

启动实验中收集基线反应时的必要性问题^{*}

董燕萍^{**1} 袁媛²⁽¹⁾广东外语外贸大学英文学院, 广州, 510420) ⁽²⁾广东外语外贸大学国际商务英语学院, 广州, 510420)

摘 要 启动实验的有效条件之一是, 控制组和实验组的目标词在基线反应时上相同, 但在实际操作中, 不少研究没有收集基线反应时, 只是确保了两组目标词在频率上相同。本研究实验结果显示: 对于母语而言, 如果词频很高, 相同频率的词得到相同反应时的机率很大; 相反, 机率就比较低; 对于外语而言, 这种机率相对更低。实验说明, 相同词频的词不一定能得到相同的反应时, 因此收集基线反应时是启动实验不可或缺的步骤。

关键词: 基线反应时 启动实验 词频

1 引言

启动任务(priming tasks)一般包括实验组和控制组, 每一组都由启动词(prime)和目标词(target)组成, 实验目的是比较被试对实验组和控制组的目标词的反应时。如果被试对这两组目标词的反应时不同, 说明实验组的启动词对目标词的影响在大脑存储或处理上不同于控制组; 否则, 不能说明这种区别的存在。这种比较的前提是: 两组目标词在没有加启动词时的反应时(即基线反应时)是相同的。收集基线反应时比较费时费力, 很多甚至是非常具有影响力的研究^[1-8]都采取了便捷的方法: 查找词频, 确保实验组和控制组的词频相同, 其依据是: 词频相同的词反应时也相同。

的确有不少研究^[9, 10]说明: 频率影响词汇决定(lexical decision)反应时, 高频词的反应时短, 低频词的反应时长。Foster's 的串行搜索模型试图解释这种频率效应, 认为高频词先被搜索, 低频词后被搜索, 因此后者需要更长的时间。试图解释这种频率效应的研究还有 Burgess 和 Hitch^[11]等等, 说明频率效应得到了广泛的认可。

但是, 频率效应是否意味着词频相同的词反应时相同呢? 这个结论对于高频词和低频词是否具有同样的答案? 对于母语和二语又如何? 本研究希望寻找这些答案。如果答案是否定的, 那么用词频来替代基线反应时的那些众多的研究就值得三思。

2 方法

2.1 被试

被试是 37 名中国某大学学生, 都至少学习了 8 年的英语, 对我们实验中的英语单词都熟悉。实验后获取一定报酬。

2.2 实验任务和材料

由于词汇决定任务(LDT; lexical decision task)是启动实验使用最为频繁的任务, 因此我们这里也采用了词汇决定任务。我们一共选取了 96 个词, 48 个汉语词汇, 48 个英语词汇, 并按照非词规则(符合每种语言的拼写规则但实际上又不存在的词)为两种语言各造了 48 个非词。每个被试因此

做了 192 个词汇决定。

每个语言中的 48 个词汇, 按照频率一一匹配, 构成两个小组, 一个用作实验组(为方便称作“组 1”), 一个用作控制组(称作“组 2”)。汉语的两个小组在频率上没有差异($F = 0.003$; $p = 0.954$), 平均频率分别是每百万字 150 和 152; 英语的两个小组同样没有差异($F = 0.015$, $p = 0.904$), 平均频率分别是每百万字 237 和 232。汉语频率查自《现代汉语词频》(1985); 英语的频率查自 BNC (British National Corpus, 见 <http://sara.natcorp.ox.ac.uk/lookup.html>)。

2.3 实验程序

实验程序通过 E-Prime 来控制。整个实验分为三块。第一块是练习, 目的是让被试熟悉实验操作和过程, 20 个汉语词构成汉语练习组, 20 个英语词构成英语练习组。第二、三块分别是汉语、英语词汇的主体实验, 各有 96 个词。一半被试先做汉语, 再做英语, 另外一半相反。组内词序随机呈现。实验说明在汉语块中用汉语说明, 在英语块中用英语说明。实验说明要求被试决定屏幕上出现的词汇是否是汉语/英语里的一个词, 如果是, 就按键盘上标有 YES 的键(即主键盘上的“1”键), 否则按“NO”键(即主键盘上的“0”键)。屏幕上呈现的词等待被试按键后才消失。

3 结果与分析

有 6 个被试由于出错过多或极端数据过多(超过 5 个)而被删除, 在剩下的 31 个被试的数据中, 我们分别对两种语言进行数据修理(trimming), 分别删除两个半标准差以外的数据, 这样, 汉语数据删除的比例为 3.5%, 英语数据删除的比例为 5.6%, 总体数据删除比例不超过 5%。修理后的反应时数据见表 1。

表 1 汉、英各两组材料反应时的描述性数据

组别及平均频率	<i>M</i>	<i>Max</i>	<i>Min</i>	<i>SD</i>
汉语组 1(150)	533	1170	239	134
汉语组 2(152)	540	1195	287	142
英语组 1(237)	715	1497	274	231
英语组 2(232)	635	1492	334	204

成对 *T* 检验表明, 两组汉语词汇在反应时上没有区别

^{*} 本文得到教育部人文社科重点研究基地重大项目资助(02JAZJD40022)。

^{**} 通信作者: 董燕萍, 女。E-mail: ypdong@mail.gdufs.edu.cn

($p = 0.255$), 但两组英语在反应时上有显著区别 ($p < 0.000$)。

这个结论是否可以推广? 是否和词汇的高低频率有关? 我们对每种语言中的两组材料分别进行操纵: 对于每组 24 个词汇, 先统计频率最低的 12 个词, 看看由这 12 个词组成的汉语(或英语)两个组是否在频率和反应时上有区别; 再统计频率最低的 14 个词, 依次类推, 直到每组保留频率最低的 22 个词(即删除了最高频的两个词)。然后我们来统计原来每组 24 个词中频率高的词汇, 先删除 2 个最低频率词, 对剩

下的 22 个最高频率词进行统计, 再删除 4 个最低频率词, 对剩下的 20 个最高频率词进行统计, 依次类推, 直到剩下 12 个最高频率词。

操纵后的组间匹配见表 2。检验结果说明: 频率相对比较高的两组汉语词汇, 如果在频率上一一匹配的话, 这两组词在反应时没有显著区别; 但当频率比较低时, 这两组词有可能在反应时上有区别 (p 值接近 0.05)。对于英语两组词的统计结果显示, 所有的组对都有显著区别 ($p < 0.000$)。

表 2 各种频率下每种语言中两组词汇反应时的差异(成对 T 检验)

	汉语				英语			
	平均频率		反应时差异		平均频率		反应时差异	
	组 1	组 2	t	p	组 1	组 2	t	p
12 个低频	77	80	-1.83	0.078	144	140	6.72	0.000
14 个低频	82	87	-1.44	0.160	155	150	6.17	0.000
16 个低频	89	94	-1.44	0.160	164	160	5.24	0.000
18 个低频	97	101	-1.54	0.133	174	170	5.83	0.000
20 个低频	105	108	-1.29	0.206	186	182	6.03	0.000
22 个低频	116	120	-1.11	0.278	207	201	6.08	0.000
22 个高频	161	162	-0.71	0.484	254	248	6.68	0.000
20 个高频	172	172	-0.85	0.401	268	263	6.42	0.000
18 个高频	182	182	-0.11	0.917	283	278	6.26	0.000
16 个高频	193	193	-0.27	0.791	298	292	5.83	0.000
14 个高频	206	206	-0.58	0.565	313	308	5.23	0.000
12 个高频	222	223	0.05	0.959	331	325	4.57	0.000

为了更加确切, 我们对汉、英每种语言中的两组材料重新进行更加紧密的一一匹配, 使两组材料在频率上的差异更小。重新匹配后的频率差异 p 值提高了很多, 接近 1(汉语两组频率差异的 p 值是 0.994; 英语两组频率差异的 p 值是 0.992)。按照这种分组的反应时数据见表 3。

表 3 重新分组后各组反应时的描述性数据

组别及平均频率	M	Max	Min	SD
汉语组 3(151)	531	1170	239	130
汉语组 4(151)	541	1195	258	145
英语组 3(235)	681	1497	334	236
英语组 4(235)	667	1483	274	206

显示, 不管是英语, 还是汉语, 频率相对比较高的两组词汇, 如果在频率上一一匹配的话, 这两组词在反应时没有显著区别; 但当频率比较低时, 这两组词在反应时上有显著区别。

结合表 2 和表 4, 我们只能得出这样的结论: 在频率上没有区别的, 甚至是一一匹配的两组词, 在反应时上不一定没有区别。对于外语而言, 频率一一匹配的两组词在多数情况下都不能得到相同(没有显著区别)的反应时, 但如果频率非常高, 偶然也会得到相同的反应时。对于母语而言, 频率没有区别的两组词不一定能得到相同的反应时, 但如果两组词在频率上高度匹配, 而且都是高频词, 那么多数情况下两组词的反应时没有显著区别。

按照上面同样的方法分别对汉语、英语进行统计。结果

表 4 重新分组后各种频率下每种语言中两组词汇反应时的差异(成对 T 检验)

	汉语				英语			
	平均频率		反应时差异		平均频率		反应时差异	
	组 3	组 4	t	p	组 3	组 4	t	p
12 个低频	78	80	-2.71	0.011	142	141	3.54	0.001
14 个低频	84	84	-2.5	0.018	153	152	4.2	0.000
16 个低频	91	92	-2.19	0.036	163	162	5.33	0.000
18 个低频	99	99	-2.52	0.017	172	171	3.89	0.001
20 个低频	106	107	-2.13	0.042	184	184	2.86	0.008
22 个低频	117	119	-1.86	0.072	203	205	2.1	0.045
24 个原词	151	151	-1.77	0.087	235	235	1.51	0.141
22 个高频	162	161	-1.01	0.323	251	251	0.56	0.579
20 个高频	173	171	-0.78	0.442	265	266	-0.43	0.673
18 个高频	183	181	0.1	0.925	280	281	0.35	0.728
16 个高频	194	192	-0.02	0.988	295	295	-0.5	0.622
14 个高频	207	205	-0.24	0.814	310	311	-0.63	0.534
12 个高频	224	222	0.14	0.893	327	329	-0.93	0.359

4 结论

上面的结论并不出人意料,因为影响词汇决定反应时的因素除了频率以外,还有其它因素。De Groot 等^[12]的研究说明,影响双语者词汇决定和词汇命名反应时的因素包括语义、频率(包括频率和熟悉程度)、词长、相邻度(neighborhood density)、成阻(onset)、辅音丛(consonant cluster)等等。根据 Hulme 等^[13]的文献综述,当被试回述高、低频词混合在一起序列时,第一,高频词偶然得到更好的回述^[14];第二,两者常常得到同样好的回述(Ward 等^[15]的实验 2 和 3);第三,低频词有时会得到更好的回述^[10]。这些都说明,频率高不一定就能得到更好、更快的处理。

总之,我们的实验说明相同频率的词不一定能得到相同的反应时,因此启动实验不应该省略基线反应时的收集比较这个重要步骤。

5 参考文献

- 1 Elston-Guettler, K. E. & Friederici, A. D. Native and L2 processing of homonyms in sentential context. *Journal of Memory and Language*, 2005, 52 (2): 256—283
- 2 Schiller, N. O. The onset effect in word naming. *Journal of memory and language*, 2004, 50 (4): 477—490
- 3 Finkbeiner, M., Forster, K., Nicol, J. & Nakamura, K. The role of polysemy in masked semantic and translation priming. *Journal of Memory and Language*, 2004, 51 (1): 1—22
- 4 Hino, Y., Lupker, S. J., Ogawa, T. & Sears, C. R. Masked repetition priming and word frequency effects across different types of Japanese scripts: An examination of the lexical activation account. *Journal of Memory and Language*, 2003, 48 (1): 33—66
- 5 Barry, C., Hirsh, K. W., Johnston, R. A. & Williams, C. L. Age of acquisition, word frequency, and the locus of repetition priming of picture naming. *Journal of Memory and Language*, 2001, 44 (3): 350—375
- 6 Zhou, X-L., & Marslen-wilson, W. Phonology, orthography, and semantic activation in reading Chinese. *Journal of Memory and Language*, 1999, 41: 579—606
- 7 Pexman, P. M., Cristi, C. & Lupker, S. J. Facilitation and interference from formally similar word primes in a naming task. *Journal of Memory and Language*, 1999, 40 (2): 195—229
- 8 Bodner, G. E. & Masson, M. E. J. Masked repetition priming of words and nonwords: Evidence for a nonlexical basis for priming. *Journal of Memory and Language*, 1997, 37 (2): 268—293
- 9 Lisa, S. A. & Cristina B. Neighborhood effect on nonword visual processing in a language with shallow orthography. *Journal of Psycholinguistic Research*, 2004, 33 (1): 75—95
- 10 Overschelde, J. P. V. The influence of word frequency on recency effects in directed free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 2002, 28 (4): 611—615
- 11 Burgess, N., & Hitch, G. J. Memory for serial order: A network model of the phonological loop and its timing. *Psychological Review*, 1999, 106: 551—581
- 12 De Groot, A. M. B., Borgwaldt, S., Bos, M. & Eijnden, E. Lexical Decision and Word Naming in Bilinguals: Language Effects and Task Effects. *Journal of Memory and Language*, 2002, 47 (1): 91—124
- 13 Hulme, C., Stuart, G., Brown, G. D. A., & Morind, C. High — and low — frequency words are recalled equally well in alternating lists: Evidence for associative effects in serial recall. *Journal of Memory and Language*, 2003, 49 (4): 500—518
- 14 Balota, D. A., & Neely, J. Text expectancy and word — frequency effects in recall and recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 1995, 6: 576—587
- 15 Ward, G., Woodward, G., Stevens, A., & Stinson, C. Using overt rehearsals to explain word frequency effects in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 2003, 29 (2): 186—210

The Necessity of Collecting Baseline Reaction Time in Priming Experiments

Dong Yanping¹ Yuan Yuan²

(¹ Faculty of English Language and Culture, Guangdong University of Foreign Studies, Guangzhou, 510420)

(² School of English for International Business, Guangdong University of Foreign Studies, Guangzhou, 510420)

Abstract When conducting priming experiments, many studies replaced the step to collect baseline reaction times (RTs) for the targets by selecting two groups of targets that are not different in frequency, on the assumption that equal word frequency leads to equal RTs. The present study tested this assumption in the task of lexical decision in both the L¹ and the L² of Chinese learners of English. And it was found that for the subjects' L¹, the matching of frequencies for two lists of words did not necessarily lead to equal baseline RTs. Equal baseline RTs were obtained in most cases but not in all if the two lists of words were closely matched (one-to-one match). And if they were words of high frequencies, the situation in the corresponding L² was even more uncertain. The results show that the step to collect baseline RTs cannot be omitted in priming experiments, especially for L².

Key words: baseline reaction time, priming experiment, word frequency