

影响课外读物阅读难度的文字特质

苏启祯

摘要

学生语文能力的培养，有赖于课外阅读。课外读物的选择，除考虑内容之外，语文的深浅也必须适当控制。文本的深浅，由于缺乏使用的工具，多依赖教师个人的语文要求及经验，有嫌过于主观。本文探讨客观分析文本特质的途径，并建议采用最近建构的华文第二语文可读性公式。

关键词：课外阅读 语文特质 可读性公式

苏启祯，南洋理工大学新加坡华文教研中心，联络电邮：kaycheng.soh@sccl.sg。

课外读物提供学生操练阅读能力的机会，是增强学生阅读能力的重要工具，它也扩大学生阅读的范围，间接提高阅读的兴趣。因此，课外读物的选择和推介，在正规的课堂教学之外，是语文教师重要的辅助工作。选择课外读物时，除了内容（文字所描述的事物和情态），教师必须考虑文字本身的要求，使学生能够在阅读能力范围内得到课外阅读的好处和乐趣。

要判断课外读物对学生阅读能力的要求，一般教师信赖本身或同事的教学经验。这样的做法，效果会因人而异，因为教师所接触过的学生不同，教师本身的语文能力和对学生的要求也不同，主观成分强，影响所作的判断和决定，是自然必然的事。

以英文来说，阅读公式（reading formula）用来估计读物的难度。从最早 1948 年的 Flesch Reading Ease 指数到较近 2006 年的 Strain Index，有不少于 23 个不同的阅读公式（Bravos, 2010）。这些阅读公式的基本概念大体上相同，大都采用多元回归法，采用数个语文特质为变项，估计读物对阅读能力的要求。例如 Flesch Reading Ease 公式如下：

$$\text{阅读易度} = 206.835 - 1.012 * \text{平均句长} - 84.6 * \text{平均音节}$$

这公式表示，在英文篇章里，-1.012 这系数表示句子越简单，文章就越容易理解。同理，-84.6 这系数表示单字越容易发音，文章就越容易理解。这些应该很合常理。Flesch Reading Formula 将文章分为 7 个等级，得分越高，越容易读懂，例如 90-100 表示非常容易；反之，0-29 表示很容易混乱。

又如 Strain Index，所用的方法非常简单。要估计文章的阅读难度，用文章开始的 3 个句子，计算这 3 个句子所包含的音节总数，将所得的总数除以 10。如果所得的商数在 5.1 以下，文章很容易阅读。如果商数在 15.3 以上，文章就很难了解。

再如 Spache Reading Formula 考虑的是句长和字频两个文字特质。句子越长，而且生僻的字越多，文章越难。Spache 的公式如下：

$$\text{等级} = 0.121 * \text{平均句长} + 0.082 * \text{非常用字与篇幅的比例} + 0.659$$

阅读公式那么多，要估计文章的难度并非易事，也反映研究者的不同重点。可想而知，阅读公式受研究者所采用的文本的影响，关键在于建立公式时所选用的文章是否能代表读者经常接触的文章类型。

华文的阅读公式研究不多，较早的有 Yang (1971) 在美国 Wisconsin 大学的博士论文，所考虑的变项包括难词比率、句数、平均笔画数。后来，孙汗银 (1992) 设计的汉语为母语的阅读公式，变项包括字均笔画数、难词比例、句均字数、句均词数。

像英文一样，有哪些因素影响或决定华文读物的难易，研究者言人人殊，都有个别不同的看法。王蕾 (2008, 页 51) 综合 1971 年至 2005 年的 5 项研究结果，认为纳入公式或者已被证实能够预测汉语阅读材料难度的因素主要有 5 项：字均笔画数、词表 / 词频、词数、句均字数、和句字数。然而，从对外汉语的观点看，王蕾 (2008, 页 50) 指出，在汉语作为第二语言教学领域中，虽然 20 世纪 90 年代新加坡华文教学领域中已经有人提出可以使用“可读性公式”衡量华语文本的阅读难度 (谢译文, 1994; 苏启祯, 1994)，但至今为止尚未出现一个较为客观的、可以对文本进行量化的可读性公式，这方面的相关研究也非常匮乏，研究成果甚至比汉语母语领域还少。

宋曜廷等人 (2013) 采用 386 篇各年级教科书中的文章，采用逐步回归分析 Stepwise multiple regression 和支向量机 Support vector machine 两种不同的分析，然后再以另外的 96 篇文章检验所得的模式，并加以比较。结果显示逐步回归的分析的，难词数、单句数比例、实词频对数平均、和人称代名词是重要的预测变项。至于支向量机分析的结果，重要的预测变项有难词数、二字词数、字数、和中笔画字元数等。逐步回归分析的准确性为 55.2%，而支向量机的准确性则高达 72.9%。此外，两种模式预测低年级文章比预测高年级文章准确。这项研究也同时反映采用新近电脑科技处理阅读难度的效益，以后的更进发展值得拭目以待。

最近，程勇等 (2020) 以中小学语文教材语料库为数据基础，分别从字、词、句子、篇章等影响阅读难度的因素进行了统计分析，主

要涉及到频率、长度、词类、词义类别、等 17 类 53 种不同因素，构建线性回归方程去检验这些因素与难度级别的预测效度，筛选有效参项并建立文本易读性评价公式，其拟合度达到 0.71，表明字频、词义丰富度、连词比例、动物、动作、关联等词义类比例，以及句长变化度等是区别文章难度级别的关键因素。

这方面的研究在新加坡缺乏。谢世涯与苏启祯（1992）的研究以新加坡高中 2 年级学生为对象，采用辨字测验、选词填充、成语选择、及完形填充收集数据。完形填充测量学生阅读理解能力，以当地华文报章为主要内容。多元回归分析的结果显示，阅读理解的变异，有 29% 由单字预测，23% 由词语预测，5% 由成语预测；3 个语文特质总共预测了阅读能力的 57% 变异。其余未预测的 43% 可能和学生的生活经验、文化知识、阅读技巧、作答态度等有关，有待进一步研究。

文本来源

新加坡学校中，华文为第二语文，进行单科教学。因此，华文课外读物的选择更是棘手问题。新加坡华文教研中心的语料库是为在线华文教师开发的平台，平台的功能包括单字询、词语查询、句型查询、文本分级、全文查询、及教师论坛等（林进展等，2015）。

教师只要输入打算要用的文章，平台就能根据原本分析过的语料的等级，将教师所选的文章加以比对，估计选文的等级是否配合目标年级的课文，并进一步提供有关的数据，例如字种、词种，超纲字，超纲词等，方便教师做更有系统和比较科学化的判断，减少主观的影响。

本文所分析的文本是这平台中的 34 篇文章。它们是从课外读物选出的故事，文体基本上是记叙文，内容大部分涉及虚构的事故，有小部分涉及人生、历史、家庭、和环境，例如《乌鸦和百灵鸟》、《鸡学游泳》、《老办法不管用了》、《一块大，一块小》等。

平台根据新加坡的华文第二语文水平，将所用的文章分为 5 级；

从小一到中四，每两学年为一等级，如小一和小二为等级 1，等等。平台提供每篇文章以下数据。表 1 展示各级文本的平均数和标准差。

1. 篇幅：文章的总字数；有随等级加长的趋势。
2. 字种：文章采用不同字的数目；有随等级增加的趋势，但等级 3 和等级 4 相同。
3. 词种：文章采用不同词的数目；有随等级增加的趋势，但等级 3 和等级 4 非常接近。
4. 句数：文章所有的句子数目；有随等级增加的趋势，但是等级 3 和等级 4 反而降低。
5. 句长：文章句子的平均字数，有随等级增加的趋势，但是变化不大。
6. 超纲字：超出等级水平的字的总目；有随等级下降的趋势，但是变化不大。
7. 超纲词：超出等级水平的词字的总；变化没有明显的规律。

表 1：各级文本的平均数和标准差

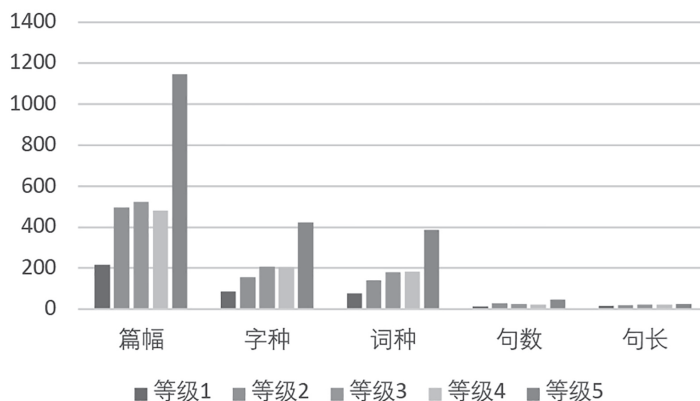
		篇幅	字种	词种	句数	句长	超纲字	超纲词
等级 1 (8 篇)	平均数	217	87	76	12	17	6	13
	标准差	100	30	29	6	5	3	6
等级 2 (8 篇)	平均数	495	156	141	28	19	6	16
	标准差	396	64	66	28	4	7	11
等级 3 (9 篇)	平均数	525	207	181	25	23	7	22
	标准差	305	51	54	21	4	5	8
等级 4 (4 篇)	平均数	482	206	184	22	23	2	9
	标准差	199	56	56	12	5	2	6
等级 5 (5 篇)	平均数	1,146	423	387	48	25	4	23
	标准差	411	112	118	21	5	2	10

注：为方便教师使用，数据都用整数。

图 1 显示，篇幅在第二、三、四级为第一级的双倍，而第五级突然大幅度增加。这现象也出现在字种和词种。其他两项（句数和句长）

变化不大。

图 1. 变项的变化趋势



相关系数

文字的特质对文章难度有怎样的影响，反映在两者之间相关程度的大小。从表 2 可以看出，等级和 7 个变项中的 5 个有中等到高度的相关，包括篇幅、字种、词种、句数、和句长。相关系数在 $r=.414$ 至 $r=.810$ 之间，表示文章难度有 17% 至 66% 因这些变项而起。值得注意的是，篇幅、字种、词种、句数之间有相当高的相关，但是句长和这 4 个变项相关程度很低。

至于超纲字和超纲词，它们和等级的相关很不大 ($r=-.204$, $r=.231$)，表示文章的难度只有约 5% 受这两个变项的影响。超纲字有随等级下降的趋势，因此，相关系数为负数。此外，超纲字和句长有可观的相关 ($r=.456$)，而超纲字和超纲词也有可观的相关 ($r=.658$)。这些相关系数表示句数多的文章有多用超纲字的倾向，而且，超纲字多了，超纲词也跟着多了。

表 2：相关系数

	等级	篇幅	字种	词种	句数	句长	超纲字	超纲词
等级	1.000	.620	.810	.781	.414	.561	-.204	.231
篇幅		1.000	.930	.945	.938	.081	.284	.581
字种			1.000	.996	.783	.277	.112	.517
词种				1.000	.515	.220	.133	.505
句数					1.000	-.177	.456	.592
句长						1.000	-.201	.176
超字							1.000	.658
超词								1.000

注：黑体的相关系数 $p < .05$ （双尾）。

多元回归

为探讨个别变项预测等级的功能，采用逐步多元回归分析（stepwise multiple regression），由电脑根据变项的预测能力决定输入的先后。最先选入的变项，预测功能最大，以后的逐步减低。如表 3 所示，字种最先被选入，其次是篇幅，再次是句长，然后是超纲字，而最后是句数。至此，电脑停止选用变项，词种和超纲词未被选用。

表 3 中，模式 1 只有字种一个变项，而它单独预测等级的 64.6% 变异。模式 2 有字种和篇幅两个变项，预测等级的 77.4% 变异，增值 12.8%。模式 3 有字种、篇幅、和句长三个变项，预测等级的 80.9% 变异，增值 3.5%。模式 4 和 5 先后选用超纲字和句数，预测的等级变异略有提升，增值为 2.3% 和 1.9%。共 5% 左右。

从以上的结果看来，显然，模式 3 可说是既经济又有效的模式，三个变项就能预测等级变异的 81%。据此，用原来的数据，不经过标准化，模式 3 的方程式如下：

$$\text{等级} = -.002 * \text{字种} + .015 * \text{篇幅} + .063 * \text{句长} -.455$$

如果用标准化的数据预测等级，则模式 3 的方程式则改为：

$$\text{标准化等级} = -.678 * \text{字种} + 1.376 * \text{篇幅} + .235 * \text{句长}$$

这方程式意味等级高的文章，总字数（篇幅）虽然较大，但是采用比较少不同的字（字种）。然而，等级较高的文章篇幅比较大，句子也比较长，反映所表达的概念和情景比较复杂。反之，等级低的文章采用比较多不同的字，其篇幅比较短，而且句子也比较短。值得注意的是，从上列标准化数据的方程式看来，篇幅的影响约为字种的 2 倍，句长的 6 倍。

表 3：逐步多元回归分析

	模式 1	模式 2	模式 3	模式 4	模式 5
定数	.896	.611	-.455	-.353	-1.182
字种	.009 (.810)	.019 (1.725)	-.002 (-.678)	.014 (1.268)	.017 (1.545)
篇幅		-.003 (-.984)	.015 (1.376)	-.002 (-.462)	-.005 (-1.419)
句长			.063 (.235)	.076 (.284)	.116 (.435)
超字				-.030 (-.206)	-.041 (-.281)
句数					.049 (.779)
校正 R ²	.646	.774	.809	.832	.851
增值	.646	0.128	0.035	0.023	0.019

注：标准化回归系数 Standardized coefficients 在括弧内。

优势分析

鉴于由字种、篇幅、和句长组成的模式 3 能预测等级变异的 80.9%，这模式显得精简可用。如果增加超纲字和超纲词，成为模式 5，也只能多预测等级变异的 4%。然而，由于这 3 个变项之间有相关（表 1），而且篇幅和字种的相关系数居然高达 $r=.930$ ，3 变项的相对预测效能有进一步检查的必要，因此，进行优势分析（dominance analysis），其确定个别变项的确实预测效能。

优势分析是由 Budescu (1993) 创用的。它从单一变项的预测效能开始，逐步增加变项数目，直到所有被考虑的变项都纳入多元回归模式为止。进一步计算每个新添的变项增加的预测轨范变异有多少。最后，计算每个变项的平均预测效能。因为每个变项的预测效能会受同

模式中其他变项的影响，优势分析把这点考虑在内，所得的结果也就比较可靠。

如果只用一个变项预测等级，字种单独能预测 64.6%，篇幅单独能预测 36.5%，而句长单独能预测 26.3%。这样的结果。似乎和以上方程式所示不一致。但是，这里是每个变项个别考虑，不受变项之间的相关所影响。

如果用两个变项预测等级，字种和篇幅联合能预测 77.4%，字种和句长联合能预测 76.5%，而篇幅和句长联合能预测 62.4%。这样的结果显示，增加一个变项，能预测的等级变异有明显甚至很大的增加，在 4% 至 19%。

如果用 3 个变项预测等级，字种、篇幅、和句长联合预测等级变异的 80.9%。这变异被分解为字种预测 52.4%、篇幅预测 26.3%、和句长预测 21.3%。这 3 个变项的相对预测效能是字种的影响最大，其次是幅度，最后是句长。值得注意的是，字种的预测效能等于篇幅的 2 倍，是句长的 2 倍多，而篇幅的预测效能只是比句长略大，而非如前的 6 倍。

优势分析考虑变项之间的相关程度，所显示的比例比较可靠。因此，以上的结论比一般多元回归的结果准确，比较可以信赖。

表 4：优势分析

		增值		
R ²		字种	篇幅	句长
k=0 平均数		.646	0.365	0.293
字种	.646		.128	.119
篇幅	.365	.409		.259
句长	.293	.472		
k=1 平均数		.441	0.230	0.189
字种，篇幅	.774			0.035
字种，句长	.765		0.044	
篇幅，句长	.624	0.185		

[接上表]

K=2 平均数		0.185	0.044	0.035
字种, 篇幅, 句长	.809			
总平均数		0.424	0.213	0.172
百分比		52.4%	26.3%	21.3%

应用

李教师看中伊索寓言《狮子与老鼠》，打算给她的小学三年级学生当课外的补充读物。首先，她认为内容适合学生的兴趣，而且有教育意义。这还不够，她还要考虑文字的难度。第一个方法是用下列公式计算等级：

$$\text{等级} = -.002 * \text{字种} + .015 * \text{篇幅} + .063 * \text{句长} -.455$$

她需要 3 个数据：

字种：计算文章用了多少个不同的字。(115)

篇幅：计算文章总共用了多少字。(203)

句长：计算每个句子的字数，然后计算平均数。(8)

然后，把这三个数据套入公式中：

$$\text{等级} = -.002 * (115) + .015 * (203) + .063 * (85) -.455 = 2.9$$

所得的可读性是 2.9。这表示《狮子与老鼠》适合二年级末（很接近三年级）的学生。于是，李老师放心她的三年级学生能够读懂。假如李老师的原意不是把它当作课外读物，而是要把它作为理解测验的文本，根据内容设计几个选择题，这也是适合的。

上列可读性公式是用新加坡华文为第二语言课文设计的，但研究结果，第二语文和第一语文之间的差距为 2.0 年（苏启祯，2020）。因此，如果选文是要给华文为第一语文的学生读，可从公式 [1] 所得的结果减去 2.0。上例公式所得的 2.9，减去 2.0 得 0.9，表示所选的文章适合华文为第一语文的二年级第一学期学生作为补充读物。

第二个方法是考虑个别的文字特质。假设教师要小 3（等级 2）的学生阅读，采用表 1 的平均数和标准差，设计评估表（表 4），以不超过（平均数 + 标准差）为准。但是，这方法只适合华文为第二语文。

表 5：文章难度的评估

文字特质	文章数据	平均数 （标准差）	结论
篇幅	205	495 (396)	低于轨范；可用
字种	110	156 (64)	低于轨范；可用
句长	15	19 (4)	低于轨范；可用

如表 5 所示，教师所选的文章，在篇幅、字种、和句长这 3 方面，都低于等级 2（小 3 至小 4）的轨范（标准），因此，适合小 3 学生阅读。假如某个数据高于轨范而差数不大于标准差，仍然可用；例如教师所选的文章，假如篇幅是 538，这比平均数 495 高出 43，这差数小于标准差 (396)，因此，篇幅还是合适。余者类推。

这两个方法，各有千秋。前一个提供综合的数据，用起来比较方便，而后一个提供细节，计较精细。

结语

课外阅读对学生的语文学习有重要的功能，它提供练习的机会，充实阅读的内容，使学生的语文能力发展不局限于课堂上所学的课文，从而建立真正使用语文的能力。除了内容之外，课外读物的选择还必须照顾到学生的阅读能力，文字太浅和太深的课外读物都难以使学生感到兴趣。教师选用课外读物，要考虑文字本身的特质，必须尽量客观，因此，也就需要评估课外读物语文水平的科学化方法。

华文在这面，由于过去的研究不多，尚无定论，使文章水平的评估成为棘手的工作。本文根据新加坡华文教研中心的语料库中的资料，分析结果显示字种、篇幅、和句长对文章的等级有预测的效能。华文教师要评断补充材料的语文要求，可从文章的这三个文字特质着手，做到虽不中也不远的地步。

参考资料

- 程勇、徐德宽、董军 (2020)：基于语文教材语料库的文本阅读难度分级关键因素分析与易读性公式研究，《语言文字应用》，1(113)，132-143。<http://doi.org/10.16499/j.cnki.1003-5397.2020.01.014>
- 林进展、赵春生、洪瑞春、吴福焕、和王志豪 (2015)：基于语料库的新加坡华文教学资源平台开发，《华文学刊》，13(1)，1-15。
- 宋曜廷、陈茹玲、李宜宪、查日稣、曾厚强、林维骏、张道行、和张国恩 (2013)：中文文本可读性探讨：指针选取、模型建立与效度验证，《中华心理学报》，55(1)，75-106。
- 苏启祯 (1994)：华文阅读难度的测定，辑于谢泽文编《新加坡华文教学论文集》，北京，北京语言学院出版社。
- 苏启祯 (2020)：新加坡与中国中小学华文课文语文水平比较，《中文教学现代化学报》，9(9)，100-106。
- 孙汗银 (1992)：《中文易懂性公式》，北京，北京师范大学硕士论文。
- 王蕾 (2008)：可读性公式的内涵及研究范式——兼议对外汉语可读性公式的研究任务，《语言教学与研究》，6，46-53。http://jwc.tsu.edu.cn/sdqcg/lunwen/08/2008/2008_06_07.PDF
- 谢世涯、苏启祯 (1992)：认字辩词与于都理解能力，《语言教学与研究》，4，128-140。
- 谢泽文 (1994)：阅读难度的测量，辑于谢泽文编《新加坡华文教学论文集》，北京，北京语言学院出版社。
- Bravos, K. (2010). Readability tests and formulas. Ideosity. Retrieved January 14, 2021, <http://www.ideosity.com/ourblog/post/ideosphere-blog/2010/01/14/readability-tests-and-formulas>
- Budescu, D. V. (1993). Dominance analysis: A new approach to the problem of relative importance of predictors in multiple regression. *Psychological Bulletin*, 114, 542 - 551.
- Yang, S. J. (1971). *A readability formula for Chinese language*. [Doctoral dissertation, University of Wisconsin].

Language Features Influencing Readability of Extra-class Reading Materials

SOH, Kay Cheng

Abstract

The development of students' language proficiency depends on extra-class reading. Besides the content, the language demand of such materials needs be properly controlled. Due to lack of objective evaluation tool, the language demand is often judged subjectively relying on the teacher's experience. This study explores an objective method of using language features and suggests a readability formula for Chinese as a second language.

Keywords: extra-class reading, language features, readability formula

SOH, Kay Cheng, Singapore Centre for Chinese Language, Nanyang Technological University, Singapore.