

中文可讀性公式試擬

陳世敏

綱要

第一章 緒論

第一節 研究動機

第二節 什麼是可讀性

第三節 什麼是可讀性公式

第二章 中文可讀性公式的要素

第一節 英文可讀性公式的要素

第二節 中文可讀性公式的要素

第三節 句的長短

第四節 字的難易

第五節 常用字標準

第三章 中文可讀性公式

第一節 句和字的研究分析

第二節 中文可讀性公式

第三節 中文可讀性公式實際應用一例

第四章 結論

第一節 可讀性公式的評價

第二節 中文可讀性公式的檢討和建議

第一章 緒論

第一節 研究動機

最先注意到可讀性(readability)研究的，是教育工作者。在一九三〇年代中期之前，美國的一些教

中文可讀性公式試擬

育工作者，設計了各種不同的方式，用計量的方法研究兒童讀物。其中若干較爲科學的計量方法——也就是可讀性公式（readability formula）——曾經引起了相當大的注意力。大眾傳播學者施蘭姆（Wilbur Schramm）即認爲，到一九五六年結束的二十年間，可讀性公式確曾影響了許多人的寫作，使得寫作者開始注意讀者的閱讀興趣和閱讀理解的問題。雖然，公式本身未盡完美（註一）。

可讀性公式有什麼重要性？

一九三〇年代中期的美國報紙，新聞的導言，字數通常是六、七十個字，有時候多達一百以上。原因是：記者寫稿，把新聞的五W和一H擠在導言裏成了習慣，同時這也是記者的一種自衛方式（Self-preservation）（註一）。賀亨柏（John Hohenberg）說：『五W一H擠在導言裏，編輯無法抱怨記者遺漏了那一項。如果敵對的報紙有了不同的標題，記者可以辯稱，那一點，在他的導言裏有了，沒有拿這一點做標題，是編輯的責任。』

但，今天的情形不同了。新聞的導言，很少超過二十個字或二十五個字，有時候更短；而且，導言裏最重要的多半是「誰」和「爲什麼」這兩個因素（註三）。

這是第二次世界大戰以來，專家呼籲提高報紙可讀性，報紙新聞的最大改變。新聞寫作趨向使用短句、短句和短段落，使得受過九年教育程度的人看得懂。美聯社（The Associated Press）在一九五八年成立了「美聯社寫作委員會」（The AP Writing Committee）聘請傅萊區（Rudolf Flesch）爲顧問，介紹他的可讀性公式。合衆社（The United Press）繼之。

根據麥杜葛（Curtis D. MacDougall）的說法，美聯社在傅萊區介紹了可讀性公式之後，新聞句子

的長度，由每句二十七個字降爲二十三個字；導言的長度，由一點七四個音節，降爲一點五五音節。合衆社在接受甘寧（Robert Gunning）的建議後，簡化了新聞的句子，使得新聞適合於受過十一點七年教育程度的人閱讀；以前，合衆社的新聞，則需有十六點七年教育程度，才能看得懂（註四）。

在這兩大通訊社的倡導之下，新聞寫作趨向「可讀」（readable），一時蔚爲風尚。今天，我們很難在一般的新聞報導裏發現擠滿了五W一H的導言。而報紙銷數始終不斷增加。原因固然很多，但是新聞更爲可讀這個因素，是不能夠忽略的。如果我們假設「報章體」（journalese）源於新聞界應用可讀性公式改良新聞寫作，似乎不爲過。可讀性公式影響了新聞寫作，這是第一點。

可讀性公式的影響還不止此。新聞界人士開始以可讀性公式研究新聞讀物與讀者的關係，甚至研究報業歷史。這一類的實用研究，數量至少在數百種以上，下面是兩個例子：

——莫奈特（James Moznette）和拉瑞克（Galen Parick）發現美國報紙一般的國外新聞，需有十四年的教育程度才看得懂，社論則僅需十二年的教育程度，這與一般人認爲社論較新聞難讀，是因為讀社論需要較高的教育程度這種觀念有異（註五）。他們的發現，跟一九四六年「編輯人和出版人」的調查結果相符。

——史蒂文生（Robert L. Stevenson）研究美國保守派報紙和激情派報紙的可讀性的演變，發現保守派報紙的可讀性，一直在穩定的提高；激情派報紙的可讀性，却隨黃色新聞的發展而提高，也隨黃色新聞的沒落而降低。不過，激情派報紙的可讀性，一直是高於保守派報紙（註六）。

美國新聞界人士的研究，從上述例子看，都具有極大的啓發性。簡單的說，這是新聞工作者走出「

個人報業」，從事「知己知彼」的工作的開始。再加上其他的研究成果，新聞工作者研究可讀性的主要發現，大體可以歸納成兩點結論（註七）：

第一、報紙或其他媒介的內容，對他們的閱聽人是太過艱難了。

第二、傳播內容的可讀性愈高，讀者率（readership）也愈高。

寫作，只是整個傳播過程的一部份，寫作者，包括新聞資料的寫作者，都不能夠操縱傳播過程的對象——讀者。新聞從業員必須瞭解他的讀者，一方面又必需設法力求自己筆下的文字，爲人所瞭解。讀物本身如果不可讀（unreadable）則不論他的立論如何驚世駭俗，他的內容如何精采動人，都不會達到傳播的效果（註八）。

對新聞界而言，問題並不這麼單純。美國第一個報團的創始人，也就是現在合衆國際社（United Press International）的鼻祖史克利浦斯（Edward Wyllys Scripps）說過：『我們唯一的事，就是爭取讀者。』

史克利浦斯的話，跟另一位報人派克（Robert E. Park）的見解相當。派克說：『報紙的爭生存，就是爭發行。』

新聞讀物的可讀性與讀者率，有積極的相關，已由克萊爾（George R. Klare）鮑爾斯和卡爾（Richard D. Powers and Bryant Kears）艾倫（Charles L. Allen）（註九）等人的研究，獲得了證實。

測量可讀性固然有許多方法（例如，憑個人經驗判斷），但是較客觀、較正確可靠的途徑，是利用

可讀性公式。到目前爲止，國內新聞界還沒有人利用這種客觀而有效的方法，去測量印刷讀物的可讀性，也似乎沒有重視可讀性與報紙發行數量的關係。

可讀性公式影響了報紙的發行量，此其二。

其次，儘管讀物應該深奧典雅，還是通俗易讀，仍多爭辯，但新聞應該通俗易讀，似已成爲定論。這一目標是沒有止境的。國際新聞學會（International Press Institute）在一九五三年出版的「新聞的流通」（The Flow of News）一書中，認爲目前的情況尙不能令人滿意。書中呼籲各國新聞從業員，在增加人情味新聞之外，新聞寫作仍須力求簡單易讀，以增加國際傳播的效果。

已知的英文可讀性公式，至少有幾十種，還有幾種其他語文的可讀性公式。這些公式的研究，始於教育工作者，繼由新聞工作者加以發揚光大，它的應用範圍，却已經擴大到包括下列各領域（註一〇）：

- 一、教育。
- 二、工商業。
- 三、新聞學和大衆傳播學。
- 四、法律和政府文件。
- 五、心理測驗和問卷。
- 六、寫作。
- 七、演說。
- 八、外國語文。

就新聞學（或泛稱大眾傳播學）這一項來說，十四年前，施蘭姆即曾指出，測量可讀性是新聞學三個專有的研究領域之一——其他的兩項是讀者率調查和內容分析。除此之外，新聞學的研究，均係承襲或借用鄰近學科（Neighboring disciplines）的研究方法（註一一）。

這三個領域的研究工作，在我國，都還停留在萌芽階段。因此，徐佳士呼籲：除了基本研究外，新聞界和新聞教育機構應該加強中文報業實際問題之研究。他認為可讀性公式，是新聞學術頗為重要的「基本研究」之一，此其三。（註一二）。

可讀性公式的研究，對於中文報業還有特殊的意義。

中文的新聞寫作，在慶幸脫離文言而採用白話之際，不幸却結合了歐化句法，形成今日所謂的「報章體」。這種「報章體」跟新聞學上崇尚清晰、簡潔、通俗和標準化等原則的「報章體」不同。它的最特色，可以用一句話來概括：用字不文不白，亦文亦白；文體是不中不西，亦中亦西！

這種現象，不但在國際新聞版上的外電見之，也在國內新聞版上記者寫的新聞見之，甚至出現在廣播和電視的新聞報導中。

在編輯桌上，在新聞採訪寫作課上，沒有人提到這個問題嗎？

有的。中央日報的「編寫手冊」，關於寫作部份，繼「一般規則」後，即直率指出不要受傳統「報章體」的束縛，接著指出了段落、句子、用字所需注意的事項。

彭歌在「新聞文學」一書中，對於報紙要辦得成功，「它不僅要使它所報導的新聞能够被讀者瞭解，而且要使讀者不致誤解」，指出了幾個原則。開宗明義第一原則是：在新聞寫作中，我們應該用大多

數人容易瞭解的字。

我們拿上述的標準，來衡量時下一般的外電翻譯和新聞寫作，當中顯然還存在著相當大的差距。原因是：第一、舊有的習慣牢不可破；第二、寫作者缺乏客觀的標準。換句話說，新聞從業人員對於衝破「報章體」的束縛，只有泛泛的原則，而沒有具體可行的規格。倘若能夠訂出客觀而具體的標準，或許能夠拋棄舊有習慣，解決「報章體」的部份問題，使得新聞報導更為可讀。

這個標準必需是簡單易行，讓動筆寫新聞的人，使用這樣的標準來衡量他的文章的時候，不需費很多的时间和很大的力氣。

一個實用的可讀性公式，或許能夠達到上述的要求。

為什麼一定非用可讀性公式不可呢？對於語言文字，硬生生地用數字的標準去衡量，豈不是矯枉過正？

傅萊區在「說白話的藝術」(The Art of Plain Talk)一書裏，有一段話說得十分透徹，值得引述以釋疑(註一三)：

『一個談話者或寫作者，最重要的事，莫過於讓他的閱聽人能夠瞭解。關於這一點，你只能憑猜測。……由於你只是猜測，你無法知道你猜測得對不對。……因為你沒有標尺(Yardstick)。』

可讀性公式對寫作者的最大貢獻，就是提供了客觀的標尺。本文即打算替中文寫作者，設計這樣的一根尺。

第二節 什麼是可讀性

一、廣義的可讀性

克萊爾 (George R. Klare) 在一九六四年出版的「可讀性的測量」(The Measurement of Readability) 一書中，為可讀性下了下列的三項定義：

- 一、筆跡和印刷的清晰度。
- 二、由於興趣——價值 (interest-value) 或作品生動，以致易於閱讀。
- 三、由於寫作風格 (style of writing) 所導致的易於瞭解和易於理解的情況。

克萊爾指出，晚近的可讀性研究，偏重在第三項，以可讀性公式測量理解程度。根據克萊爾的記載，葛瑞 (W. S. Gray) 和賴利 (B. F. Leary) 曾經分析了影響一篇文章可讀性的因素，共得六十四種，他們刪去了其中若干較不重要、較不易測量的因素，採用了四十四種。這四十四種影響可讀性的因素，歸類起來，不外是：

- 一、內容：文章寫些什麼。
- 二、外觀：包括墨水、紙張、印刷、插圖等項。
- 三、組織：就報紙而言，指的是標題、小標題、字體和嵌條的種類。
- 四、風格：指寫文章的方式。

卡爾 (B. F. Kears) 進一步指出，可讀性的研究，除寫作風格外，尚應包括 (註一四)：

- 一、文章的意見趨向。
- 二、作者的個性和目的。
- 三、出版物的外觀。
- 四、讀者的期望——他期望這篇文章是沉悶還是有趣。

此外，施蘭姆認爲用圖片的目的，是在配合新聞，增加其正確性和可讀性。徐佳士更提出了改進中文報紙版面以增加可讀性的詳細辦法（註一五）：

- 一、減小欄數。
 - 二、統一標題字體。
 - 三、標題形式簡化。
 - 四、加大行與行間的空白，減少不必要的線條和減少廣告套色。
- 這些措施，無非在減少版面的「噪音」，提高可讀性和增加傳播效果。
- 從上述幾位學者的意見，我們擬爲廣義的可讀性，歸納出兩項內涵：
- 一、文章外型因素，這是屬於機械方面（mechanical）的因素，是印刷與編輯上的問題，訴諸人們的視覺，包括了版面、字體、印刷、紙張、油墨、標題、嵌條等外型的條件。
 - 二、文章本身的因素，這是屬於非機械方面（non-mechanical）的因素，純粹是寫作的問題，包括了內容是否有趣和風格是否可讀等條件在內。

二、狹義的可讀性

在大多數字典裏，可讀性的意義，是「容易了解」和「讀起來有趣」。前者是讓人唸起來不費力，後者是能使人繼續讀下去，而樂在其中。

然而截至目前為止有關可讀性的研究，幾乎全部是研究「容易讀」的問題，而不研究「讀起來有趣」的問題——唯一的例外，是傅萊區在「可讀寫作的藝術」(The Art of Readable Writing)一書中所提出的可讀性公式。

正如克萊爾所說，大多數研究工作認爲，可讀性公式的主要用途，在於預測文字資料的容易了解的程度 (the comprehensibility of written material)。換句話說，可讀性的概念，最接近「容易被了解的程度」。

像絕大多數研究可讀性問題的人士一樣，本文所指的可讀性，只限於寫作風格，不談文章外型的因素，也不談文章本身的內容是否有趣。這些因素不是對可讀性不重要，只因為它們很難用可讀性公式測度出來，很難得到一個客觀的標尺。文章的內容，要怎麼樣才生動有趣，以及那一類內容的文章，最受那一類人的喜愛，中外學者大體上已有共同的結論(註一六)，不另贅。本文所謂的可讀性，是指讀物容易被了解的程度而言。

第三節 什麼是可讀性公式

文章要具有可讀性，有沒有規則可循？傅萊區認爲沒有，不過人們曾經以科學方法分析過句子，使寫作者了解一般美國的讀者，句子應以何種長度最便於閱讀。用這種科學方法所得到的結果，有些人名

之爲「方程式」(equation)，又有人稱爲「方法」(method)、「測量」(measurement)、「技術」(technique)，有的人甚至以「計量的關聯研究」(quantitative associational study)來代替「公式」(formula)一詞，實際上都是一回事。

基於此，本文擬採納克萊爾(George Klare)替「可讀性公式」所下的定義(註一七)：

『可讀性公式是一種測量的方法，用以預測作品風格的困難程度，提供計量的、客觀的估量，而不需去測驗讀者。』

關於本文的討論範圍，還有幾點說明：

一、人情趣味(human interest)是造成可讀的一個重要因素，是許多專家一致的見解。瓊斯(John Paul Jones)甚至主張，記者應多在法院的判決當中，尋找埋藏在法院判決書後面的人情趣味因素(註一八)。

不過，人情趣味的測量，可能比文字的風格的測量要困難得多，可能需要有完全不同的公式。本文有關可讀性的研究，是嘗試性的，故不擬包羅在內。

二、可讀性公式的實用價值，大於學術價值，或許不能夠拿邏輯學、語言學和語意學的標準去分析、度量它。

三、一篇讀物，對於一個人是可讀的，未必對另一個人也是可讀的。可讀性公式只是估計一般人的閱讀材料，對個人間的差異，並不在意。

四、中文可讀性公式打算分析的對象，是現代中國人使用的語言，包括口頭說的和筆下寫的東西，

也包括翻譯的著作在內。文言文和詩詞歌謠，不適用。這一點，就跟舊式英語不適用英文可讀性公式，道理是一樣的。

第一章註解

- (註 一)•Wilbur Schramm, "Twenty Years of Journalism Research," *Public Opinion Quarterly*, Spring 1957, P. 105.
- (註 二)•John Hohenberg, *The Professional Journalist*, 台北虹橋書局出版 1969, P. 73.
- (註 三)•*Ibid*, P.74.
- (註 四)•Curtis D. MacDougall, *Interpretative Reporting*, Macmillan, 1957, PP. 87—88.
- (註 五)•James Moznette and Galen Rarick, "Which are More Readable: Editorials or News Stories?" *Journalism Quarterly*, Summer 1968, P. 319.
- (註 六)•Robert L. Stevenson, "Readability of Conservative and Sensational Papers Since 1872," *Journalism Quarterly*, Spring 1964, P. 203.
- (註 七)•George R. Klare, *The Measurement of Readability*, Iowa State University Press, 1963, P. 95.
- (註 八)•Robert Gorham Davis, *Direct Communication*, D. C. Heath and Company, Boston, 1943, P. 3.
- (註 九)•1. Klare, *op. cit.*, PP. 143—48.

2. Richard D. Powers and Bryant E. Kears, "Readability and Display as Readership Predictors," *Journalism Quarterly*, Spring 1968, P. 117.

(註一〇)·Klare, *op. cit.*, P. 4.

(註一一)·Schramm, *op. cit.*, P. 104.

(註一二)·徐佳士,「報業與新聞教育」,新聞學人一卷四期第一〇頁。

(註一三)·Rudolf Flesch, *The Art of Plain Talk*. Harper & Brothers 1946, PP. 1—2.

(註一四)·Bryant E. Kears, "A Closer Look at Readability Formulas," *Journalism Quarterly*, Autumn 1948, P. 346.

(註一五)·徐佳士,「中文報紙版面應該怎樣變」,報學四卷一期,第六頁至第七頁。

(註一六)·艾偉,國語問題,台北中華局出版,第一三頁至第一四頁。

(註一七)·Klare, *op. cit.*, P. 3.

(註一八)·John Paul Jones, *Modern Reporter's Handbook*, Rinehart & Company Inc., New York, 1949. P. 8.

第二章 中文可讀性公式的要素

第一節 英文可讀性公式的要素

爲了瞭解英文可讀性公式最經常採納的,是那些影響可讀性要素,本文擬分析克萊爾(George

R. Klare)在「可讀性的測量」(*The Measurement of Readability*)一書中所載的英文可讀性公

式，共有卅一種，另有十種是根據某一公式加以修正改造，合計四十一種，都是在一九二三年到一九五九年間發表的（註一）。

表一 可讀性公式要素統計表

數次	種類
6	字體具與字象抽
18	表彙字字用常
18	長句均平
3	數語片詞系介
8	數字單
5	數子句單簡
4	數指度字難
12	數節音
1	數尾字頭字
7	數字字定特
1	數句子
1	數母字
3	數字稱人
16	他其
103	計合

從表一可以看出：英文可讀性公式中，最常用的要素，是一、依照常用字字彙表計算難字或易字的字數，二、求每百字的句子平均長度和三、求每百字多音節字或單音節字的字數。

到目前為止，可讀性和可讀性公式的研究，還脫不出「經驗的」(empirical)範圍。學者決定那幾種要素影響可讀性最大，據此而擬出可讀性公式，完全視個人的看法以爲斷。如何決定適當的可讀性公式的要素，誠是此一公式良窳的決定力量。

從上述分析，句的長短和字的難易，是最受注意的兩項因素。倘若我們爲了得一簡單而實用的可讀性公式，究竟這兩項因素是否即是最重要的因素？這兩項因素，是不是決定中文可讀性高低的重要因素

？這是草擬中文可讀性公式之前，必需解決的問題。

第二節 中文可讀性公式的要素

上面說過，可讀性的研究，主要是研究閱讀理解力的問題。爲了明瞭一篇文章的可讀性，用客觀的方法求得計量的資料，首先必須解決那些因素影響了中文的可讀性，這些因素中何者能夠用計量的方法表示，其中那些因素最爲重要等等問題。

不過，分析影響中文可讀性因素的工作，十分浩繁，本文擬採取另一種較爲可行的方式，以分析英文可讀性公式的要素，作爲研判中文可讀性公式應有那些因素的參考。正如傅萊區（Rudolf Flesch）表示，各種語文，基本上都有相通之處，如果一種可讀性公式能夠測量一種語文，則亦必能夠測量別的外國語文（註二）。

影響可讀性的因素，可以大別爲兩類（註三）：

一、閱讀材料方面。

二、讀者方面。

（一）直接因素：包括認字的速度和記憶範圍。

（二）間接因素：包括教育程度與對某類讀物的特別經驗，以及學習動機（set to learn）和最省力原則（principle of least effort）。

由於讀者特性千殊萬別，我們只能夠假定讀者的條件完全相同，然後能夠研究不同閱讀材料的可讀

性。換句話說，可讀性的問題，只限於材料方面的問題，與讀者方面的主觀條件無關。

大體上說，除了字頭字尾和音節等少數情形外，中文也可以分析成與英文相似的因子，雖然，這不一定表示某一個因子對中文的可讀性和英文的可讀性，有同樣的左右力。例如，漢字的筆劃多寡，約略與英文單字字母的多寡相當，但是一個英文長字對英文可讀性的影響力，是否就跟一個漢字筆劃多的字一樣？

這純粹是語言特質的問題，也是中文不能夠直接借用英文可讀性公式來測量中文可讀性的癥結所在。楊孝濬的中文可讀性公式，曾經考慮計算「對稱字」（例如，「晶」「朋」等字），就是基於中文特質的理由。再舉個例子，就可以知道中文特質的複雜性：

中文是不是單音節 (monosyllable) 字？

很多外國人說是，很多中國人也說是。耶魯大學的甘迺迪 (George Kennedy) 表示反對，他說中文完全是個「謎團」(myth)，所以應稱之為「單音節謎團」(monosyllable myth)。趙元任在中央研究院歷史語言研究所成立四十週年紀念會演講，曾經進一步詮釋了這個問題。他認為中文確是「單音節謎團」，因為，事實上，你要看某一個音節是能不能夠獨用(註四)。比方說，「長生果」、「降落傘」這一類的三音節字，分開使用，成什麼話？

如果中文確是「單音節謎團」，我們自然可以問：中文的音節，跟可讀性有沒有關係？

中文「音節」與可讀性的關係，也是有待來人解答的一個問題。

中文的這些特質，依常識判斷，可能或多或少影響了中文的可讀性，但是，除非某一特質極具左右力，否則，似不應列入一個實用的可讀性公式中。英文可讀性公式由繁而簡，裁汰較不重要的因子，考慮到實用，是最大的關鍵。只要把握住最重要的因素，即可達到可讀性公式用來「粗略估計文章的可讀性」的目的。

基於此，本文假定句子的長短和用字的難易，是影響中文可讀性的兩個最重要因素。

第三節 句的長短

中國自古，語文殊軌，文字變成另一套獨立的表達符號。自「我手寫我口」的運動倡議語文結合以來，一般人的造句、用語趨於繁冗。言曦說（註五）：

「……………兩百餘年前紅樓夢中寶玉向黛玉傾訴的一段長篇自白，又何其簡潔、鏗鏘而有力？（他每句話都不超過十個字），乃至民國初年的巨著擊海花中的對話，亦尚保留簡淨爽利的特色。降至近廿年來所見的小說，則多繁冗不堪卒讀的對話，作者固疏於剪裁擷取，但多少亦反映現代人說話的方式在向拼音文字的「自找麻煩」的方向變。實際上，這樣做反而貶損了中國語文原有的表達力。」

中國語文，句子由短而長，結構由簡而繁，是今日傳播效力不彰的原因之一。動筆寫作的人沒有一個客觀的「標尺」，不知道句子長到什麼程度，便會使人難以卒讀。傅萊區研究了英文句子的長度跟文章難易的關係，然後提出下面的結果（註六）：

非常容易

八個字以下

容易	十一個字
還算容易	十四個字
標準句子	十七個字
還算困難	二十一個字
困難	二十五個字
非常困難	二十九個字以上

傅萊區的標準，並非全然不可更動，但是他認為，一個句子以十九個字為宜，越長的句子越不好懂，反之，則易懂。

薛曼(L. A. Sherman)在一八八〇年主持英國文學研討會的時候，發表報告。略謂，伊麗莎白時代的英國文學著作，句子平均的長度約為五十個字，到他的時代，句子長度已降為二十三個字(註七)。薛曼同時指出，短句不一定因為短而容易懂，長句也不一因為長而混淆。不過，一般言，長句確是理解的一個障礙。魯巴金也有同樣的看法。

莫奈特(James Moznett and Galen Parick)等人發現，不論是新聞報導還是社論，最初的一百個字，要比最後的一百個字難讀，因為導言通常太長、太複雜。

根據于宗先在民國四十八年的研究，台灣地區學生了解新聞的程度，六年級學生平均為百分之六十六，九年級學生為百分之八十八，十二年級學生為百分之九十(註八)。這個數字，如果除去可讀性極低的經濟新聞，則台灣地區學生了解新聞的程度，以十二年級學生為例，至少達百分之九十七以上。

如果這項調查能够代表台灣報紙的新聞的可讀性，那麼，我們似可斷定：十二年前的台灣報紙的新聞可讀性，似不在外國著名通訊社之下。

可惜，調查中沒有說明用來測驗的新聞實例，句長多少，難字若干，無法得悉台灣報紙新聞報導是不是也在向拼音文字的方向變。

拼音文字賴文法維持句中各部份的關係，除倒裝句法外，句子長，是拼音文字的一個特色。中文一直沒有縝密的文法結構，對短句的需要，尤其殷切。因為句子一長，上下文的關係便容易弄混，所以近人主張多用句點之外，還主張多用逗點——最好是在平常講話停頓的地方都用逗點。這是增加文章條理和增加可讀性的有效措施，也是中文恢復到「簡潔、鏗鏘而有力」的途徑。

短句有什麼好處？傅萊區說得很清楚（註九）：

『最好的途徑是寫短句，讀者和聽衆因而有時間喘口氣，不致於爲了字裏行間似幻似實的語意關係而迷惑。』

瓊斯（John P. Jones）從另外的角度，進一步說明短句的重要（註一〇）。他認爲一個句子以長而有修飾語的子句或片語開始，有礙於閱讀的了解。如果句子的結構，複雜到讀者必需回頭再讀一次，則文章實無可讀性可言。

結構簡單未必就是短句，但短句通常結構簡單。可讀性的要求，句子方面，要力求簡而短，「像A BC一樣」。上述原則，中文自可適用。

第四節 字的難易

東漢許慎的「說文解字」，有漢字九千餘，到唐朝顏真卿著「韻海鏡源」，得二萬七千字，漢字在六百餘年間，增加了兩倍；民國初年編成的「中華大字典」，單字有四萬五千個。一千八百年間，字量增加了四倍。張其昀等人近年編著的「中文大辭典」，搜羅了四萬九千八百八十八個單字，比康熙字典多出了七千七百一十四個字。

近五萬個漢字中，大都是林語堂所說的「死字」，實際上經常用上的，不過二、三千字。以國立編譯館主編的「國民學校常用字彙研究」爲例，三千八百六十一個常用字，出現次數的百分比，高達百分之九十九點六四（註一一）。換句話說，一般讀物中，非常用字出現的機會，每一百個字裏，只有零點三六個字。

語文既以實用爲主，自宜多用常用字，少用非常用字。但，什麼是常用字？常用字一定是容易瞭解的字嗎？

常用字是人們經常見到、讀到或用到的字（註一二）。提高可讀性，在用字方面，應盡可能使用熟習的字（常用字）外，有人還主張多用簡單的字（中文是筆劃越少越好，英文是音節越少越好）。常用字是一般人熟習的字，大致不會引起爭論，但是簡單的漢字却不一定是一般人熟習的字，例子很多，不需枚舉。

至於一個字是不是容易爲人所了解，跟這個字是不是常用字和簡單字，似乎沒有絕對的關係。例如

，英文「和平」(Peace)這個字，在戴爾—柴爾(Dale-Chall)三千常用字彙表上，應該是四年級學生懂得的。賀亨柏(John Hohenberg)談到哥倫比亞大學的教授們，爲了替「和平」下定義，曾經多次集會討論，前後四年，最後，大部份人勉強同意這樣的定義：和平就是沒有戰爭(Peace is the absence of war.)。簡單字像「遺傳基質」(id)，認識這個字的人，恐怕更少。Peace是常用字，id是簡單字，它們都不是容易了解的字。

特別是簡單字這一部份，值得進一步解釋，以明瞭中英文的差異，同時用以說明本文關於「字的難易」，是以一個字是不是常用爲標準，不以它的外型是不是簡單爲標準。

首先以字的繁簡作爲測量讀物的難易標準的，是貝爾(M. V. Bear)。她發現(註一三)：

一、單音節字的多寡，能够作爲讀物難易的指引(index)。

二、字的長度與這個字的使用頻數，有高度的相關。

魯威格(Merritt C. Ludwig)以「華萊士的農夫」(Wallaces' Farmer)的訂戶爲對象，研究難字與讀者率的關係，發現一則新聞經過改寫，音節較多的一則，比音節較少的一則，讀者率低(註一四)。

傅萊區在一九五八年客席「美聯社寫作委員會」的時候，提醒美聯社記者，「如果可能，切勿使用每天談話中通常用不到的字。記住，美聯社的工作不在於增加人們的字彙。如果你必需使用一個可能爲普通讀者所不熟習的字，你得加以解釋。」

傅萊區主張字的零件要愈少愈好。dis-approve有兩個字頭，等於是文字多了兩個零件。標準的

文章，字頭字尾（*affixes*）以三十七個字爲宜，二十二個以下爲極易讀，五十四個以上爲極難讀。詹姆士（King James）版聖經的字頭字尾數，平均每百字約爲二十個字（註一五），故極易讀。傅萊區後來以音節的多寡代替字頭字尾數，本質上仍然未脫離「字越長越難」的觀念。

以計算音節數作爲可讀性高低的公式甚多。漢字能計算筆劃，沒有拼音文字的音節問題。漢字是不是筆劃愈多愈難？

這個問題，幾十年來的研究，已有定論：答案是否定的。一般言，簡字（筆劃少的字）除了書寫上比較省時外，在學習上，無論是認識、記憶或書寫，都屬繁體字比較容易（註一六）。

本文試擬的中文可讀性公式，在字的難易方面，不計字的筆劃，理由在此。

較可行之道，還是常用字。一個字如果經常被使用，比起冷僻的字眼，它大體上是比較容易瞭解的字。問題是，中文可讀性公式，應該以那一套常用字作標準較好？

第五節 常用字標準

由於每一套常用字彙編製的根據不同，方法不同，利用他們來測量可讀性，必需先明瞭某套常用字彙的來源和特性。

國人中，最先從事漢字常用字彙研究的，是陳鶴琴。民國二十四年，教育部編「小學分級暫用字彙」，計二千七百一十一個字。民國三十五年，蔡樂生在成都進行字彙的研究，目的在編訂一般民衆適用的最基本字彙。他先就歷代字彙遞增的情形、各種適用字典所收字數，以及各家研究字彙的結果，綜合

加以分析。他認為，最常用的兩千字既然佔普通讀物中總字量的百分之九十七，而「三民主義」一書的生字，也只有二千一百三十四個字，足見「兩千個字如果選得好，在社會上已經大致够用。一般民衆如能學會兩千字，那麼就可文盲掃除、教育普及了」（註一七）。

鑒於上述資料陳舊，未必合於時宜，教育部乃在民國五十二年，命國立編譯館成立「國民學校常用字彙釐訂委員會」，聘請文字、語言、教育、心理、統計等方面的專家，從事國民學校常用字彙的研究，歷時四年始告完成。

這是我國常用字研究規模最大的一次，分析的對象，包括報章、國民學校課本、兒童作品、課外讀物、憲法、國歌、廣播稿和民衆讀物，共七十五萬三千九百四十字，得單字四千八百六十四個，以出現頻數十次以上的，列爲常用字，計三千八百六十一個，其他五百七十四個字，列爲次常用字，四百二十九個字列爲備用字。

這項研究，曾就各單字的出現頻數、部首、音順，列表附在書後，極具參考價值。而規模大、選樣具有代表性、統計精細而科學，爲其他常用字彙的研究所不及。

其他的幾套常用字彙，較不重要。必需一提的，是四套新聞常用字彙。因爲，我們希望中文可讀性公式，能夠用來改進目前的新聞寫作，而新聞報導常用的字彙，多少與一般的讀物不同，新聞常用字的研究，乃有事實上的需要。

- 一、「聯合報中文自動鑄排機常用字彙表」。
- 二、「中央日報自動鑄排機常用字彙表」。

三、日本「寫真植字機研究所自動照相排字機中文字彙表」。

四、「世界中文報業協會新聞常用字彙表」。

上面提到的常用字彙研究，都可以作為可讀性公式衡量難易字的標準，但，究竟選取那一套常用字，還需根據實際狀況來決定：

一、各套常用字彙表，由於根據不同，對象不同，致優劣互見。在十全十美的字彙表未出現之前，應用可讀性公式的時候，究竟那一套常用字彙最適用，應視個人的需要。如果是為了測量新聞文字的可讀性，自以後四種新聞常用字彙為宜；如果是為了測量兒童作品的可讀性，則以「小學分級暫用字彙」較好。

二、常用字字量的選取，要能夠顯示測量對象的差異。如果數篇可讀性有別的文章，非常用字（難字）同樣太少或同樣太多，則顯不出文章的難易，失去了以難字多寡作衡量可讀性高低的原意。

基於這兩點考慮，本文擬以「國民學校常用字彙表」作為常用字的標準。一來是國立編譯館的這項研究工作，規模大、選樣具代表性、統計過程合乎科學原則；二來是資料新，可避免若干常用字晚近變為「明日黃花」的弊病；三來是資料的編列排比，有多種方式，參考的時候極為方便。最主要的是，這套常用字適用於「一般公民」，不限於國民學校學生。這一點，可以從樣本的廣泛，得到證實。

剩下的問題是常用字的字數問題。我們擬以「國民學校常用字彙」出現頻數在二十次以上的字，為常用字，計二千三百二十五個字。粗略估計，這二千三百二十五個字的累積出現次數，約佔百分之九十五，換句話說，一般的文章，平均每一百個字約有五個難字，有較多的餘地，以顯示文字難易之別。

克萊爾批評洛奇借用戴爾的七百六十九個字的常用字彙表，因其字彙太少，致預測可讀性較不準確，尤以困難度較高的讀物爲然。所以，常用字量太少，實非所宜。

我們以「國民學校常用字彙」出現頻數在二十次以上的字，爲常用字，除了它較能夠顯示難易程度之別外，另一個用意，是它的字量，與「聯合報自動鑄排機常用字彙表」和「中央日報自動鑄排機常用字彙表」的二千三百七十六個字，只相差五十一個字。中文可讀性公式如果以這兩種字彙表作難易的標準，大致可以用來測量新聞文字的可讀性。

第二章註解

- (註 一)•George R. Klare, *The Measurement of Readability*, Iowa State University Press, 1963, PP.75—80。
- (註 二)•Rudolf Flesch, *The Art of Plain Talk*, P.5.
- (註 三)•George Klare, *The Measurement of Readability*, Iowa State University Press, 1963, P.7.
- (註 四)•趙元任，「中文音節和體裁的關係」，五十七年十二月號綜合雜誌，第一八頁。
- (註 五)•言曦，思印集，文星書店出版，第一〇五頁。
- (註 六)•Rudolf Flesch, *The Art of Plain Talk*, P. 38.
- (註 七)•Kearl, *op. cit.*, P.347.
- (註 八)•于宗先，「台灣報紙可讀性之研究」，報學二卷六期，第二三頁。
- (註 九)•Flesch, *op. cit.*, P.33.

(註一〇)·John P. Jones, *The Modern Reporter's Handbook*, Rinehart & Co., 1954, P.12.

(註一一)·國立編譯館主編，國民學校常用字彙研究，台灣中華書局出版，第三〇頁。

(註一二)·羊汝德，新聞常用字之整理，台北市新聞記者公會出版，第二八頁。

(註一三)·Klare, op. cit., P.32.

(註一四)·Merritt C. Ludwig, "Hard Words and Human Interest: Their Effects on Readership,"

Journalism Quarterly, Summer 1949, P. 171.

(註一五)·Flesch, op. cit., P.42.

(註一六)·杜學知，「簡字難學」，五十七年七月二十八日，中央日報副刊。

(註一七)·國立編譯館主編，國民學校常用字彙研究，台灣中華書局出版，第五頁至第一〇頁。

第三章 中文可讀性公式

第一節 句和字的研究分析

我們已經假定了句的長短和字的難易，為影響中文可讀性的有力因子。從第三章的分析，我們決定以這兩個因子，作為擬議中的中文可讀性公式的要素。目前僅有而十分重要的問題是：如何證明句的長短和字的難易，確能影響中文的可讀性？句和字，在可讀性公式中，應佔怎樣的比重？

研究方法

本文打算先用「變異數分析」(analysis of variance)的方法，分析句和字這兩個「自變數」(

independent variable)，對「閱讀理解分數」這個「應變數」(dependent variable)，是不是如同我們預期的一樣，有相當顯著的影響力。

上面說過，可讀性的問題，實際是「理解力」(comprehension)的問題，只是不計閱讀者的個別差異而已。倘若句和字確曾顯著影響了閱讀理解分數，那麼，我們最後的目的，在藉統計學「迴歸分析」(analysis of regression)的原理，求得「迴歸方程式」(regression equation)。

這個迴歸方程式，就是我們所要的可讀性公式。從方程式中的兩個自變數，可以求得應變數的值——也就是可以從句子的長短和難字的多寡，推斷出某篇文章可讀性的高低。這就是「迴歸分析」。「迴歸分析」是繼「變異數分析」之後，求得中文可讀性公式所必需的途徑。

選擇樣本

一、從現代中國作家的作品中，選出「句長字易」和「句長字難」的文章各一篇。文章的長度，以能表達一個完整的觀念為原則。改寫這兩篇文章，把句子特別長的部份，在合理的地方添上標點符號，增加句子數，成為「句短字易」和「句短字難」兩篇文章。原文跟改寫文唯一不同的地方，是句子的長短。

二、用同樣的方式，選擇「句長字難」和「句短字難」的文章各一篇。改寫這兩篇文章，把難字部份的一半，換以簡單的同義字，亦即減少了原文的難字數，成為「句長字易」和「句短字易」兩篇文章。原文跟改寫文唯一不同的地方，是難字的多寡。

四篇原文的作者和出處是：

- 一、「莫泊桑研究」，費海璣作，摘自國語日報社出版「書和人」第三冊，第四一三頁。
- 二、「梭羅及其不服從論」，彭歌作，摘自純文學社出版「改變歷史的書」第八五頁和八九頁。
- 三、「燕園之憶」，王聿均作，摘自國語日報社出版「書和人」第二冊，第二三六頁。
- 四、「學問與趣味」，梁實秋作，摘自文星書店出版「秋室雜文」第五八頁。

問卷製作

四篇文章經改寫，得八篇文章，分成量表甲和量表乙，每一量表有四篇文章，均打字印妥。量表甲的第一篇文章，和量表乙的第一篇文章，只是原文與改寫文不同，測驗的題目相同。

每篇文章後，附十個四選一的問題。其中第一和第二個問題，是解釋包含有難字的詞彙的意義；第十個問題問「你認為這篇文章按什麼題目最好？」

其他的問題，有答案甚為明顯者。例如——

幫助莫泊桑成爲名作家的是：一、左拉。二、哈代。三、叔本華。四、尼采。

有需相當的思考者。例如——

梭羅的哲學思想：一、強調國家重要。二、強調個人重要。三、強調社會重要。四、強調民衆重要。

大體說來，題目的答案，多半需理解而後可得，一望可知答案的問題極少。

選擇測驗對象

測驗的對象，是在學的初中學生。因爲目前我國國民受教育的平均年齡，大約是七、八年，相當於

初中一、二年級的程度；同時，初中階段是閱讀能力增加最快的時候（註一），他們剛剛脫離國民小學那種有明顯答案的閱讀測驗方式，而理解力還未達到高中學生的成熟狀態，對文字的駕馭未臻穩定，以他們作對象，最能夠看出文章特性對他們的閱讀理解力的影響。

以初中學生為測驗對象的另一個理由是：中文可讀性公式的終極目標之一，在於測量中文新聞寫作的可讀性，甚至用來作為改進新聞寫作的「標尺」。初中程度的讀者，構成報紙、雜誌讀者的大多數。因此，新聞學家和新聞工作者，都呼籲新聞寫作的目標，是要使初中程度的人讀得懂（註二）。我們對於初中程度的讀者，自宜研究他們的閱讀理解力以及字和句給予他們的影響。

為了方便比較，決定以初中二年級的兩組（班）學生作對象。他們的閱讀理解程度，頗難斷定，較客觀的方法，就是以他們的國文程度代表閱讀理解力。經平均三次月考和一次期考的成績後，決定以台北市南門國中二年級兩班女生，為測驗的對象。

結果分析

表二 句和字與閱讀理解分數的關係表

四		三		二		一		別	篇
乙	甲	乙	甲	乙	甲	乙	甲	別	組
6.04	6.25	6.91	8.58	4.86	5.33	6.39	7.95	數	分驗測
330	330	237	237	309	302	307	318	數	字總
33	33	20	20	29	18	19	30	數	字總
10	10	12	12	10.6	17	16.2	10.6	長	句均平
9	18	16	8	12	12	7	7	數	字難
2.76	5.52	6.75	3.38	3.97	3.97	2.2	2.2	比	分百字難
句短字易	句短字難	句長字難	句長字易	句短字難	句長字難	句長字易	句短字易	質	性
改寫	原文	原文	改寫	改寫	原文	原文	改寫	別	類
文明說		文絃記		文明說		文絃記		記	附

從表二四篇文章比較甲乙兩組的測驗分數，可以看出，甲組均高於乙組，可見甲組的國文程度高於乙組。表二顯示，甲乙兩組在測驗分數上的差異，以第一篇和第三篇較為明顯。換句話說，甲組的測驗分數優於乙組，在記敘文方面較為明顯，說明文方面較不明顯。

句字長度和用字難易這兩個「自變數」，對測驗分數這個「應變數」，有沒有顯著的影響呢？表三至表六，是四組問題的變異數分析，經電子計算機處理後，結果如下：

一、第一篇測驗原為「句長字易」，總句數由十九改為三十，成為「句短字易」。用這兩篇文章測驗兩組學生，發現甲組的測驗分數比乙組高出百分之十五點六（見表二），句子的長短，對測驗分數有顯著的影響（見表三）。換句話說，句子愈短，測驗成績愈高。

二、第二篇測驗，目的在驗證第一篇測驗的正確性。乙組的「句短字難」，測驗分數似應比甲組的「句長字難」高；但乙組的理解力較差，乙組的測驗分數雖因短句而縮小與甲組的差距，不過平均測驗分數仍低於甲組百分之五點七（見表二）。兩組的分數差異，在統計上沒有顯著性（見表四四）。除上述甲乙兩組測驗對象本身的差異之外，說明文較記敘文難懂，（不論甲組或乙組，第二篇的測驗成績都低於第一篇，第四篇的成績都低於第三篇），可能是兩組測驗分數沒有顯著性的原因之一。

表六 難字數的變異數分析(四篇)

類別	平方和	平方均數	自由度
組間	0.50	0.50	1
組內	139.45	3.09	45
總和	139.95		46

F 值=0.162
顯著水準=0.690 (不顯著)

表三 句子長度的變異數分析(第一篇)

類別	平方和	平方均數	自由度
組間	28.83	28.83	1
組內	112.43	2.49	45
總和	141.27		46

F 值=11.542
顯著水準=0.001 (顯著)

表七 句長與難字數的總變異數分析

類別	平方和	平方均數	自由度
組間	168.04	168.04	1
組內	631.91	14.04	45
總和	799.95		46

F 值=11.967
顯著水準=0.001 (顯著)

表四 句子長度的變異數分析(第二篇)

類別	平方和	平方均數	自由度
組間	2.52	2.52	1
組內	59.94	1.33	45
總和	62.46		46

F 值=1.896
顯著水準=0.175 (不顯著)

表五 難字數的變異數分析(第三篇)

類別	平方和	平方均數	自由度
組間	32.76	32.76	1
組內	85.65	1.90	45
總和	118.42		46

F 值=17.213
顯著水準=0.000 (顯著)

三、第三篇測驗是「句長字難」，在難字百分比由百分之六點七五降為百分之三點三八後，甲組測驗成績高出乙組百分之十六點七（見表二），難字很明顯的影響了閱讀理解測驗的分數，達到統計上的顯著水準（見表五）。換句話說，難字百分比愈高，測驗分數愈低；難字百分比愈低，測驗分數愈高。

四、第四篇測驗，是在驗證第三篇的正確性。甲組測驗成績僅高出乙組百分之二點一（見表二），同樣的，測驗成績並沒有達到統計上的顯著水準（見表六），原因與第二點說明相同。

表七是就甲乙兩組四篇測驗的總成績，分析句長以及難字百分比與測驗成績的關係。從統計上具有顯著水準，可以知道：句子的長度和難字百分比這兩個「應變數」，顯著影響了「自變數」。從理論上說，這兩個「應變數」，應該可以作為中文可讀性公式的因子。下一步驟，就是用迴歸方程式來表示中文可讀性公式。

第二節 中文可讀性公式

茲根據表二，分析成表八資料：

表八 迴歸方程式分析表

類別	迴歸係數	迴歸係數標準差	偏相關係數	T值	顯著水準
字數	7.81044	2.91112	0.768	2.68297	0.0437

句長	-0.07830	0.19208	-0.179	-0.40766	0.7004
難字百分比	-0.09914	0.33144	-0.133	-0.29913	0.7769

估計標準誤=1.3885

決斷係數=0.0422

複相關係數=0.2055

依表八，我們可以得到下述的方程式： Y （閱讀測驗分數） $= 7.81044$ （常數） $- 0.07830X_1$ （每句平均字數） $- 0.09914X_2$ （難字百分比）

這個方程式的意思是說，倘若本文所作的測驗，具有代表性，則我們知道一篇文章的句長和難字百分比，就能夠推斷出閱讀測驗分數。

這個方程式，就是根據客觀資料求得的可讀性公式。它的有效性（Validity）如何？

遺憾的是，由於缺乏必需的「基本研究」，使得本文的嘗試，並沒有令人滿意的成果。

一、從統計顯示，這一公式的「決斷係數」（coefficient of determination）為百分之四點二二。

換句話說，用句長和難字百分比來推測測驗成績，有百分之四點二二的正確度。

二、「應變數」和「自變數」的交互相關，如表九：

表九 句長和難字百分比與閱讀測驗分數相關表

變數別	句長	難字百分比	測驗分數
句長	1.000		
難字百分比	-0.171	1.000	
測驗分數	-0.158	-0.102	1.000

表九可以看出，三個變數之間，有負的單相關（simple correlation）且相關的程度，均未達顯著水準。三個變數的複相關係數（multiple correlation coefficient），為〇點二〇五五，也沒有顯著性。

由於正確性和顯著性都不高，這個中文可讀性公式，仍有待進一步的研究、修正，才能成爲有效的公式。

但就初步的研究言，本文既已證明了句子的長度和難字百分比，是影響閱讀理解分數的有力因子，則依循這個方向，擬出一個實驗性的中文可讀性公式，似乎仍不失爲值得嘗試的目標——至少，它可避免後人在從事進一步研究過程中，作無謂的摸索。

可讀性公式既以實用爲貴，以能測度真正的難易程度爲目標，則分析影響可讀性的因子，「科學」

和「客觀」之外，「經驗」和「判斷」似乎不能忽略。合乎科學原則而不能不便應用，當非研究可讀性公式的本意。這是本文打算以現有資料，試擬中文可讀性公式的理由。

這一公式，必需達到兩項要求：

一、預測正確：句長和難字百分比既為試擬中的中文可讀性公式僅有的兩個因子，則公式的正確與否，完全繫於因子的係數——也就是因子影響可讀性的比重。在本公式中，句長的係數是○點○七八三；難字百分比的係數是○點○九九一四。這兩個係數，得自測驗分析，故正確性似不成問題。

二、使用方便：為了使用方便，係數的小數位數愈少愈好，用一位數已足，前已述及。我們打算把兩個因子的係數各乘以十，然後尾數四捨五入。句長係數得○點八，難字百分比係數得一。

我們以「教育程度」(grade placement)或「含霧指數」(fog index)來代表可讀性分數的高低。句子短、難字少，則可讀性高，也就是可讀性分數低，理解某篇文章所需的教育年齡低；句子長、難字多，則可讀性低，也就是可讀性分數高，理解某篇文章所需的教育年齡高。公式中僅計句長和難字百分比兩因子，常數項無特殊意義，不予計入。

按上述兩項要求修正後的中文可讀性公式如下：

$$Y(\text{可讀性分數}) = 0.8X_1(\text{每句平均字數}) + X_2(\text{難字百分比})$$

比方說，Y等於七，表示某篇文章，一般受過相當於七年教育的人可以看得懂。這篇文章的可讀性分數就是七。

第三節 中文可讀性公式實際應用一例

下面有個例子說明應用中文可讀性公式的步驟。

一、選樣：從一本書或一篇文章，用隨機或間隔抽樣的方法，選出一段或數段文字，每段的字數約在二百左右，以意思完整的地方為終結之處。如果可能的話，以一本書或一篇文章的全部份，作為樣本，正確性更大。樣本的字數決定以二百字左右，因為一百字太少，誤差大；三百字太多，計算費時費事。要避免選取第一段。避免顯然極為特殊的段落。避免文言或有詩、詞的段落。

本文選取的，是民國六十年五月二日台北各中文日報刊載的一則中央社專電「周外長駁斥麥高文謬論」。全文約七千字，茲以第五段作例子：

「除此以外不可能有別的解釋了。因為撇開其修辭及虛構的史實不談，麥氏三月廿四日所發表的演說，可以用一句話來概括：「放棄自由中國」——不計傳統的友誼或條約義務；不顧道德或正義，抹煞一切原則和聯合國憲章；把中華民國攆出聯合國，將台灣的一千四百八十萬人民交給毛澤東割宰，把中國大陸上仍然寄望從共產暴政下解救過來的億萬人的生路加以杜塞。

替麥氏寫稿的人很清楚，把這些建議直截了當地提出來，公正的美國人，不論其政治信仰如何，必將難予接納。」

二、計算總字數：字數包括標點符號在內，每一個標點符號算一字，但書名號、私名號、括號和引號不計。破折號和刪節號佔兩個格子，仍以一個標點符號計算。本段有一百九十一個字，加上十六個標

點符號，共二百零七個字。

三、計算句數：高聲朗誦時需要停頓的地方，都算一句，也就是有標點符號的地方是一句。引號所在之處，因不需停頓，僅加強語氣，它的功用顯然與本文所謂的「句子」不同，故排除在外，共得十六句。

四、計算每句平均字數（句長）：以句數除總字數，為每句平均字數。得十二點九七字。

五、計算難字數：按照國立編譯館主編「國民學校常用字彙研究」，列出的前二二三二五字，逐字核對。不在常用字表上的字，為難字，得六字。這一部份工作，初看繁雜，如果先把「常用字彙表」大略先看幾遍，則對於絕對有把握的字，自然無需查表。原則是稍一有懷疑就查表，可防錯誤。

這段樣本，有「撇」、「括」、「抹」、「煞」、「攢」、「塞」六個難字。

六、求難字百分比：以總字數除難字數，得百分之二點八九。

七、求可讀性分數：代入 $Y = 0.8X_1 + X_2$ 公式，得 $Y = 0.8 \times 12.97 + 2.89$ ，計十三點二六。換句話說，這篇文章的可讀性分數是十三點二六，一般相當於受過十三點二年教育的程度的人，可以看得懂。

第三章 註解

（註一）：于宗先，「台灣報紙可讀性之研究」，報學二卷六期，第二三頁。

（註二）：中央日報，編寫手冊，第三頁。

第四章 結 論

第一節 可讀性公式的評價

唐朝詩人白居易的詩，號稱「老嫗能解」，似乎可懸為可讀性的最高境界。

可讀性公式最大的貢獻，即在破除「知而不能行」的毛病，影響寫作很大——特別是新聞寫作。賀亨柏（John Hohenberg）說，美聯社和合衆國際社，注意到可讀性的問題，「是抓住了新聞事業的首要原則，這是破天荒之舉」（註一）。

然而，可讀性公式在風行之際，也遭到不少反對者的批評。

新聞界詰難傅萊區等人的理由是：爲什麼在動手寫新聞的時候，要受到短字、短句的限制？爲什麼要迎合低水準的讀者？如果新聞中的五個W和一個H附屬於人情味之下，讀者如何吸收新聞中的主要事實？

芝加哥每日新聞的伐布德（Robert Faberty）也有同樣的論點。他批評可讀性公式是給新聞報導穿上了一件「緊身衣」（註二）。

爲「世界百科全書」（The World Book Encyclopedia）寫「閱讀」一節的辛普森（Elizabeth A. Simpson）在此書中批評得更爲尖刻：

『使用簡單的語言，對於有異於平常的觀念和意見，表達上可能較爲困難，不若複雜字眼正確。用

來左右句子結構的寫作型式，可能缺乏變化和趣味——儘管句子變得較為可讀。作者爲了適應可讀性公式的要求，可能因打算調整作品的型式而失去若干創意。」

對可讀性公式的主要批評，在於它無法測量讀物的思想內容。婦女家庭雜誌跟哲學著作之不同，不在文章中句子的長短和用字的難易，而是在它們的思想深度（註三）。

這些批評，多半起於他們對可讀性公式的本質瞭解不够所致。可讀性公式不是寫作的萬靈丹，克萊爾指出了可讀性公式先天上的限制（註四）：

- 一、可讀性公式只能測量寫作的一部份：風格。
- 二、公式只能測量風格的一部份：難易程度。
- 三、公式甚至不能完全測量難易程度。
- 四、一篇糟透了的文章，可讀性分數可能很低，也可能很高。如果是很高，你不能說它是篇好文章。

所以，傅萊區在「美聯社寫作手冊」（Writing for The AP）中，再三說明不要盲目遵循可讀性公式。他在「可讀寫作的藝術」一書中，對這問題，有更深入的解釋：

『我只能重複我在「說白話的藝術」的序言中所說的話：我擔心若干讀者會希望有個寫出好文章的神奇公式，因而對我這種簡單的標尺感到失望。其他渴望求得語言正確性的人，會沉迷在這些簡單的規則和計算中，無視於白話文的原則。我所希望的是，讀者不要把這個公式看得太認真，除了把它當做粗略的估計外，不要希望太多。』

實地運用過可讀性公式的人，他們對公式的價值，意見仍甚分歧。

麥基 (Blaine K. Mckee) 曾經向美國「幸福」雜誌 (Fortune) 所列的美國最大的六十家公司的公共關係部門，發出問卷，其中三十二家有反應。答覆問卷的公共關係人員當中，有百分之七十五的人，認為可讀性公式有價值。十二家公司使用過可讀性公式，四家正使用中。有價值與否，跟他們是不是使用過可讀性公式，有直接的關係。使用過公式的十二家公司的公共關係人員，均認為有價值 (註五)。

麥基在例舉正反兩方的意見之後，替可讀性公式的功過，下了一個頗為持平的結論：

一、對於需要改進寫作的人，可讀性公式是有用的工具，但不可一意盲從。比方說，短句並非絕對好過長句。

二、較長的句子，讀者需有較高的教育程度才能了解，有統計數字為證。美國四年級學生的造句，平均長度是十一點一字，九年級是十七點三字，大學畢業班是二十一點五字。

三、寫作者一旦了解字、句複雜所造成的影響，可讀性公式的任務即已達成，寫作者自不需再亦步亦趨，受公式的拘束了。

照這種說法，可讀性公式的用途，是在寫作之後，而不是在寫作之前，可見使用公式的技術和態度，實係決定公式功過的關鍵。如何趨利避害，克萊爾認為，第一次稿不應考慮到可讀性的問題，因為此時拘泥於公式，可能「導致機械式的寫作，喪失了若干個人寫作的風格」。

克萊爾主張，第一次稿完成之後，才以公式衡量可讀性。這樣便能夠知道他的文章是否適合他「心目中的閱聽人」 (intended audience)。

第二節 中文可讀性公式的檢討和建議

鮑爾斯 (Richard D. Powers) 和卡爾 (Bryant E. Kears) 曾表示，目前最靈敏的英文可讀性公式，對於讀物的困難度，只能測度出一半 (註六)。可見英文可讀性公式的研究，仍然存有很多困難。中文可讀性公式，自不例外。

最大的困難，在於缺乏基本研究，致無法確定中文那些要素，影響可讀性最大。這樣的研究，必需是大規模的、持久性的。

例如，我們必需確定一個句子有多少的單字，才是長句或短句？那些個字，是真正的常用字？漢字筆劃的多寡，對可讀性有沒有影響？五聲、六書呢？還有，中文的「音節」跟可讀性有什麼關係？換句話說，中文的特質是什麼？

由於上述的問題，未獲解決，本文所擬出的中文可讀性公式，應該是屬於試探性的。它是不是具有實用的價值，仍有待廣泛的應用予以肯定。

至少，句子的長度和用字的難易，直接影響可讀性，已在本文獲得了證實。

下面提出的幾項建議，可以作為研究中文可讀性公式的方向，也是達到較完美公式必須先解決的問題：

一、中文需要一套難易不同的文章，作為衡量可讀性的標準。這一套標準，宜由專家選定具有代表性的文章，依其難易次序排列，並註明那一篇文章是適合於某種教育程度的「典型文章」。有了這樣的

標準，才能够比較文章的可讀性。

二、「迴歸分析」是否爲求得可讀性公式唯一方法？「填字測驗」(Cloze Procedure)既已在英文、日文和韓文的可讀性研究中，有了相當的成績，這種方法，是否能用來測量中文的可讀性？

三、句長與難字百分比，在中文可讀性公式中，應該佔怎樣的比重才正確？

四、特定讀者與一般的讀者之間，對某一內容是否感到興趣，與句長、難字百分比有什麼關係？

五、「人情趣味」(human interest)在中文可讀性公式中的地位如何？據研究，過份強調「人情趣味」可能趕走讀者而非吸引讀者(註七)，中文是否有同樣的現象？

六、應用公式的時候，樣本的大小，以何種程度爲宜？

七、常用字彙表的字數，以何種程度最能測出文章的難易差別？「字」(word)並不是語言的一個正式單位(formal unit)，因此，計算字數頗引起某些人的批評。麥奇(Paul Mckee)指出了計算字數有下列的缺點(註八)：

(一)着重視覺上的生與熟，而非觀念上的意義，也就是着重「視覺字眼」(sight vocabulary)而非「意義字眼」(meaning vocabulary)。

(二)出現頻數，並非是字的難度萬無一失的指標。

(三)不能指出一個字許多其他的意義。

(四)鼓勵剔除不熟習的字眼，是很值得商榷的作法。

基於這四點原因，梅爾(Marie Mehl)主張計算「概念」(concepts)·杜克(Edward Dolch)

主張計算「事實」(facts)來取代計算字的出現頻數。杜克在一九四九年的建議，仍值得我們擬定中文常用字彙表的參考：

- (一)字彙表字量的多寡，應視應用者的需要而有不同。
- (二)不在字彙表上的字，應加以研究。
- (三)應明瞭字彙表的來源。
- (四)特殊讀物的字彙，不應被忽視。

這些問題，同樣可能發生在中文常用字彙表上。中文除極少數「一音節字」外，均非語言的正式單位。在澳大利亞東方大學研究院任教的柳存仁，正在研究常用詞彙。將來如果以詞彙代替字彙，或許能夠加強中文可讀性公式的正確性，又得以避免字彙表的缺陷。

八、發現影響可讀性的新因素，並修正現有公式。

九、計算筆劃、字彙和核對字彙表，手續麻煩，是否能夠用電子計算機代勞？

本文所擬出的公式，如果有相當程度的價值，則它最終的目的，應該是應用在新聞寫作方面。那麼，下一個步驟，就是實際應用公式，以了解中文新聞寫作的現狀。

第四章 註解

註 一•John Hohenberg, *The Professional Journalist*, 台北虹橋書局出版 1969, P.8.

註 二•錢震，新聞論，中央日報發行，第二七四頁。

註 三•Hohenberg, op. cit., P. 82.

- 註 四••George R. Klare, *The Measurement of Readability*, Iowa State University Press 1963, PP.24-25.
- 註 五••Blaine K. McKee, "Readability Formulas can Aid Good Writing," *Public Relations Journal*, July 1967, P.8.
- 註 六••Richard D. Powers & Bryant E. Kears, "Further Directions for Readability Research," *Journalism Quarterly*, Fall 1958, PP.427-32.
- 註 七••Merritt C. Ludwig, "Hard Words and Human Interest: Their Effects on Readership," *Journalism Quarterly*, Summer 1949, P.171.
- 註 八••Klare, op. cit., PP.87-88.

