

汉语阅读中先行词词频对代词加工影响的眼动研究^{*}

王文静 闫国利 白学军

(天津师范大学 心理与行为研究中心, 天津 300074)

摘 要 采用眼动分析法考察先行词词频对代词加工的影响。结果表明,先行词的词频对其代词加工的难易没有影响,即当先行词为高频词时对代词后区域的阅读时间与低频条件没有差异。研究结果支持了代词加工的词条重新通达假说,该理论认为读者在加工代词时只需通达先行词的部分词汇信息。

关键词 代词 照应语 先行词 词频 眼动

中图分类号 B842.1

文献标识码 A

文章编号 1003-5184(2006)03-0030-06

1 前言

代词的加工受到许多因素的影响。以往的研究大多集中在高水平的语言学因素对代词加工的影响上,例如动词的隐含因果关系、语义以及句法对代词加工的影响^[1-3],而鲜有研究探讨先行词低水平的词汇因素对代词加工的影响。本研究旨在考察其中的一种低水平因素——词频对代词加工的影响。

目前先行词词频对代词加工的影响存在三种理论观点。第一种为完全重新通达假说(full reaccess hypothesis),该理论认为当读者对代词进行加工时需要通达其先行词的全部词汇信息,就像读者重新加工先行名词一样。因此,该观点认为先行词的词频会影响代词的阅读时间,与加工其先行名词类似,当先行词为低频词时其代词的加工时间会长于高频条件。Cacciari 等人以及 Garnham 等人在罗曼斯语中发现读者在进行词汇通达时可以重新通达代词先行词的词性(gender)^[4,5]。第二种观点为词条重新通达假说(lemma reaccess account),该观点认为一个单词包括词条(lemma)和词位(lexeme)两种信息,语义和句法都属于词条信息,而词频则属于词位信息。读者在加工代词时只需要重新通达先行名词的词条,而非词位。由于词频属于词位信息,所以该观点认为先行词的词频不会影响代词的加工时间。由 Simmer 等人所做的实验结果支持了该理论,在他们的实验中,被试自控速阅读包含一个代词的句子(整句呈现),代词的先行词由一幅图片来呈现,图片所表征的要么是一个高频词要么是一个低频词,结果

发现词频对句子的阅读时间没有影响^[6]。然而,这一结果可能并不可靠,因为首先研究者测量的是整个句子的阅读时间,这个因变量不够灵敏,其次用图片来表征高低频词可能会由于图片本身的特征而掩盖词频所带来的效应。

第三种观点为突出性假说(saliency account),该观点认为词频会影响一个单词的突出性,低频词比高频词更突出更容易引起读者的注意,因而更容易从记忆中提取。Malmberg 等人的研究证明了低频词的识别要快于高频词^[7]。当读者加工代词时低频先行词应该比高频先行词更容易重新通达,因此低频先行词代词的阅读时间会短于高频先行词的代词。van Gompel 等人的研究支持了该理论^[8],但该实验中并未控制句子的合理性以及可预测性可能会对关键区域阅读时间所造成的影响。

针对 Simmer 实验中存在的问题,实验所采用的眼动技术可以较为准确地获得代词的阅读时间,先行词包含在实验材料的第一个句子中,从而避免了图片呈现所带来的问题。另外,实验控制了句子的合理性及可预测性从而避免了 van Gompel 实验中的不足。实验旨在考察:1)汉语阅读中先行词的词频是否会对代词加工造成影响。2)第二次重复出现的名词是否也存在词频效应。就第二个实验目的,根据完全重新通达假说,不论是照应语的词汇通达还是对先行词的重新通达都能够导致重复名词的词频效应,根据词条重新通达假说,如果该效应存在,这可能是由于照应语的词汇通达所造成的,如果不存

^{*} 基金项目 教育部社会科学重点研究基地重大项目(01JAZJDXLX003, 02JAZJDXLX003)

在则可能是由于对先行词的重新通达。最后,虽然突出性假说认为低频先行词的重新通达相对容易,但是这种效应会被照应语的词汇通达所掩盖,所以即使是第二次出现,高频词的阅读时间也会短于低频词。

2 研究方法

2.1 被试

天津师范大学的 62 名大学生,视力或矫正视力正常,其中 6 名被试的数据因为眼动记录不合格而被删除,有效数据 56 份。

2.2 实验材料

表 1 高频与低频词的词频、笔画数(括号内为标准差)

	频率(次/百万)	首字笔画数	尾字笔画数	整词笔画数
高频词	360.6(308.48)	6.7(2.39)	7.1(3.81)	13.9(4.55)
低频词	10.7(9.77)	6.4(2.88)	7.7(3.29)	14.2(3.93)

请 15 名大学生对实验材料进行 7 点合理性评定(7 为非常合理),评定结果均在 4.4~7 之间。另外请 15 名大学生对照应语指代的明确性进行 7 点评定,评定结果在 4.73~7 之间,保证所有代词的指代都是清晰明确的。插入 32 个与实验句结构差异较大的填充句,以及 20 个判断题,以保证被试认真阅读句子,对这些题目判断的正确率在 95% 以上。

表 2 实验材料举例

实验条件	实验句
高频先行词-重复名词	经理喊住路过的学生帮忙推车。 经理特别感谢学生的热情帮助。
低频先行词-重复名词	经理喊住路过的学员帮忙推车。 经理特别感谢学员的热情帮助。
高频先行词-代词	经理喊住路过的学生帮忙推车。 经理特别感谢他们的热情帮助。
低频先行词-代词	经理喊住路过的学员帮忙推车。 经理特别感谢他们的热情帮助。

2.3 实验设计

2(先行名词的词频:高频,低频)×2(照应语类型:重复名词,代词)二因素重复测量实验设计,先行词词频和照应语类型均为被试/项目内因素。另外,被试/项目分组为被试/项目间变量^[11]。

2.4 实验仪器和程序

采用 Eyelink II 眼动记录仪记录被试的眼动数据。Eyelink II 眼动记录仪的采样次数为每秒 250 次,刷新频率为 120Hz。

依据《现代汉语频率词典》^[9]和《现代汉语分类词典》^[10]选择满足实验条件的高低频近义词(词频、笔画数见表 1)。编制 32 组句子为实验材料(见表 2),每组句子包含两个分句,两个分句之间有逻辑联系,句长均控制在 13~16 个字。第一个分句包含先行名词,第二个分句包含照应语,指代第一个分句中的先行名词,照应语或者是一个代词或者是一个重复名词。先行词/照应语均位于句子中央,以避免首尾效应。两个分句的句子结构均相同。将实验材料分为四个词单,每个词单包含每组句子的其中一种实验条件,每个被试只看每组句子的其中一种条件。

具体步骤为:1)让被试熟悉实验室环境并进行眼校准程序;2)说明指导语,通过练习,确认被试完全理解实验要求后,正式实验开始。整个实验大约需要 20 分钟。

2.5 兴趣区划分

实验划分了 3 个兴趣区,兴趣区 1 为第一个分句中的先行名词,兴趣区 2 为第二个分句中的照应语,兴趣区 3 为第二个分句中紧跟在照应语后的词(见例 1)。

例 1 经理喊住路过的【学生】帮忙推车。

①

经理特别感谢【他们的】热情帮助。

②

③

3 实验结果

研究采用三个眼动指标:1)首次注视时间,指落在兴趣区的第一个注视点的持续时间;2)凝视时间,指从注视点第一次落入兴趣区开始,到离开该区域为止这段时间内,被试对该区域的注视时间的总和;3)总注视时间,指对兴趣区注视时间的总和,包括回视的注视时间。不同条件下的平均阅读时间见表 3。对结果进行被试分析(F_1)和项目分析(F_2)。由于仅对词频效应以及它和照应语的交互作用感兴趣,因此不报告照应语的主效应,因为如果照应语的主效应显著,这可能是由于代词和名词之间的词类或长度差异造成的。

表 3 被试在三个兴趣区上的平均注视时间(ms)和标准差(括号内为标准差)

	兴趣区		
	兴趣区 1	兴趣区 2	兴趣区 3
首次注视时间			
高频先行词 – 重复名词	229.66 (74.81)	213.35 (78.27)	240.57 (91.13)
低频先行词 – 重复名词	249.15 (84.47)	223.06 (79.15)	235.22 (90.36)
高频先行词 – 代词	232.01 (81.13)	216.92 (83.59)	232.51 (86.58)
低频先行词 – 代词	243.17 (81.20)	211.74 (77.02)	231.95 (91.82)
凝视时间			
高频先行词 – 重复名词	285.11 (138.29)	251.56 (120.28)	293.02 (147.48)
低频先行词 – 重复名词	337.43 (164.22)	254.21 (118.20)	307.36 (161.03)
高频先行词 – 代词	287.45 (137.41)	240.71 (113.67)	278.12 (140.09)
低频先行词 – 代词	334.47 (157.32)	227.99 (96.20)	273.65 (142.12)
总注视时间			
高频先行词 – 重复名词	510.70 (306.25)	596.59 (327.71)	624.87 (363.00)
低频先行词 – 重复名词	585.51 (340.68)	671.23 (352.76)	601.18 (334.08)
高频先行词 – 代词	513.50 (293.44)	520.67 (319.23)	609.95 (328.82)
低频先行词 – 代词	612.63 (365.21)	494.62 (295.13)	643.27 (367.94)

3.1 兴趣区 1

对于首次注视时间、凝视时间和总注视时间 ,词频主效应均非常显著。首次注视时间 : $F_{(1,52)} = 19.47, p < 0.01$, $F_{(1,28)} = 13.93, p < 0.01$;凝视时间 : $F_{(1,52)} = 46.87, p < 0.01$, $F_{(1,28)} = 32.08, p < 0.01$;总注视时间 : $F_{(1,52)} = 46.98, p < 0.01$, $F_{(1,28)} = 16.20, p < 0.01$ 。对高频词的阅读时间显著低于低频词。

3.2 兴趣区 2

对于首次注视时间、凝视时间和总注视时间 ,词频主效应均不显著。首次注视时间 : $F_{(1,52)} = 0.27, p = 0.61$, $F_{(1,28)} = 0.46, p = 0.50$;凝视时间 : $F_{(1,52)} = 0.76, p = 0.39$, $F_{(1,28)} = 0.56, p = 0.46$;总注视时间 : $F_{(1,52)} = 1.21, p = 0.28$, $F_{(1,28)} = 1.91, p = 0.18$ 。

对于首次注视时间 ,词频与照应语类型的交互

作用仅在被试分析时显著 , $F_{(1,52)} = 4.22, p < 0.05$, $F_{(1,28)} = 3.05, p = 0.09$ 。进一步简单效应分析发现 :当照应语为重复名词时 ,对高频词的阅读时间在数量上小于对低频词的阅读时间 ,但并未达到显著水平 , $F_{(1,52)} = 3.31, p = 0.07$;当照应语为代词时 ,无词频效应 $F_{(1,52)} = 1.01, p = 0.32$ 。对于凝视时间 ,词频与照应语类型的交互作用在被试分析和项目分析上均不显著 $F_{(1,52)} = 1.60, p = 0.21$, $F_{(1,28)} = 1.28, p = 0.27$ 。对于总注视时间 ,词频与照应语类型的交互作用在被试分析和项目分析上均非常显著 $F_{(1,52)} = 26.24, p < 0.01$, $F_{(1,28)} = 9.99, p < 0.01$,进一步简单效应分析发现 :当照应语为重复名词时 ,词频效应非常显著 $F_{(1,52)} = 9.11, p < 0.01$, $F_{(1,28)} = 8.42, p < 0.01$,即使是重复出现 ,对高频词的阅读时间仍然会小于对低频词的阅读时间 ;当照应语为代词时 ,词频效应在被试分析上显著 ,在项目

分析上不显著 $F_{(1,52)} = 25.04, p < 0.01, F_{(1,28)} = 1.41, p = 0.25$ 。

3.3 兴趣区3

对于首次注视时间、凝视时间和总注视时间,词频主效应均不显著。首次注视时间: $F_{(1,52)} = 0.38, p = 0.54, F_{(1,28)} = 0.67, p = 0.42$;凝视时间: $F_{(1,52)} = 0.46, p = 0.50, F_{(1,28)} = 0.66, p = 0.42$;总注视时间: $F_{(1,52)} = 0.18, p = 0.67, F_{(1,28)} = 0.12, p = 0.74$ 。词频与照应语类型的交互作用也不显著。首次注视时间: $F_{(1,52)} = 0.16, p = 0.69, F_{(1,28)} = 0.22, p = 0.64$;凝视时间: $F_{(1,52)} = 1.66, p = 0.20, F_{(1,28)} = 1.90, p = 0.18$;总注视时间: $F_{(1,52)} = 2.39, p = 0.13, F_{(1,28)} = 3.94, p = 0.06$ 。这一结果支持了词条重新通达假说。

4 讨论

首先证实了词频效应,高频名词的首次注视时间、凝视时间和总注视时间都短于低频名词,这与以往词频效应的研究相一致^[12]。另外,对于重复出现的名词,高频词和低频词的首次注视时间和凝视时间之间没有差异,而在总注视时间上高频词少于低频词,这说明词频仅对重复出现名词的晚期加工有影响,而这种影响可能主要来自于回视或对重复的名词的再次加工。特别需要指出的是,在照应语为代词的条件,当先行词为高频名词时,对代词后区域(兴趣区3)的阅读时间与低频时的没有差异,这说明先行词的词频对代词的加工没有影响。这些数据支持了词条重新通达假说,由于词频属于词位信息,而在加工代词的时候只需要重新通达先行词的词条信息,因此词频不会对代词的阅读时间造成影响。

本研究之所以在代词后区域(兴趣区3)考察词频对代词加工的影响,主要是基于以往的研究表明,对代词的加工会相对滞后,即最早的效应也会反应在代词之后的区域^[13]。这种对代词加工的轻微滞后可能是由于通达语篇表征所需要的时间要长于低水平的词汇学因素,因此在考察重复名词的词频效应时,需要关注兴趣区2,而在考察代词加工是否存在词频效应时,则需要关注兴趣区3,因为对代词的加工要求通达前文。在兴趣区2,并未发现对重复名词的早期加工存在词频效应,这可能是由于对先行词词条信息的重新通达所导致的,而在兴趣区3,照应语为重复名词的条件下也未观察到词频效应,再

次支持了词条重新通达假说。

研究结果与另外两种假说均不一致。如前文所述,完全重新通达假说认为对代词的加工需要重新通达先行词的所有信息,包括语义和词频信息,就像重新加工先行名词一样,实验的结果显然与该观点的预测不一致,词频对代词的加工没有影响,说明对代词的加工并不需要通达先行词的全部信息。另外,突出性假说认为低频的先行名词比高频的先行名词更显著,更容易引起读者的注意,获得更多的加工资源从而更容易被提取,因此在加工代词时,低频先行词应该比高频先行词更容易重新通达,从而减少阅读时间,而实验结果与该观点的预测不一致,即低频这一特征并不能降低对代词的加工时间。

5 结论

5.1 词频效应显著。对于首次加工的名词,对低频词的阅读时间要长于对高频词的阅读时间。

5.2 先行词的词频对代词加工没有影响,支持代词加工的词条重新通达假说。

5.3 对重复名词的早期加工不存在词频效应,低频词与高频词的阅读时间没有差异。

参考文献

- 孙燕,舒华,周晓林,等.动词隐含因果关系对代词加工的影响.心理科学,2001,24(1):39-41.
- 缪小春,宋正国.动词语义和句子语法对代词加工的影响.心理科学,1995,18(4):197-200.
- 缪小春.影响代词加工的语义和语法因素研究.心理学报,1996,28(4):352-357.
- Cacciari C, Carreiras M, Cionini C B. When words have two genders: anaphor resolution for Italian functionally ambiguous words. Journal of Memory and Language, 1997, 37: 517-532.
- Garnham A, Oakhill J, Ehrlich M-F, et al.. Representations and processes in the interpretation of pronouns: new evidence from Spanish and French. Journal of Memory and Language, 1995, 34: 41-62.
- Simner J, Smyth R. Phonological activation in anaphoric lexical access(ALA). Brain and Language, 1999, 68: 40-45.
- Malmberg K J, Nelson T O. The word frequency effect for recognition memory and the elevated - attention hypothesis. Memory and Cognition 2003, 31(1): 35-43.
- van Gompel R P G, Majid A. Antecedent frequency effects during the processing of pronouns. Cognition, 2004, 90: 255-264.
- 北京语言学院语言教学研究所.现代汉语频率词典.北京:北京语言学院出版社,1986.

10 董大年. 现代汉语分类词典. 上海: 汉语大词典出版社, 1999.

11 Pollatsek A, Well A D. On the use of counterbalanced designs in cognitive research: a suggestion for a better and more powerful analysis. Journal of Experimental Psychology: Learning Memory and Cognition, 1995, 21(3): 785 – 794.

12 Rayner K. Eye movement, cognitive processes, and reading. 心理与行为研究, 2004, 2(3): 482 – 488.

13 Rayner K, Pollatsek A. The Psychology of Reading. NJ: Lawrence Erlbaum Associates, Publishers, 1989. 268 – 289.

Antecedent Frequency Effects during Processing of Pronouns
On Eye Movements in Reading Chinese

Wang Wenjing Yan Guoli Bai Xuejun
(Research Center of Psychology and Behavior ,Tianjin Normal University ,Tianjin 300074)

Abstract :The present experiment investigated whether the processing of pronouns is affected by the lexical frequency of their antecedents. The data showed reading times following pronouns with infrequent antecedents were as same as following pronouns with frequent antecedents. This was consistent with a lemma reaccess account which claimed that reader accessed part of the lexical properties of the antecedent during the processing of pronouns.

Key words :pronouns ; anaphors ; antecedents ; lexical frequency ; eye movement

(上接第 9 页)

The Past ,Present and Future of Chinese Psychology

Yan Guocai
(Psychology Department of Shanghai Normal University ,Shanghai 200234)

Abstract :The author of this paper briefly analyses the development of Chinese psychology ,included three questions :the history of Chinese psychology, the present of Chinese psychology and the trend of Chinese psychology.

Key Words : Chinese psychology ,the past of Chinese psychology ,the present of Chinese psychology ,the future of Chinese psychology

(上接第 22 页)

Is There Only High and Low Division in Self – Esteem ?
—Heterogeneity of High Self – esteem and Its Revelation

Zhang Xiangkui Tian Lumei
(Department of Psychology , Northeast Normal University , Changchun 130024)

Abstract :Traditional view proposed that people with high self – esteem had better psychological and social adaptation , but more and more researchers believed this dichotomy of self – esteem too simple because of differences in nature in each level of self – esteem especially in high self – esteem , i. e. heterogeneity of high self – esteem. The article introduced it simply and suggested the important revelation for future self – esteem study and education.

Key words :self – esteem ; heterogeneity of high self – esteem ; revelation