即读者注视量词时，注视点右侧目标词的预视加工可能会影响量词上的结果。量词位置上的预期性效应可能是 POF 效应。因此需要延长量词和目标词名词间的距离，进一步探究目标前词量词(N-1)位置上是否存在预期性效应。

实验 2 延长量词和目标词的距离，在量词和目标词间添加一个汉字，进一步探究量词位置上是否存在预期性效应。

大多数研究者认为读者只能预视加工到注视点右侧一个字 N+1 的预视信息，但小部分研究者认为读者能预视加工到注视点右侧两个字 N+1 和 N+2 的词汇信息，并且对 N+2 预视加工仅停留在词汇层面，较难达到较高水平的语义层面。在量词和目标词间添加一个字，读者在注视量词时仍然有较小的可能预视加工到目标词的信息，量词上的结果仍有可能是 POF 效应。为避免注视量词时，预视加工到目标词对结果的影响，需要进一步延长量词和目标词间的距离，再次探究目标前词量词(N-1)位置上是否存在预期性效应。

实验 3 进一步避免预视加工对目标前词量词(N-1)上预期性效应的影响，在量词和目标词间添加两个汉字，再次探究量词位置上是否存在预期性效应。

3 研究意义

3.1 理论意义

词汇的预期性是影响词汇阅读方式的最重要因素之一(Staub, 2015)。本研究采用自然句子阅读任务,探究在目标词整合效应出现前,即在目标前词量词上是否能预期到目标词的语言信息,即目标前词(N-1)位置上是否存在预期性效应,加深了对预期性效应的认识,有利于解决目标词上预期性效应与整合效应的争论,有助于预期性相关理论和模型的构建与完善,为中文阅读模型(Li & Pollatsek, 2020)提供借鉴。

3.2 实践意义

预期性普遍存在于日常生活中。学生对预期性非常敏感,在阅读过程中,学生不需要一直对文本内容进行逐字阅读,能够根据当前文本提供的信息,利用头脑中积累的知识经验对文本内容和文章脉络进行预期,猜测作者的意图和文章走向。当预期内容与文章内容相符,能够减轻认知负荷,提高阅读效率,进行快速阅读。当预期内容与文章内容不一致时,能够加深学生印象,便于学生识记知识点。

教师可以结合预期性对学生进行阅读指导。引导学生认真阅读文章目录,寻找关键词,并据此预期关键词的含义,培养学生带着自己的预期和问题进行跳读。通过快速浏览大致了解文章大意,看是否符合自己的预期及是否对此感兴趣。当文章内容符合预期,即可略读了解大概内容。当文章内容不符合预期时,可以做好标记,进行精读,寻找与预期不一致的原因。重视阅读内容的背景介绍,人物性格等背景因素,注意作者提前埋下的伏笔,启发学生再次根据文本内容和文章脉络对下文进行合理预期,将学生对关键词的预期及阅读时的预期与正文内容进行对比,寻找不同点,为高效阅读提供方法论。学生也可以记录发现与预期不一致文章内容的想法,学习作者的写作手法,从而提升写作能力。

在进行教材编制时,编制组成员可以提炼关键词,突出显示,放在教材的开头或章节末尾,引导学生提前预期课程重点,从而在学习时更有侧重点,提高学习效率。这在一定程度上也能唤起学生的好奇心,激发学生的学习兴趣,从而主动参与学习。教材编制者也要注意营造丰富的背景信息,为下文内容提前做好铺垫,便于学生根据背景知识和经验对下文内容进行预期,从而更快更好地掌握知识。

第三部分 实证研究

实验 1：中文阅读理解中，目标前词量词位置上的预期性效应

1 实验目的与假设

在目标词整合效应出现前，探究目标词的预期性效应是否能在量词(N-1)位置上体现。根据 Chow 和 Chen(2020)与 Kwon 等人(2017)，假设目标前词量词(N-1)和目标词(N)上均存在预期性效应，高预期词的注视时间更短，跳读率更高。

2 研究方法

2.1 被试

使用 PANGEA(v0.2)(Westfall et al., 2014)计算被试人数。根据 Brysbaert(2019)采用 $d = 0.4$ 作为效应量的良好估计，当固定效应量值 $d = 0.4$ ，项目数量为 70，统计检验力为 0.8 时，需 38 名被试。随机招募 76 名来自山东某大学的学生参与实验，招募人员母语均为汉语，视力或矫正视力正常且没有阅读障碍。被试均自愿参加本实验，不清楚实验目的，实验结束后根据表现获得一定的报酬。

2.2 实验仪器

采用 SR Research 公司研发的 EyeLink 1000 Plus 眼动仪，采样率为 1000Hz。显示器尺寸为 21 英寸，分辨率 1024×768 像素，刷新率为 120Hz。实验句和填充句在显示器上以白底黑字呈现，判断对错问题用白底红字呈现。被试眼睛与显示器的距离为 60cm，所有文本材料的字号为宋体 28 号，每个汉字约占 1.2°的视角。记录被试右眼的眼动数据。

2.3 实验设计及材料

实验设计为单因素两水平(目标词预期性水平：高 vs 低)被试内实验设计。自变量为目标词高预期，低预期两种水平。因变量为被试的注视时间和跳读率。

在现代汉语语料库中选取 70 个单字词作为目标词，根据量词-名词的语义一致性，选取与目标词语义一致的量词，构建 70 个量名词组合。研究表明，目标词词频，文本限制性均会影响预期性效应(Frission et al., 2017; Hand et al., 2010; Rayner et al., 2004)。目标词的笔画数也会显著影响读者的阅读加工(Liversedge et al., 2014; 彭聃龄, 王春茂, 1997)。虽然实验材料控制了高、低预期两种条件量词和目标词的词频和笔画数，但不能排除视觉复杂性

等其他无关变量对结果的影响，每组句子框架内，两种水平的量词和目标词仍然存在笔画数、视觉复杂性等其他水平的差异。为更好控制无关变量对量词和目标词的影响，构建两组句子框架，框架(A)中的高预期词同样出现在框架(B)中，每个目标词既出现在高预期句子中，也出现在低预期句子中。为控制文本属性(文本限制性)，每组句子框架除量词和名词不同，其余内容均相同，共 70 组句子框架。实验材料举例如下(目标词加粗呈现)：

实验 1 材料：

框架 A：高预期：美丽的新娘手里拿着一束**花**笑得开心。

低预期：美丽的新娘手里拿着一颗**糖**笑得开心。

框架 B：高预期：爱吃甜的小孩手里拿着一颗**糖**笑得开心。

低预期：爱吃甜的小孩手里拿着一束**花**笑得开心。

从 Cai 和 Brysbaert(2010)的语料库中计算目标词和量词的词频，高预期条件与低预期条件目标词词频无显著差异($t(34) = -1.18, p = 0.25$)，量词词频也无显著差异($t(34) = -0.03, p = 0.99$)。目标词的笔画数在两种条件下无显著差异($t(34) = -0.28, p = 0.79$)，量词笔画数也无显著差异($t(34) = -1.12, p = 0.27$)。

表 1 量词及目标词词汇属性统计(括号内为标准差)

	高预期		低预期	
	句子框架 A	句子框架 B	句子框架 A	句子框架 B
量词词频(次/百万)	263.31(612.87)	265.46(619.13)	265.46(619.13)	263.31(612.87)
目标词词频(次/百万)	108.69(99.02)	191.31(401.93)	191.31(401.93)	108.69(99.02)
量词笔画数	7.66(3.38)	8.49(2.73)	8.49(2.73)	7.66(3.38)
目标词笔画数	8.69(3.23)	8.89(3.38)	8.89(3.38)	8.69(3.23)

2.4 材料评定

2.4.1 预期性评定

依据 Taylor(1953)使用的填充测验任务对文本的预期性进行评定，呈现量词及目标词前的文本，由 30 名不参加正式实验的被试在空白处填写最先想到的量词和单字名词。根据句子框架进行拉丁方平衡，材料评定共分为 2 个版本，每个版本各包含 70 句实验句和 25 句低预期填充句。被试填完后，计算量词和目标词的填充概率。结果表明，高预期条件下，名词的填充概率为 81%，两组句子框架的填充概率分别为 83%和 79%；量词的填充概率为

69%，两组句子框架中量词的填充概率分别为 71% 和 68%。低预期条件下，名词的填充概率为 4%，两组句子框架分别为 3% 和 4%；量词的填充概率为 5%，两组句子框架分别为 5% 和 5%。统计结果显示，高预期水平名词的预期性显著高于低预期目标词($t(69) = 36.34, p < 0.001$)；高预期水平量词的预期性同样显著高于低预期水平($t(69) = 21.94, p < 0.001$)。

表 2 材料评定

	高预期			低预期		
	句子框架 A	句子框架 B	合并	句子框架 A	句子框架 B	合并
量词预期性	0.71(0.21)	0.68(0.19)	0.70(0.20)	0.05(0.10)	0.05(0.10)	0.05(0.10)
目标词预期性	0.83(0.14)	0.79(0.13)	0.81(0.14)	0.03(0.07)	0.04(0.09)	0.04(0.08)
实验 1 句子合理性	4.08(0.39)	4.00(0.40)	4.04(0.39)	3.96(0.37)	3.93(0.44)	3.94(0.40)
实验 2 句子合理性	4.07(0.35)	4.04(0.44)	4.06(0.40)	3.98(0.55)	3.96(0.40)	3.97(0.48)
实验 3 句子合理性	4.05(0.30)	4.02(0.34)	4.04(0.32)	3.97(0.39)	3.95(0.37)	3.96(0.38)

2.4.2 合理性评定

为确保实验句语义合理，逻辑通顺，避免量词上整合效应的出现，另外招募 36 名被试对实验材料进行合理性评定。合理性问卷共两个版本，每个版本各包含 70 句实验句和 25 句不合理填充句。添加不合理填充句是为了检验被试态度是否认真。规定被试对合理性进行 5 级合理性评定，采用正向、反向两种计分方式，数据收集完毕后，将反向计分的结果转换为 1 = 非常不合理，5 = 非常合理，并进行统计分析。评定结果显示，高预期水平的合理性平均分为 4.04，两组句子框架的合理性分别为 4.08 和 4.00。低预期水平的合理性平均分为 3.95，两组句子框架的合理性分别为 3.96 和 3.93。统计结果显示，高低预期两种条件的合理性无显著差异($t(69) = 1.65, p = 0.11$)。

2.5 实验程序

被试进入实验室，在被试机前就坐。要求被试调整座椅距离，将下巴放在下颌托上，尽量保持头部不要移动。被试准备就绪，讲解指导语：“同学你好，欢迎参与本次阅读实验！实验过程中会阅读不同的句子，请认真阅读并理解呈现的每个语句，不需要刻意记忆。读完一句话后，按手柄左侧的任意键翻页。某些句子后会出现关于前一句的问题，请根据前一句的表述判断对错，表述正确按手柄右侧的蓝色键，表述错误按绿色键。另外，每句话呈现前，屏幕左侧会出现一个小黑点，请盯准小黑点的圆心进行注视。”确认被试熟悉

指导语后进行三点校准，确保眼动仪精确记录被试的眼动轨迹。实验过程中时刻关注被试的注视情况，若发现注视点漂移，立刻进行重新校准。

实验共分为两个系列，每个系列都包含高、低预期两种实验条件，被试随机分配到其中的一个系列。正式实验每个系列包括 70 个目标句，35 句填充句及 35 个判断对错的问题。此外，有 6 句填充和 2 个问题作为练习。整个实验大约需要 25 分钟。

3 实验结果

所有被试正确率均在 75% 以上，平均正确率为 94.5%。删除 3 个标准差以外的数据，共有 1% 的数据被删除。实验结果分析使用的线性混合效应模型(Linear Mixed-effects Models)来自 R 软件(Version 3.6.3; R Core Team 2020)中的 lme4 软件包(Version 1.1-23, Bates et al., 2015)。固定效应为预期性水平，随机效应为被试和项目，为使数据更接近正态分布，对注视时间进行 log 对数转换。

由于量词较容易加工，量词的跳读率较高，且量词常与数词组合出现，构成数量词，因此将量词和量词前的数词合并为数量词 Nc 兴趣区。分析数词(N-2)，量词(N-1)，数量词(Nc)及目标词(N)四个兴趣区的数据。纳入分析的眼动指标包括：首次注视时间(First Fixation Duration, FFD)，指首次阅读兴趣区第一个注视点的注视时间；单一注视时间(Single Fixation Duration, SFD)，指兴趣区内有且只有一次的注视时间；凝视时间(Gaze Duration, GD)，指首次注视当前兴趣区到首次离开该兴趣区的持续的注视时间；回视路径时间(Regression Path Duration, RP)，指从某个兴趣区的第一次注视开始，直到注视点移动到兴趣区右侧为止，兴趣区内所有注视点持续时间的总和；首次跳读率(Initial Skipping Rate, ISP)，首次阅读时兴趣区被跳读的概率(闫国利 等, 2013)。

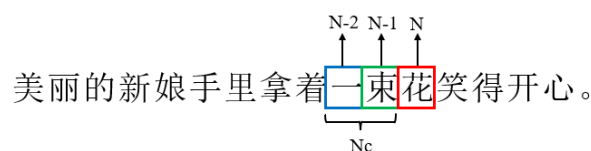


图 1 实验 1 兴趣区示例

各兴趣区的眼动数据结果见表 3。有关各兴趣区的推断统计结果，表 4 给出了效应系数(b)、标准差(SE)、 t 值或 z 值和 p 值。

在目标词(N)上, 预期性的主效应在所有注视时间指标差异显著。与低预期条件相比, 高预期条件下的首次注视时间、单一注视时间、凝视时间和回视路径时间更短。在首次跳读率上, 高预期条件和低预期条件差异不显著。

在量词(N-1)上, 预期性的主效应差异显著, 高预期条件下的首次注视时间、单一注视时间和凝视时间更短。两种条件下的回视路径时间和跳读率均无显著差异。

在数量词(Nc)上, 预期性主效应在首次注视时间、单一注视时间和凝视时间差异显著, 低预期条件下的首次注视时间、单一注视时间及凝视时间更长。两种条件在回视路径时间和首次跳读率上无显著差异。

在数词(N-2)上, 预期性的主效应在各眼动指标上均不显著。

表 3 实验 1 各兴趣区的眼动数据均值与标准差(单位: ms)

IA	预期性	FFD	SFD	GD	RP	ISP
N 目标词	高	223(86)	223(86)	232(95)	391(317)	0.53(0.50)
	低	235(87)	235(87)	245(96)	410(310)	0.52(0.50)
N-1 量词	高	206(73)	207(70)	208(73)	295(219)	0.71(0.45)
	低	216(71)	217(69)	220(73)	287(188)	0.70(0.46)
Nc 数量词	高	190(66)	191(63)	201(76)	273(201)	0.53(0.50)
	低	199(67)	198(65)	209(80)	282(205)	0.53(0.50)
N-2 数词	高	171(56)	172(55)	173(56)	218(137)	0.80(0.40)
	低	175(57)	175(57)	176(58)	236(171)	0.80(0.40)

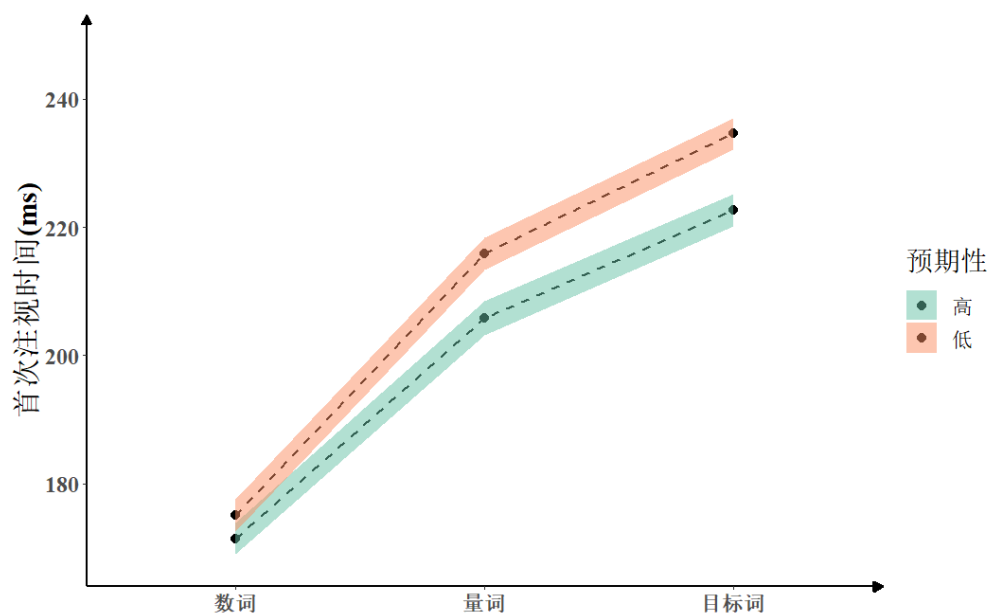


图 2 实验 1 不同兴趣区首次注视时间图

注: 彩色透明阴影表示标准误

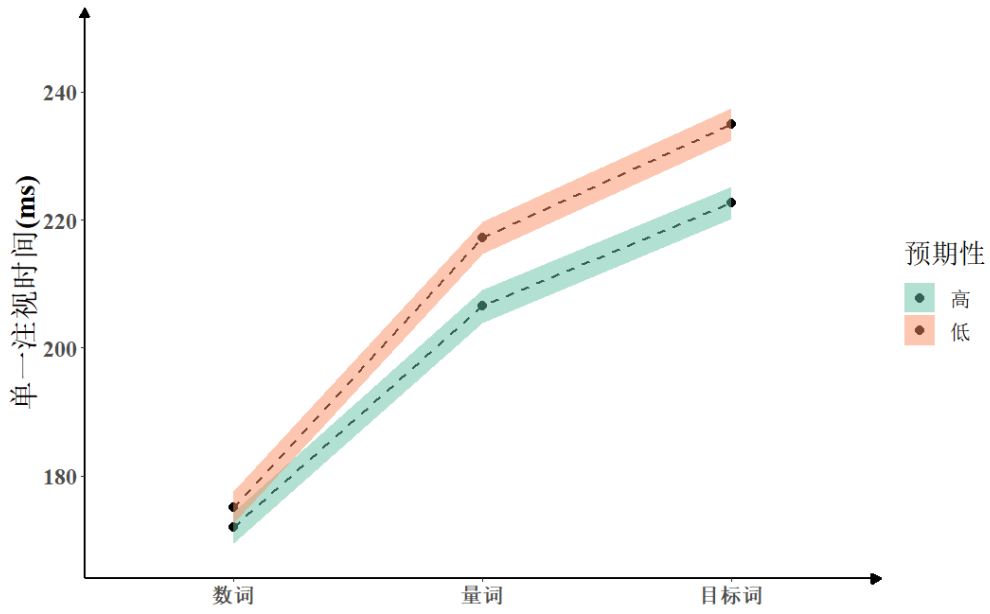


图 3 实验 1 不同兴趣区单一注视时间图

注：彩色透明阴影表示标准误

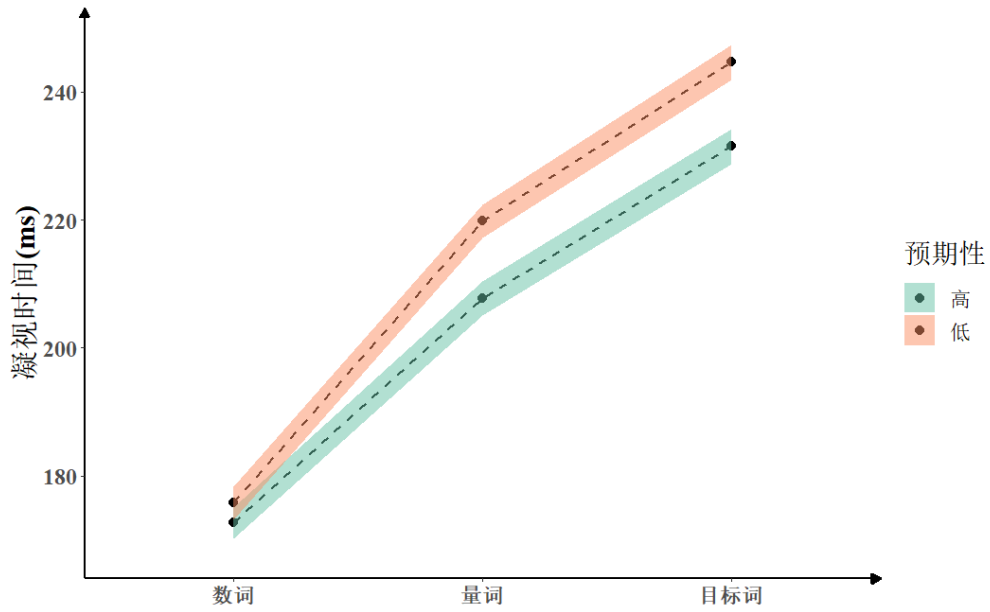


图 4 实验 1 不同兴趣区凝视时间图

注：彩色透明阴影表示标准误

表 4 实验 1 各兴趣区的眼动指标固定效应值

兴趣区	眼动指标	<i>b</i>	<i>SE</i>	<i>t/z</i>	<i>p</i>
N 目标词	首次注视时间	0.06	0.02	3.73	< 0.001
	单一注视时间	0.06	0.02	3.82	< 0.001
	凝视时间	0.06	0.02	3.95	< 0.001
	回视路径时间	0.06	0.02	2.39	0.02
	首次跳读率	-0.04	0.06	-0.71	0.48
N-1 量词	首次注视时间	0.07	0.02	3.22	< 0.01
	单一注视时间	0.06	0.02	3.21	< 0.01
	凝视时间	0.07	0.02	3.48	< 0.001
	回视路径时间	0.01	0.03	0.50	0.62
	首次跳读率	-0.06	0.06	-0.96	0.34
Nc 数量词	首次注视时间	0.05	0.02	3.15	< 0.01
	单一注视时间	0.03	0.01	2.07	0.04
	凝视时间	0.03	0.02	2.13	0.03
	回视路径时间	0.03	0.02	1.57	0.12
	首次跳读率	-0.02	0.06	-0.37	0.71
N-2 数词	首次注视时间	0.009	0.02	0.38	0.71
	单一注视时间	0.003	0.02	0.14	0.89
	凝视时间	0.003	0.02	0.13	0.90
	回视路径时间	0.03	0.03	0.97	0.33
	首次跳读率	0.05	0.07	0.65	0.51

注：加粗字体表示效应达到显著

4 讨论

实验 1 探究了目标词整合效应出现前，目标词的预期性效应是否能够在目标前词量词(N-1)位置上体现。研究操纵了目标词的预期性水平为高、低预期两种水平，采用自然句子阅读任务，控制句子合理性，控制目标词的词汇属性和文本属性，构建两组句子框架。结果发现，目标词(N)上存在预期性效应，高预期条件下的注视时间显著长于低预期条件，表明阅读过程中的预期性效应十分稳定(Chang et al., 2023; Cui et al., 2022; Li et al., 2018; Liu et al., 2020; Rayner et al., 2005; Rayner & Well, 1996; Staub, 2011; Xu et al., 2022)。但目标词上的预期性效应存在与整合效应的混淆。目标前词位置上的预期性效应可以反映在目标词整合效应出现前的预期性效应。

目标前词量词(N-1)位置和数量词(Nc)位置上均存在预期性效应，高预期词的首次注视时间、单一注视时间和凝视时间均短于低预期条件。证明目标前词量词位置上存在预期性效应，与前人研究结果一致(Chow & Chen, 2020; Kwon et al., 2017)。

读者在阅读过程中，不仅能够获取当前注视点的信息，也能从注视点右侧词获取预视信息(Engbert et al., 2005; Rayner, 1975)。当前注视词的加工容易程度，会受到注视点右侧词的词汇属性，如词频、词长、笔画数、预期性等的影响(白学军 等, 2009; Kliegl et al., 2006; Pynte & Kennedy, 2006; Rayner et al., 2003; Yang et al., 2009)。读者在注视量词时，会预视到右侧的目标词的信息，注视点右侧目标词的预期性会影响当前量词的加工。量词上，高预期词的注视时间更短，可能是因为读者预视到注视点右侧高预期目标词的预期性信息，进而对当前注视点量词的加工产生影响，即量词位置上的预期性效应有可能是 POF 效应。

考虑到预视加工对目标前词量词结果的影响，需要延长量词和目标词名词的距离，进一步探究目标前词量词(N-1)位置上是否存在预期性效应。

实验 2：量词和目标词间添加一个汉字，目标前词量词位置上的预期性效应

1 实验目的与假设

通过在量词和目标词间添加一个汉字的实验操作，从而实现延长量词和名词间距离。探究在目标前词量词(N-2)位置上是否存在预期性效应。根据 Chow 和 Chen(2020)与 Kwon 等人(2017)，假设目标前词量词(N-2)和目标词(N)上均存在预期性效应，高预期词的注视时间更短，跳读率更高。

2 研究方法

2.1 被试

使用 PANGEA(v0.2)(Westfall et al., 2014)计算被试人数。根据 Brysbaert (2019)采用 $d = 0.4$ 作为效应量的良好估计，当固定效应量值 $d = 0.4$ ，项目数量为 70，统计检验力为 0.8 时，需 38 名被试。另随机招募 76 名来自山东某大学的学生参与实验，实验结束后获得报酬。被试要求同实验 1。

2.2 实验仪器

同实验 1。

2.3 实验设计及材料

实验设计为单因素两水平(目标词预期性水平: 高 vs 低)被试内实验设计。自变量为目标词高预期, 低预期两种水平。因变量为被试的注视时间和跳读率。

在实验 1 材料的基础上, 在量词和名词之间添加一个汉字, 实验材料举例如下(目标词加粗呈现):

实验 2 材料:

框架 A: 高预期: 美丽的新娘手里拿着一束**鲜花**笑得开心。

低预期: 美丽的新娘手里拿着一颗**软糖**笑得开心。

框架 B: 高预期: 爱吃甜的小孩手里拿着一颗**软糖**笑得开心。

低预期: 爱吃甜的小孩手里拿着一束**鲜花**笑得开心。

2.4 材料评定

2.4.1 预期性评定

同实验 1。

2.4.2 合理性评定

评定过程同实验 1。评定结果显示, 高预期水平的合理性平均分为 4.06, 两组句子框架的合理性分别为 4.07 和 4.04。低预期水平的合理性平均分为 3.97, 两组句子框架的合理性分别为 3.98 和 3.97。统计结果显示, 高低预期两种条件的合理性无显著差异($t(69) = 1.44$, $p = 0.16$)。

2.5 实验程序

同实验 1。

3 实验结果

所有被试正确率均在 75% 以上, 被试阅读理解的平均正确率为 91.5%。删除 3 个标准差以外的数据, 共有 0.98% 的数据被删除。分析数词(N-3), 量词(N-2), 添加一个字(N-1), 目标词(N)及数量词(Nc)五个兴趣区的数据。

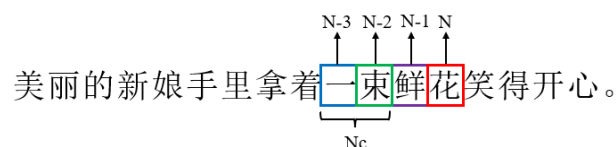


图 5 实验 2 兴趣区示例

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：<https://d.book118.com/508036031063007012>