

國小二、三年級學童朗讀流暢能力之成長模式探究*

張郁雯

國立臺北教育大學
教育學系

朗讀流暢性測驗常用於設定學童的成長目標以及監控其學習進展。先前以課文為測試文本的研究指出朗讀流暢能力的成長符合線性模式。本研究以標準文本探討國小二、三年級學童朗讀流暢能力的成長模式並估計其成長率。研究參與者為新北市兩所小學的二年級學生 57 名，三年級學生 51 名。每三週個別施測一次朗讀流暢測驗，一個學年共施測 13 次。以階層模式分析資料，結果發現分段成長模式比線性成長模式與資料的適配度較好，意即學童在同年度的上、下學期有不同的成長速度。二上、二下、三上以及三下每個學期可以進步的字數分別為 13、18、20 和 6 個字。二年級上學期初，每分鐘朗讀的正確字數平均為 109 個字；三年級上學期初約為 137 個字，下學期期末平均為 165 個字。二年級學童在學期初的朗讀流暢能力與其成長率之間沒有顯著之相關。然而，在三年級上學期，朗讀流暢能力與其成長率相關值為 0.32 達顯著水準。下學期成長率與朗讀流暢初始能力卻呈現顯著的負相關。最後，作者根據研究結果對朗讀流暢性研究與教學實務之意涵提出討論。

關鍵詞：朗讀流暢能力、分段成長模式、成長斜率、標準文本

* 1. 通訊作者：張郁雯，yuwenc@mail.ntue.edu.tw。

2. 本研究承蒙科技部補助（計畫編號：NSC 99-2410-H-152-011-MY3），特此致謝。作者感謝參與本研究之所有國小學童與教師。

閱讀歷程包含認字與理解，而流暢的認字技能則是閱讀理解的必要條件（陳明蕾，2018；蘇宜芬等人，2016；LaBerge & Samuels, 1974; Stanovich, 1980），也是連結識字與理解間的重要橋樑（Hudson et al., 2009; Kuhn et al., 2010）。朗讀流暢能力的測量是請學童朗讀文章，計算其 1 分鐘能唸出的正確字數。美國閱讀諮詢委員會（National Reading Panel, 2000）回顧超過 10 萬筆以上的研究後，認定朗讀流暢度是閱讀教學的五項重要基石之一。實徵研究也得到朗讀流暢能力與閱讀理解表現有顯著的正相關（張郁雯，2018；葉靖雲，1998；Arnesen et al., 2017; Reschly et al., 2009; Shin & McMaster, 2019; Wayman et al., 2007）。

國內蘇宜芬等人（2016）以朗讀流暢測驗檢測學童認字效能，並探討使用該測驗篩選閱讀困難學童的準確度。他們主張當朗讀流暢度的 z 分數為 $-.5$ 時，命中率（即正確篩出閱讀困難個案以及正確排除非閱讀困難學童佔整體受檢學生數之比率）可達 $.70$ 以上，仍能維持一定的敏銳度（即閱讀困難個案能被測驗正確篩出的比率）以及可接受的錯誤接受率（即非閱讀困難學生被錯誤篩選為閱讀困難個案的比率），但錯誤拒絕率（即閱讀困難個案被錯誤排除的比率）偏高。國內過去相關研究（如王梅軒、黃瑞珍，2005；葉靖雲，1993，1998）提出了朗讀流暢測驗的效標關連效度證據，此一研究則提出測驗做為篩檢工具的效度證據。

朗讀流暢性測驗也能用於監控與評估學生閱讀能力之學習進展，以適時調整教學。研究發現一年級時的朗讀流暢成長率對小學一年級及升到三年級時的閱讀理解有最大解釋力，高過聲韻覺識、字母命名、詞彙等技能。對二年級和三年級的學童，則是初始的朗讀流暢能力最能預測其三年級的閱讀理解能力。此一結果意味著朗讀流暢能力是整合低階技能而形成的較為高階的能力（Kim et al., 2010）。當用於監控閱讀之學習進展時，閱讀教學介入後每分鐘的朗讀正確字數，以及介入期間學生的朗讀流暢的成長率是學習進展監控觀察的兩項指標（Burns et al., 2016）。國內多篇朗讀流暢能力的研究提供了國小學童各年級的朗讀流暢平均數值（王梅軒、黃瑞珍，2005；吳明隆等人，2013；張毓仁、吳明隆等人，2011；張毓仁、邱皓政等人，2011；葉靖雲，1993，1998；蘇宜芬等人，2016）。然而，除了蘇宜芬等人（2016）的研究，其餘沿用課程本位評量的作法，以教科書的課文做為測驗的文本（Fuchs & Deno, 1994）。吳明隆等人（2013）除了運用不同版本教科書中的課文做為朗讀流暢性的施測文本，也以非學校採用版本的教科書課文做為測驗的文本，並將後者文本所測得的朗讀流暢性能力，稱為一般成就朗讀流暢性評量（general outcome measurement oral reading fluency），有別於課程本位評量。其研究的對象為國小一年級的學生，每個學期測量 2 次，為期 2 年，共測量 7 次，涵蓋國小一、二年級學童的朗讀流暢性成長階段。結果發現以課程本位評量測得的朗讀流暢性比以一般評量測得的，每分鐘的朗讀正確字數多 50 個字。顯示朗讀流暢能力的估計受到文本難度的影響（Hintze & Christ, 2004）。以課程本位文本施測，學童所朗讀的是課堂中學習過的文本。以標準文本施測，學童所朗讀的則是未學習過的文本，此種方式更能測出學童朗讀新文本的表現。同時，評量是學習的一環，Reed 等人（2019）指出朗讀不同文本比重複朗讀相同文本更能提升四年級學童的朗讀流暢性，對於中低成就者，朗讀不同文本的效果更為顯著。以標準文本測量學童的流暢性，提供學童朗讀不同文本的機會，也有助於其朗讀流暢性的提升。

國內對於朗讀流暢能力成長率的探究僅見於吳明隆等人（2013）的研究。國外雖有不少研究，但中、西方文字系統不同，國外朗讀流暢之成長率數值，恐難以做為國內研究與實務之參考借鏡。

吳明隆等人（2013）的研究指出朗讀流暢性能力符合線性成長模式。成長率方面，課程本位評量測得的成長率為半個學期進步 9.62 字。若使用一般評量，則其成長率為 11.05 字。該研究為國內第一篇探討朗讀流暢性成長率之研究，在研究設計上仍有兩項進步的空間。該研究的第一個問題是未能控制施測文本的難度。吳明隆等人的研究使用教科書課文，未能確保文本難度的一致性。每分鐘朗讀正確字數及其成長率之估計均會受到朗讀文本難度之影響（Hintze et al., 1994; Hintze & Shapiro, 1997）。第二個問題為施測間隔較長，且間隔時間不等。該研究兩年內測量七次，其中一年級下學期於 6 月下旬施測，二年級上則到 11 月上旬才施測，此次間隔時間幾近 5 個月，其他的間隔時間多半為 10 週，在估計成長率時並未將時間間隔長短納入考量。根據 Chall（1996）的閱讀發展階段理論，國小一、二年級為識字期，而二、三年級則為流暢期。為探討朗讀流暢性之關鍵成長期的成長率，本研究以二、三年級學童為研究對象，採用標準文本的朗讀流暢性測驗（張郁雯，2018），事先已知各文本的難度，並縮短學期內之施測間隔，期待研究結果能提供國內研究人員與小學教師對學童朗讀流暢能力發展之理解與實務監控之運用。以下回顧朗讀流暢能力成長模式、成

長的速率、以及初始能力與成長率關係之相關文獻。

(一) 成長模式與成長率

在探討朗讀流暢性的成長率時，學者通常以年級為分析單位。瞭解學年內多次測量的朗讀流暢成長率是否能以線性模式加以描述。將成長模式分為線性 (linear) 與非線性 (nonlinear)，而非線性模式最常使用二次曲線模式。後來，部分學者發現學期內的成長為線性，但不同學期有不同的成長率，主張分段成長模式 (piecewise growth model) 最能描述朗讀流暢能力之成長 (Christ et al., 2010)。

國外以英文文本進行朗讀流暢的成長趨勢研究，早先認為線性模式之適配度優於非線性模式。兩個常被文獻引用的研究指出各年級學生的成長趨勢為線性，雖然不同年級成長斜率不同 (Deno et al., 2001; Fuchs et al., 1993)。然而，近年挪威學者 Arnesen 等人 (2017) 則指出成長模式不必然為線性，依年級不同而不同。二、三年級學童的朗讀流暢呈線性成長，但四年級秋季到冬季的成長慢，冬季到春季的成長快。五年級秋季一開始比四年級春季得分低，到冬季時達高峰，然後下降。因此，四、五年級的朗讀流暢性呈現非線性成長樣態。Arnesen 等人並未以分段成長模式分析資料。同一年級，年度內不同學期的成長率是否相同，文獻上呈現不一致的結果。Christ 等人 (2010) 指出學童同一學年度，成長斜率依季節而有所不同。從秋季到冬季的成長率不同於冬季到春季的成長率。秋季的朗讀流暢能力之成長率高於春季成長率。例如，小學二年級在秋季時，其成長率為每週增加 1.71 個詞，在春季則是每週增加 1.02 個詞。三年級秋季的成長率為每週增加 1.38 個詞，春季則是每週 0.97 個詞。這種秋季優勢出現在所有年級。隨著年級上升，朗讀流暢能力成長率逐漸趨緩，秋季與春季成長率的差異也逐漸靠近。例如，二年級時兩個學期成長率差異為 0.69 個詞，到六年級減為 0.14 個詞的差異。Christ 等人稱這種季節性成長模式為分段成長模式。Zvoch (2016) 的研究支持分段成長模式，同時也發現冬季到春季的成長率較低。Graney 等人 (2009) 的研究雖然也支持分段成長模式，但卻發現春季的成長幅度大於秋季。這可能是因為他們的研究過程中有方案介入，使得春季有較大的成長幅度。

國內吳明隆等人 (2013) 中文文本的研究得出臺灣國小一、二年級朗讀流暢能力的成長趨勢較符合線性，較不符合非線性模式。成長率方面，課程本位評量測得的成長率為一個學期進步約 19 個字。若使用一般評量，則其成長率為 22 個字。這項結果與預期似乎不合，推論可能是因為課程本位測量容易出現天花板效應，成長幅度受限。另一個可能原因為兩種測量的文本都取自教科書，但就朗讀流暢性而言，其文本難度可能是不同的，造成以一般文本所估得的成長率較高。該研究並未進一步檢驗資料與分段成長模式的適配度，可能是受限於該研究每個學期只有兩個測試資料點。研究測驗間隔大約為 10 週，但一下到二上間的施測間隔卻有 4 個月餘。吳明隆等人未於學期初測量學童的朗讀流暢能力，研究顯示若暑假期間，學童沒有參與教育活動，開學初學童之學業成就會下降 (Kim, 2004)。測驗的間隔長短不同，在資料分析時，未能適時調整，可能讓成長率的估計產生偏誤。Jenkins 等人 (2009) 的研究指出以三週為間隔進行朗讀流暢性測量，能得到較接近真實數值的成長性斜率。參考他們的研究建議，本研究採取三週的施測間隔，縮短先前研究之測驗間隔，並蒐集更多的資料點，有助於探討學童朗讀流暢性的成長模式。本研究將檢視線性模式與分段成長模式何者與資料的適配度更好。若為分段成長模式，則上下兩個學期的成長率呈現何種樣態？

(二) 朗讀流暢之初始能力與成長率的關係

Walberg 等人 (1984) 主張發展早期若個體擁有優勢的教育經驗，則個體能更有效的由新的教育經驗獲得成長。幼時就擁有高的智能與動機，在之後學習的成長更為快速，而良好的表現又獲得進一步增強。例如，詞彙量高的兒童會閱讀較多的書，更多的閱讀經驗又能獲得更豐富的詞彙，這種交互促進關係使得「富者越富，貧者越貧」，即所謂的「馬太效應」(Matthew effects)。Stanovich (1986) 回顧閱讀文獻，指出造成閱讀發展的馬太效應之主要因素包含學童的閱讀量以及學童主動選擇與塑造環境。

可以預期的是，學童詞彙量大者其朗讀流暢能力會較好。依照馬太效應可以推論朗讀流暢起始能力佳者，其成長率會較快。然而，吳明隆等人（2013）的研究卻指出學童朗讀流暢的起始能力佳，容易產生天花板效應，進而壓縮後續的成長幅度。不過，在他們的研究中並未報告朗讀流暢起始能力與成長率之相關。若朗讀流暢能力不同的學童其成長速度不同，則在教學介入時，要依初始能力設定適性的學習成長目標。反之，若成長率並不因起始能力有所不同，則介入時每位學童朗讀流暢能力之增加幅度可設為相同。因此，探討學童朗讀流暢起始能力與成長率之關係，是該測驗實務應用的重要議題。

具體而言，本研究聚焦研究問題為：

1. 國小二年級與三年級學童的朗讀流暢能力之成長模式為何？
2. 國小二年級與三年級學童的朗讀流暢能力每個學期的成長幅度為何？
3. 國小二年級與三年級學童朗讀流暢能力之成長幅度是否與其學期初的能力水準有關？

方法

（一）參與對象

本研究的對象來自新北市兩所小學，二年級與三年級的學生各一班。研究者首先取得兩所學校校長的同意，徵求願意參與研究的導師，再經由導師發放家長同意書。家長同意參加的學童二年級為 57 名，三年級為 51 名，共 108 名學生。其中，女生 53 名，男生 55 名。

（二）研究工具

本研究的測量工具為朗讀流暢性測驗（張郁雯，2018），適用於二、三年級。每個年級有 19 篇文本，文長為 350—400 個字。複本信度介於 .80— .89，二、三年級單一題本施測之類化係數為 .84 和 .85；若每次以三個文本施測求平均表現，則二、三年級的類化係數可達 .94 和 .95。二年級與三年級學童朗讀流暢分數與閱讀理解測驗之相關分別為 .60 和 .67。測驗之信、效度佳。

（三）研究程序

自 101 年 10 月初開始以三週為間隔，施測朗讀流暢性測驗，在隔年的 6 月進行最後一次測試。學生總共在 13 個時間點接受個別測試。參考 Jenkins 等人（2009）的研究，在學年初（前測）、上學期末（期中）和下學期末（後測）均施測三個文本，以期獲得較精確的成長率估計值。

文獻指出施測時學生的第一篇文本的唸讀通常會低估學生的流暢性，故建議應先給學生一篇文本練習（Jenkins et al., 2009），此一練習文本所得到的數值，並不納入計分。本研究除了前測、期中以及後測外，其他次測試時，將前一次測試的文本作為該次測試的練習測驗，接著測該次測驗的文本。例如，前測使用三篇文本，下一次測試時，先以前測的第三篇文本作為練習文本，接著給予正式測驗文本；第三個施測時間點，先給予學生唸讀第二個施測時間點的文本（練習文本），接著給予第三次測試的施測文本（正式文本）。前測、期中與後測都測量 3 個文本，其餘 10 次，測量 2 個文本（練習文本與正式測驗文本），故每位學生共有 29 個朗讀流暢性測驗成績。其中 19 個成績是用於成長分析，另外 10 次測驗分數則視為文本間隔三週的再測分數，用於評估文本的再測信度。10 次中除了三次的再測信度低於 .90，分別為 .85、.87、.88，其餘皆在 .90 以上。顯示 10 篇文本的再測信度良好。

八位施測者經過半天之施測訓練。訓練內容除了介紹測驗操作與計分方式並由施測者兩兩進行施測演練，以確保其施測正確性。施測時全程錄音，以便進行之後的計分確認以及施測者信度的檢驗。作者從所有的錄音檔抽檢 15% 的檔案，由另一位施測者進行複評，結果顯示評分者信度達 .99。國外朗讀流暢測驗研究，評分者間信度在 .95 以上（Christ et al., 2010）。本研究的評分者一致性高。

施測時，由施測者說明指導語，請受測學童從文本的第一個字開始朗讀，施測者以碼錶計時一分鐘，在記錄紙上以劃記紀錄各類錯誤（停頓超過三秒字、發音錯誤、省略字、字序顛倒）。施測三個題本時，請受測學童依照相同方式朗讀下一篇文本。施測時全程錄音，施測結束後，將學童朗讀的總字數扣除錯誤字數，是為朗讀的正確總字數。

(四) 資料分析

研究資料分析的步驟如下：

1. 將學生在 19 個題本得分，依照 Santi 等人（2016）所提出的線性轉換方法以及先前研究同時施測 19 個題本的資料加以等化。再分別計算前測、期中以及期末三次測驗結果的平均值，做為該次測量的成績。因此，每位學生總共有 13 個時間點的測試資料。
2. 進行各次朗讀流暢性正確字數描述統計，瞭解流暢性能力變化情形。
3. 以 Raudenbush 與 Bryk（2002）階層線性模式（hierarchical linear model, HLM）分析朗讀流暢性之成長變化。以隨機係數模式為例，第一層次的模式為：

$$ORF_{it} = \beta_{0i} + \beta_{1i}T_{it} + e_{it}$$

ORF_{it} 是第 i 個人在第 t 個時間點之朗讀流暢性分數， $i = 1, \dots, n$ ； $t = 1, \dots, ti$ 。 T_{it} 是第 t 個時間序的編碼， e_{it} 是第 i 個人時間點 t 的殘差。分析時，時間編碼是以週為單位。 β_{0i} 和 β_{1i} 是隨個人變動的截距與斜率。第二層次的模式為：

$$\begin{aligned} \beta_{0i} &= \gamma_{00} + \gamma_{0i} \\ \beta_{1i} &= \gamma_{10} + \gamma_{1i} \end{aligned}$$

γ_{00} 是該年級朗讀流暢能力的平均， γ_{10} 則為該年級朗讀流暢成長率之平均。 γ_{0i} 是個別學生與該年級平均朗讀流暢能力之差異， γ_{1i} 個別學生與該年級平均斜率之差異。考驗 γ_{10} 是否為零，可以知道整體學生的平均朗讀流暢性是否隨時間成長，考驗 γ_{1i} 可以知道該年級個別學生的朗讀流暢能力成長率是否不同。檢視 HLM 各參數的估計標準誤，則可瞭解個人與年級層次的成長率之信度。

4. 以上述的模式為基礎，分別計算零模式（M0）、隨機截距模式（M1）與隨機係數模式（M2）、分段成長模式（M3）以及二次成長曲線模式（M4）與資料的適配度。各模式設定說明如下：

- (1) 零模式未納入時間訊息，將學生 13 次施測的總平均值進行變異數分析。模型如下：

$$\begin{aligned} \text{第一層次：} & ORF_{it} = \beta_{0i} + e_{it} \\ \text{第二層次：} & \beta_{0i} = \gamma_{00} + \gamma_{0i} \end{aligned}$$

- (2) 線性成長模式——隨機截距模式（M1）：加入時間變項，將成長軌跡設為線性，且假設所有學生的成長斜率都相同，而其起始能力不同。以及在下述模式中將截距設為隨機效果而斜率設為固定係數。模型如下：

$$\begin{aligned} \text{第一層次：} & ORF_{it} = \beta_{0i} + \beta_{1i}T_{it} + e_{it} \\ \text{第二層次：} & \beta_{0i} = \gamma_{00} + \gamma_{0i} \\ & \beta_{1i} = \gamma_{10} \end{aligned}$$

- (3) 線性成長模式——隨機係數模式（M2）：與 M1 模式相似，唯一的不同是容許每個學生的成長率可以不同，亦即 $\beta_{1i} = \gamma_{10} + \gamma_{1i}$ 。

- (4) 分段成長模式（M3）：

$$\begin{aligned} \text{第一層次：} & ORF_{it} = \beta_{0i} + \beta_{1i}T1_{it} + \beta_{2i}T2_{it} + e_{it} \\ \text{第二層次：} & \beta_{0i} = \gamma_{00} + \gamma_{0i} \\ & \beta_{1i} = \gamma_{10} + \gamma_{1i} \\ & \beta_{2i} = \gamma_{20} + \gamma_{2i} \end{aligned}$$

學校的學習因寒暑假而中斷，學習因而停滯或倒退。傳統的單一成長軌跡模式可能無法適當捕捉朗讀流暢性的成長樣態。分段成長模式適用於分析這類間斷時間序列（interrupted time-series）的資料。本研究參考 Zvoch（2016）的研究進行分段成長模式的編碼。以兩個變項 T1 和 T2 分別代表兩個學期的朗讀流暢性的測量時間編碼（以週為單位）。前測朗讀流暢分數對應的 $T1_{i1} = 0$ ，第二次測量間隔三週，故對應的 $T1_{i2} = 3$ 。餘類推。第一學期結束時的測試為第六次，離前測間隔 14 週， $T1_{i6} = 14$ 。 $T1_{i7}, T1_{i8}, \dots, T1_{i13}$ 皆設為 14。第二變項 T2 的設定如下：在第六次測量前均設為 0，即 $T2_{i1}, T2_{i2}, \dots, T2_{i6} = 0$ 。第七次和第六次間隔 6 週，故 $T2_{i7} = 6$ ，間隔三週後測第八次，故 $T2_{i8} = 9$ ，其餘類推。 β_{1i} 反映的是上學期的成長斜率， β_{2i} 則反映下學期的成長斜率。

（5）二次曲線成長模式（M4）：加入時間變項的二次項。

$$\text{第一層次：ORF}_{i_t} = \beta_{0i} + \beta_{1i}T1_{i_t} + \beta_{2i}T2_{i_t} + e_{i_t}$$

$$\text{第二層次：}\beta_{0i} = \gamma_{00} + \gamma_{0i}$$

$$\beta_{1i} = \gamma_{10} + \gamma_{1i}$$

$$\beta_{2i} = \gamma_{20}$$

本研究以卡方檢定來檢驗不同模式之離異值的差異，找出最適配的成長模式。

5. 找出最佳模式中的成長斜率，並分析朗讀流暢初始能力與成長斜率之相關。

結果與討論

（一）朗讀流暢性分數隨時間呈現上升趨勢

表 1 呈現二、三年級學童各次測驗分數經過線性等化後之平均值與標準差。表 1 的數據顯示，兩個年級的朗讀流暢性字數從 10 月初的前測持續增長，由於前測、後測以及期中（上學期末）採同一個時間施測三個題本，因此這三個時間的數據，是報告三次測量的平均值。二年級從 10 月的一分鐘 109 個字，進步到期末的 139 個字。三年級從學年初一分鐘 137 個字，進步到期末的 167 個字。二年級學期結束所測得的朗讀流暢性字數與三年級初所測的十分接近。王梅軒與黃瑞珍（2005）二年級下學期施測 108 名學童，得到的朗讀流暢性之平均為 129 個字，該研究採用教科書課文做為施測材料，作者未說明於下學期何時施測。蘇宜芬等人（2016）以「國民小學二、三年級閱讀理解成長測驗」的文章做為朗讀文本，得到二上、二下、三上和三下學童的朗讀流暢性分別為 122、128、135 和 154 字，標準差介於 34 到 37 字。此兩研究在二下所得的朗讀流暢能力數值與本研究第九次施測的數據接近，亦即在二下大約期中的時間點。由於兩篇論文並未寫出確切的施測時間點，故難以知道本研究結果與過去研究的一致性。張郁雯（2018）二下和三下學期結束前施測，學童朗讀流暢性分別為 147 字和 177 字。吳明隆等人（2013）的研究顯示二下期末的朗讀流暢能力平均為 155 個字，高於本研究的數值。由於文獻顯示每分鐘朗讀正確字數及其成長率之估計受到朗讀文本難度之影響（張郁雯，2018），不同研究使用的朗讀文本不同，所估得的朗讀流暢能力缺乏可比較性。研究結果凸顯使用標準文本測量朗讀流暢性的重要性。同時，二三年級學童朗讀流暢能力仍在發展中，同一學期不同的施測時間點所得到的數據也有所不同。

在 13 次的測試當中，偶爾還是會出現字數倒退的情形。蘇宜芬等人（2016）的研究，四上和六下朗讀流暢度的平均值也曾出現字數倒退的情形。本研究參與研究的人數較少，中間施測時只測一個題本，還有間隔時間短，都可能造成測量的穩定性不足。本研究雖施測兩個文本，但第一個題本使用前次測驗文本做為暖身，為避免練習效果，故計分時只使用單一文本。未來可擴大樣本，每次施測均採用多文本。同時，從數據上看來，間隔 6 週的測量似乎比較能穩定看到流暢性的成長，是否 3 週的改變幅度仍不夠大，未來研究可採施測兩個以上文本，探討不同施測間隔對流暢度成長穩定性之影響。

表 1
二、三年級學童各次測驗之平均值與標準差

	二年級 (N=57)		三年級 (N=51)	
	平均數	標準差	平均數	標準差
前測	109	27	137	30
2	112	26	143	38
3	111	31	146	37
4	119	29	150	38
5	114	32	151	38
期中	120	29	158	37
7	124	28	149	39
8	124	28	154	38
9	131	30	155	37
10	129	29	159	39
11	137	31	157	36
12	138	31	156	38
後測	139	32	167	37

(二) 朗讀流暢能力之發展符合分段成長模式

圖 1 呈現兩個年級學生在 13 個時間點的成長情形。本研究年級為受試者間設計，而 13 個時間點的重複測量為受試者內設計，故進行二因子混合設計變異數分析。由於資料違反球形性的假設，在統計顯著性考驗時採取 Greenhouse-Geisser 的修正方法。結果發現年級效果顯著 ($F(1, 106) = 23.7, p < .001$)，參照圖 1，可知三年級學童的朗讀流暢性分數顯著高於二年級。時間 (13 個測試平均值差異) 主要效果顯著 ($F(7.77, 824) = 44.63, p < .001$)，時間與年級的交互作用顯著 ($F(7.77, 824) = 5.28, p < .05$)。表示二、三年級學童 13 次朗讀流暢性得分的平均值線型是不同的。仔細檢視圖 1，二年級在期中之前的進步較為平緩 (105—114)，在下半學期進步幅度較大 (124—139)；三年級則是期中之前進步幅度較大 (137—151)，期中之後，進步趨於平緩 (158—167)。然而，圖 1 顯示的是二個年級各次測驗整體學生平均值的走勢。接下來將以階層線性模式檢視個人的成長模式。由於從平均值來看，時間與年級有交互作用，故將分別分析兩個年級之成長模式。

為了回答研究問題一，研究者檢視五種模式與資料的符合度。第一種為零模式 (M0) 未納入時間訊息，將學生 13 次施測的總平均值進行變異數分析；第二種模式是線性成長模式——隨機截距模式 (M1)，截距為隨機效果但斜率設為固定係數；第三種為線性成長模式——隨機係數模式 (M2)，截距與斜率均為隨機變數；第四種模式為分段成長模式 (M3)，探討上學期與下學期朗讀流暢性的成長軌跡是否不同；第五種模式為二次曲線成長模式 (M4)，與過去文獻做比較，瞭解成長率是線性還是二次曲線，在第一層次增加一個時間的二次項。根據前三種模式分析結果，隨機斜率的模式適配度較好，故分段成長模式亦將斜率設為隨機。表 2 呈現二年級與三年級資料五個模式的分析結果。

兩個年級各模式的組平均數信度平均皆達 .90 以上，M1 到 M3 的模式組平均數信度平均更達 .95 以上，顯示由 13 次的測量估計個別學生的初始朗讀流暢性數值具高信度。二年級零模式的 ICC (intraclass correlation) 為 .73。顯示 73% 的朗讀流暢分數差異來自學生間差異 (第二層次)，27% 的變異來自個體內的差異 (第一層次)。三年級的 ICC 為 .81。高的 ICC 值代表兩組資料皆為組間異質性高，使用多層次模式有其必要性 (Bliese, 2000)。

圖 1
二、三年級學童學年間朗讀流暢性成長軌跡

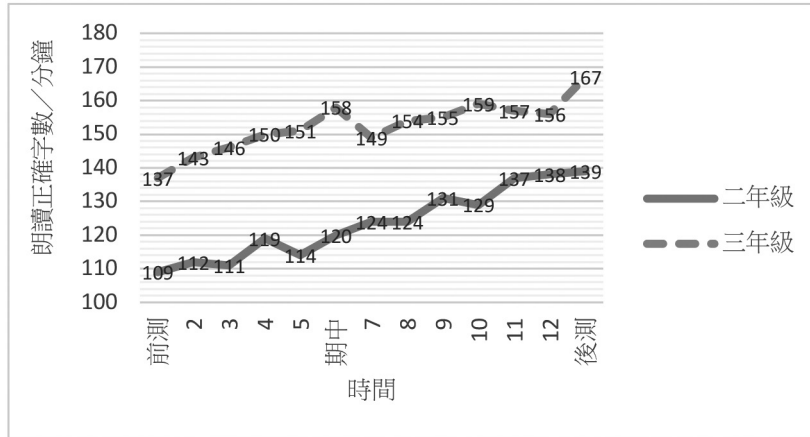


表 2
五種朗讀流暢能力成長模式的分析結果

模式	γ_{00}	γ_{10}/γ_{20}	σ_{ϵ}^2	τ_{00}	τ_{11}	τ_{22}	離異值
二年級							
M0	124***	- / -	264.68	723.4***	-	-	6436
M1	108***	0.84***/-	156.06	731.75***	-	-	6077
M2	108***	0.84***/-	142.89	664.68***	0.086***	-	6055
M3	109***	0.67***/ 0.92***	135.68	638.37***	0.41**	0.23***	6049
M4	109***	0.58**/ 0.007	142.36	664.85***	0.087***	-	6063
三年級							
M0	152***	- / -	277.36	1146.16***	-	-	5809
M1	142***	0.53***/-	233.97	1149.50***	-	-	5706
M2	142***	0.53***/-	186	1127.47***	0.32***	-	5636
M3	139***	1.01***/ 0.30***	180.33	984.93***	0.53*	0.32***	5621
M4	141***	0.85***/ -0.008	185.22	1127.69***	0.32***	-	5643

* $p < .05$. ** $p < .01$. *** $p < .001$.

檢視五個模式的適配度，兩個年級離異值最低的為分段成長模式，表示該模式與資料的適配度最好。不過，由於分段成長模式較為複雜，估計的參數較多，在評估模式適配度時，尚須考慮不同模式離異值的差異是否達統計顯著水準。二年級的分析結果顯示，隨機截距模式 M1，容許每位學生的朗讀流暢性起點不同，但保持固定斜率，能大幅提升資料與模式的適配度。離異值由 M0 的 6436 下降到 6077，兩者差異達顯著水準 $\chi^2(1, N = 57) = 359, p < .001$ 。 τ_{00} 為 731.75 顯示學生間的朗讀流暢能力有顯著的差異。隨機係數模式 M2 容許學生有不同的成長斜率，結果顯示適配度更好，和隨機截距模式 M1 的離異值差異達顯著水準， $\chi^2(2, N = 57) = 22, p < .001$ 。即容許學生有不同的

朗讀流暢起始能力與不同的成長率之模式更能符應資料。

接著研究者容許上下學期的成長率有所不同，且容許成長率因個體而異。分段成長模式 M3 的離異值下降到 6049，但並未達統計顯著水準 ($\chi^2(3, N = 57) = 6, p > .05$)。也就是說分段成長模式並未比隨機係數模式有顯著更佳的適配度。但若與隨機截距模式 M1 相比，則兩者離異值的差異達統計顯著水準 ($\chi^2(5, N = 57) = 28, p < .001$)。

最後，檢測文獻上經常用來與線性模式比較之二次曲線模式 M4。該模式的離異值高於隨機係數模式 M2 與分段成長模式 M3。因此，二次曲線成長模式並未優於線性與分段成長模式。

三年級朗讀流暢性資料與模式的適配度與二年級雷同。隨機截距模式容許每位學生的朗讀流暢性起點不同，但保持固定斜率，能提升資料的適配度，離異值由零模式之 5809 下降到 5706，兩者差異達顯著水準 $\chi^2(1, N = 57) = 103, p < .001$ 。若讓每位學生可以有不同斜率的隨機係數模式，適配度更好，與隨機截距模式的差異達顯著水準 $\chi^2(2, N = 57) = 70, p < .001$ 。分段成長模式之離異值最小，且與隨機係數模式相較，適配度顯著改善 ($\chi^2(3, N = 57) = 15, p < .01$)。二次曲線模式的離異值比隨機係數模式以及分段成長模式高。故與三年級的資料適配度最好的是分段成長模式。經由各個模式的比較，並考慮兩個年級資料解釋的一致性，本研究以下報告分段成長模式的結果。

二年級學童的朗讀流暢能力在學期初為每分鐘 109 個字，由於 τ_{00} 統計檢驗達顯著水準，表示學童間的朗讀流暢能力有顯著的差異。上學期每週提升 0.67 個字，一學期以 20 週計，則平均可提升大約 13.4 個字；亦即上學期結束時，朗讀流暢能力之字數應為 122 個字。下學期每週提升 0.92 個字，一學期提升 18.4 個字。在二年級結束時大約為 140 個字。一年可以進步 32 個字。吳明隆等人 (2013) 的研究，以課程本位朗讀流暢性所估得的成長率為一學期 19.2 個字，一年約為 38 個字。該研究並未檢驗分段成長模式與資料的符合度。然，檢視其研究資料的常模，可見一年級期末（接近本研究二年級期初）平均為 160 個字，二上期末 190 個字，二下期末為 198 個字，呈現的是上學期進步較快，下學期進步較少（30 個字 vs 8 個字）。以一般成就朗讀流暢性測驗測量，其成長率為一學期 22.1 字（一年約為 44 個字）。其研究一年級期末為 113 個字，二上期末 142 個字，二下期末為 155 個字，也是上學期進步較快，下學期進步較少。上學期為 29 個字，下學期為 17 個字。在他們的研究，兩種測量方法上學期成長約 30 個字；下學期差異較大，顯然分段成長模式比線性成長模式更符合資料。本研究的結果是二上的進步幅度略低於二下，而根據吳明隆等人研究的常模資料則顯示上學期大幅進步，下學期的進步幅度有限。除了成長趨勢上的差異，在成長字數上也有明顯的不同。本研究的文本屬性上接近吳明隆等人的一般成就朗讀流暢文本。本研究二年級一個學年之成長字數約為 32 字，而吳明隆等人的研究高達 44 個字，差距頗大。

上述研究的差異可能來自兩項研究設計上的差異。第一，朗讀流暢性文本的難度控制。吳明隆等人 (2013) 的每次的朗讀文本是取自不同版本的教科書課文 3 篇，並未考慮文本的難度問題。張郁雯 (2018) 的研究指出學童在二下期末測得的朗讀流暢性，因文本不同，平均值會在 130 到 169 字之間，文本難度的影響頗大。因此，吳明隆等人研究對成長率的估計可能受到文本難度的混淆。第二個研究設計的問題來自測試時間間隔。吳明隆等人研究兩年內檢測 7 個時間點，從一下到二上間隔有 4 個多月，比其他測試點間隔來得長（吳明隆等人，2013，頁 46）。二上的時間橫跨 7 個月，而二下只有 5 個月。該研究在資料分析時，將時間間隔以等距視之。本研究則在一個學年度內有 13 個施測時間點，施測間隔時間控制為 3 週，寒假的間隔為 6 週。除了以較多的時間點估計成長率，在資料分析時，是以實際間隔週數做為時間值，較能真實反映朗讀流暢能力在上、下學期的發展情形。

三年級學生在上學期的朗讀流暢性成長速率為每週 1.01 個字，亦即上學期一分鐘的朗讀流暢性大約成長 20 個字，和二年級下學期接近。下學期成長率為 .30，即成長減緩為 6 個字。一整年的成長字數為 26 個字，略低於二年級的 32 個字。本研究在三年級的成長幅度與蘇宜芬等人 (2016) 的研究相似都是 26 個字，但他們的研究發現三上進步 7 個字，三下成長幅度為 19 個字，呈現上學期進步較少，下學期進步較多。張毓仁、吳明隆等人 (2011) 對四、五、六年級共 232 名學童施測朗讀流暢測驗，結果發現四年級與五年級的朗讀流暢度僅成長 7 個字。過去研究在沒有文本難度等化的情形下，發現各學期的成長幅度變異大。本研究經文本難度等化，三年級成長幅度依然有著學期效應。若要運用朗讀流暢進行學習進展監控，需要有更多研究確認二、三年級學童的成長樣態，才能針對學童設立適當的教學目標。

綜合二年級與三年級的結果，似乎指向二年級到三年級上學期，是學童朗讀流暢性的主要成長期，二上、二下和三上一個學期可以進步的字數分別為 13、18 和 20 個字。二、三年級學童的識字量仍在成長中，根據王瓊珠等人（2008）指出國小一到四年級的識字量平均為 700、1200、2100、2600 個字。二到三年級是識字擴增的時期，也是識字自動化的重要時期。另外，三年級的識字量可視為脫盲的水準。吳明隆等人（2013）研究指出中文認字與朗讀流暢性之相關在 .6 左右。剛上二年級的學童，識字量不多且識字技能不夠精熟，面對陌生文本，需要耗費較多的認知資源在認字上，朗讀流暢能力成長較為緩慢。朗讀流暢性做為連結識字與理解間的重要橋樑（Kuhn et al., 2010），隨著識字量增加，識字技能越加精熟，朗讀流暢性到二下開始快速成長持續到三年級上學期。這樣的發展符合 LaBerge 與 Samuels（1974）對技能自動化發展的觀點。他們認為初學時技能尚未精熟需耗費大量認知資源，但正確率與速度會持續成長，等到技能精熟，則正確率與速度已在高水準，進步幅度變小，此時認知資源的耗費會降低。根據 Chall（1996）的閱讀理解發展理論，國小二、三年級學生為朗讀流暢期，三年級下學期的發展趨緩，雖符合理論的預期，但下降幅度頗大。此一現象值得未來研究持續加以探究。

Christ 等人（2010）發現朗讀流暢性的成長率在各年級都是秋季比春季的成長幅度大，秋季與春季間的差異隨著年級而變小。他們認為教師對低成就學生期待逐漸變低，教室管理在下學期鬆懈皆是可能的因素。然而，本研究二三年級的學期效應趨勢並不一致，學校的教學可能造成成長率的加速或減緩（Graney et al., 2009）。國內現有二、三年級學童朗讀流暢的相關研究僅有上、下兩個學期期末的測量，且未能將文本難度納入考量，期望未來更多研究能驗證本研究的學期效應結果。同時也可探討造成不同學期學習成長幅度差異之原因以及積極的介入教學對朗讀流暢能力成長的影響。

（三）朗讀流暢之初始能力與成長率的關係

表 3 呈現兩個年級朗讀流暢初始能力與兩個學期成長率之相關。在二年級時，學童在學期初的朗讀流暢能力與其成長率之間沒有顯著之相關。在這個階段，雖然學童的朗讀流暢能力有明顯的個別差異，但能力好壞與其朗讀流暢能力的發展快慢未出現顯著的線性關係。兩個分段的斜率之變異數達統計顯著水準，這表示每位學童的成長率存在著個別差異。然而，在三年級上學期，朗讀流暢能力與其成長率達顯著正相關 $r = 0.32, p < .01$ 。下學期成長率與朗讀流暢初始能力與呈顯著的負相關 $r = -.43, p < .01$ 。亦即三年級上學期，朗讀流暢初始能力較佳者在上學期快速成長，到了下學期其成長速率比朗讀流暢能力弱者來得慢。這個結果部分與吳明隆等人（2013）的研究一致，他們指出中文認字能力會負向影響朗讀流暢性的成長速率。起始能力佳的學生，屬於閱讀發展較成熟者，其識字量較多、識字技巧較佳，到下學期時若朗讀文本難度未提高，在計時測驗下，可能會產生天花板效應，使得成長速率受限。另外，迴歸到平均值現象也可能導致此一結果。Campbell 與 Kenny（1999）主張長時間測量時，若變異數保持相同，迴歸到平均值會造成起始值與成長率呈現負相關。從表 1 可以看到三年級只有前測標準差相對較小，這之後各次測試的標準差頗為接近。起始能力佳者在上學期成長較快，到下學期則是迴歸到平均值，成長幅度趨緩。而起始能力弱者可能識字量與識字技能在上學期還在緩慢成長，到下學期識字自動化發展較成熟時，朗讀流暢成長率較快。不同閱讀能力學童的朗讀流暢成長幅度是否不同值得未來研究進一步探究。

在不同年級，不同學期，學童的朗讀流暢初始能力與成長率呈現不同關係型態。此一發現提醒教師設定教學目標時，應將學生的年級以及初始能力納入考量，同時依照學生的成長，適時調整目標。未來推動閱讀策略教學時，除了深化教師對閱讀理論與閱讀信念之探討（柯華葳，2020），也可將學生閱讀起始能力的檢測，以及根據檢測結果設定教學目標納入教師研習之項目。

表 3
朗讀流暢初始能力與成長率之相關

		分段成長模式		隨機係數模式
		β_{10}	β_{20}	β_{10}
二年級	β_{00}	.11	.08	.13
三年級	β_{00}	.32 **	-.43 **	-.12

* $p < .05$. ** $p < .01$.

結論與建議

本研究以標準文本探討國小二、三年級學童朗讀流暢能力的成長模式並估計其成長率。結果發現分段成長模式與三年級資料的適配度最好，意即學童不同學期的成長速度有所不同，對二年級而言，分段成長模式與隨機成長模式與資料適配度相似，亦即就現有的資料，尚無從支持上下學期的成長率有明顯的不同。二上、二下、三上以及三下每個學期可以進步的字數分別為 13、18、20 和 6 個字，學童每週朗讀流暢能力的成長字數為 0.67, 0.92, 1.01, 0.30 字。結果顯示朗讀流暢性在二年級到三上的成長幅度較大，到三下明顯趨緩。朗讀流暢性成長率的變化可能對應識字量的擴增速度以及識字技能的熟練度的增長。研究結果提出二、三年級不同學期朗讀流暢能力的成長率，除了有助於閱讀研究者參考對照之用，對閱讀教學也有其應用價值。教學現場教師針對教學介入多半透過觀察來評估成效。若是蒐集相關數據，多半使用描述的方式來呈現學生學習表現。本研究使用分段成長模式分析縱貫資料的作法，能提供教學介入更堅實的支持證據。此外，不同學生如何以不同方式回應教學是研究者與教學實務者感興趣的議題，分段成長模式能系統性描述與評估教學實務動態的本質。研究結果顯示以朗讀流暢性做為閱讀理解能力的指標，進行閱讀教學時，針對不同年級，不同學期，宜訂定不同朗讀流暢的成長目標，而非整個學年度的每個月份均設定相同的成長目標。

二年級上學期初，每分鐘朗讀的正確字數大約在 110 個字；三年級上學期初約為 140 個字，下學期期末約為 165 個字。本研究結果可做為二、三年級學童應有朗讀流暢能力水準以及成長速度的期望值。教師可藉此瞭解學生的朗讀流暢能力的成長是否符合預期，以評估二、三年級國語文教學成效以及是否需要調整教學目標。其次，文獻顯示，朗讀流暢能力對閱讀理解有不錯的預測力。根據此一研究結果，教師可運用朗讀流暢測驗結果，快速地找出疑似閱讀落後高風險學生，給予適當的介入，並以其成長率評估教學介入效果。朗讀流暢測驗結果也能提供教師初步評估班級學生相對的閱讀能力。再者，以線性成長模式估計學童朗讀流暢能力之成長對三年級學童可能是不恰當的，應納入學期效應因素。以學童朗讀流暢性的成長量評估教學介入的成效時，若如過往使用線性成長模式，則在成長較慢的學期會高估教學成效未達標的人數；相反地，在成長較快速的學期，則會低估未達標的學生人數。

本研究的結果顯示不同年級有的在上學期成長率高，有的則是在下學期成長率高，學期效應是否可能是教師教學因素導致的？造成此種學期效應的成因值得未來研究加以探討。本研究 13 次測試過程中，並未將學童測驗結果立即回饋給教師。若教師能即時獲得學童朗讀流暢測驗的結果，並能根據評量結果，調整教學的策略，對學童的學習當能產生正向的影響，可避免學習落後的產生。回饋朗讀流暢測驗之結果並引導教師善用評量結果改善教學，能對學童朗讀流暢成長產生何種影響是未來值得深究的主題。

本研究結果僅部分支持馬太效應。學期初時，有較好的朗讀流暢能力之二年級學童與其成長速度並未顯著優於能力較差者。但三年級上學期朗讀流暢能力與成長速率兩者呈現顯著正相關。朗讀流暢能力佳者在上學期成長較快速，但能力弱者成長速度較慢。若能對二年級閱讀能力弱的學童及早介入，提升其朗讀流暢能力，也許有助於其在三年級上學期快速的成長，避免個別差異的擴大。

本研究特色之一是使用標準文本進行朗讀流暢性測驗。測驗的文本事先經過難度測試，再根據實徵難度，進行測驗分數的線性轉換。使用標準文本進行朗讀流暢性測驗有諸多的優點。首先，各個研究的結果有對話的基礎。不同研究者若使用的不同的朗讀流暢性文本，所得到的結果會陷入難以比較的困境。二、三年級學童的朗讀流暢能力應為多少才是合理的數值？即便是以教科書為文本

的研究，也因為不同版本或課文的變動所得到的結果也有比較上的困難。其次，朗讀流暢測驗作為進展監控的文本必須具備成長敏感度。現有研究顯示一個學期成長的字數最多的為 22 個字，每週成長的字數大約為 2 個字。若未能控制文本難度，成長幅度受到文本難度差異的混淆，恐怕無法真正測得成長的情形。由於每週的成長字數不高，未來可透過研究找出最佳的施測間隔，使施測結果效能最大化。雖然，國外的研究指出 3 到 9 週的間隔皆可以得到有效的成長率估計。但研究設計應納入多個時間點的資料蒐集，方能真正反應朗讀流暢複雜的成長軌跡。

由於朗讀流暢性測驗需要個別施測，限於人力與經費，本研究僅以雙北學生為對象，未來研究可將對象擴及其他縣市各年級的學童，除了驗證本研究估計朗讀流暢能力之成長率外，也能得知各年級成長率的變化情形，做為研究與實務的參考。此外，本研究採年級內縱貫研究設計，二年級和三年級學童是兩個獨立的樣本。雖然是取自相同學校不同年級的班級，仍不能排除樣本間差異對結果的影響。三年級下學期的成長幅度小，未來可對三年級下學期以及四年級的學童，進行朗讀流暢能力之成長研究，確認在此一時期的成長是否已到高原期。

本研究使用的標準文本，每個年級共有 19 篇文本。為了正確估計成長斜率，在期初、期中與期末以三個題本施測，雖然中間施測點施測兩個題本，但計分時只能採計一篇文本的朗讀流暢分數，為了提升測量的精確性，擴充標準文本的數量有其必要性。

參考文獻

- 王梅軒、黃瑞珍（2005）：〈國小課程本位閱讀測量方法之信度與效度研究〉。《特殊教育研究學刊》，29，73–94。[Wang, M.-H., & Huang, R.-J. (2005). The reliability and validity of curriculum-based reading measures on elementary school students. *Bulletin of Special Education*, 29, 73–94.] <https://doi.org/10.6172/BSE200509.2901004>
- 王瓊珠、洪儷瑜、張郁雯、陳秀芬（2008）：〈一到九年級學生國字識字量發展〉。《教育心理學報》，39（4），555–568。[Wang, C.-C., Hung, L.-Y., Chang, Y.-W., & Chen, H.-F. (2008). Number of characters school students know from grade 1 to g9. *Bulletin of Educational Psychology*, 39(4), 555–568.] <https://dx.doi.org/10.6251/BEP.20071026>
- 吳明隆、張毓仁、曾世杰、柯華葳、林素貞（2013）：〈國小低年級學童中文朗讀流暢能力的發展軌跡分析〉。《臺東大學教育學報》，24（2），33–65。[Wu, M.-L., Chang, Y.-J., Tzeng, S.-J., Ko, H.-W., & Lin, S.-J. (2013). Developmental trajectories in Chinese oral reading fluency: A longitudinal survey of 6- and 7-year-old children. *NTTU Educational Research Journal*, 24(2), 33–65.] <https://doi.org/10.3966/102711202013122402002>
- 柯華葳（2020）：〈臺灣閱讀策略教學政策與執行〉。《教育科學研究期刊》，65（1），93–114。[Ko, H.-W. (2020). Reading policy and reading instruction in Taiwan. *Journal of Research in Education Sciences*, 65(1), 93–114.] [https://doi.org/10.6209/JORIES.202003_65\(1\).0004](https://doi.org/10.6209/JORIES.202003_65(1).0004)
- 陳明蕾（2018）：〈課文本位閱讀策略教學對國小學童閱讀表現與策略使用覺知情形之影響〉。《教育心理學報》，49（4），581–609。[Chen, M. (2018). The effects of textbook-based approach to reading instruction on elementary graders' reading comprehension and awareness of using reading strategy. *Bulletin of Educational Psychology*, 49(4), 581–609.] [https://doi.org/10.6251/BEP.201806_49\(4\).0004](https://doi.org/10.6251/BEP.201806_49(4).0004)
- 張郁雯（2018）：〈朗讀流暢性測驗標準文本之發展研究〉。《教育實踐與研究》，31（2），1–23。[Chang, Y. (2018). The development of oral reading fluency passages. *Journal of Educational*

- Practice and Research*, 31(2), 1–23.]
- 張毓仁、吳明隆、胡芝妮（2011）：〈國小四、五和六年級學童國語文課程本位朗讀流暢能力之比較〉。《教育研究月刊》，210，49–61。[Chang, Y.-J., Wu, M.-L., & Hu, C.-N. (2011). Curriculum-based oral reading fluency for students in grades 4 through 6. *Journal of Education Research*, 210, 49–61.]
- 張毓仁、邱皓政、柯華葳、曾世杰、林素貞（2011）：〈聲韻覺識、唸名速度和流暢性對中文閱讀理解的影響：結構方程模式與增益效度之探究〉。《教育與心理研究》，34（1），1–28。[Chang, Y.-J., Chiou, H.-J., Ko, H.-W., Tzeng, S.-J., & Lin, S.-J. (2011). The effects of phonological awareness, rapid naming speed, and oral reading fluency on Chinese reading comprehension: A study of structural equation model and incremental validity. *Journal of Education & Psychology*, 34(1), 1–28.]
- 葉靖雲（1993）：〈課程本位閱讀測驗的效度研究〉。《特殊教育學報》，8，273–323。[Yeh, C.-Y. (1993). The validity of curriculum -based reading comprehension tests. *Journal of Special Education*, 8, 273–323.]
- 葉靖雲（1998）：〈課程本位閱讀測驗的效度研究〉。《特殊教育與復健學報》，6，239–260。[Yeh, C.-Y. (1998). The validity of curriculum-based reading comprehension tests. *Bulletin of Special Education and Rehabilitation*, 6, 239–260.]
- 蘇宜芬、張祐瑄、李孟峰、黃鈺茜（2016）：〈國小二至六年級朗讀流暢度篩檢準確度及切截點分析〉。《教育科學研究期刊》，61（4），33–57。[Su, Y.-F., Chang, Y.-H., Li, M.-F., & Huang, Y.-C. (2016). Assessment accuracy and cut-off points of oral reading fluency for grade 2-6 students. *Journal of Research in Education Sciences*, 61(4), 33–57.] [https://doi.org/10.6209/JORIES.2016.61\(4\).02](https://doi.org/10.6209/JORIES.2016.61(4).02)
- Arnesen, A., Braeken, J., Baker, S., Meek-Hansen, W., Ogden, T., & Melby-Lervåg, M. (2017). Growth in oral reading fluency in a semitransparent orthography: Concurrent and predictive relations with reading proficiency in Norwegian, Grades 2–5. *Reading Research Quarterly*, 52(2), 177–201. <https://doi.org/10.1002/rrq.159>
- Bliese, P. D. (2000). Within-group agreement, non-independence, and reliability: Implications for data aggregation and analysis. In K. J. Klein & S. W. J. Kozlowski (Eds.), *Multilevel theory, research, and methods in organizations: Foundations, extensions, and new directions* (pp. 249–381). Jossey-Bass. <https://doi.org/10.2307/3094811>
- Burns, M. K., Silbergitt, B., Christ, T. J., Gibbons, K. A., & Coolong-Chaffin, M. (2016). Using oral reading fluency to evaluate response to intervention and to identify students not making sufficient progress. In K. D. Cummings & Y. Petscher (Eds.), *The fluency construct: Curriculumbased measurement concepts and applications* (pp. 123–140). Springer Publishing Company. https://doi.org/10.1007/978-1-4939-2803-3_5
- Chall, J. S. (1996). *Stages of reading development*. McGraw-Hill. <https://doi.org/10.1017/S0142716400005166>
- Christ, T. J., Silbergitt, B., Yeo, S., & Cormier, D. (2010). Curriculum-based measurement of oral reading: An evaluation of growth rates and seasonal effects among students served in general and special education. *School Psychology Review*, 39(3), 447–462. <https://doi.org/10.1080/02796015.2010.1208>

7765

- Campbell, D. T., & Kenny, D. A. (1999). *A primer on regression artifacts*. Guilford.
- Deno, S. L., Fuchs, L. S., Marston, D., & Shin, J. (2001). Using curriculum-based measurement to establish growth standards for students with learning disabilities. *School Psychology Review*, 30(4), 507–524. <https://doi.org/10.1080/02796015.2001.12086131>
- Fuchs, L. S., & Deno, S. L. (1994). Must instructionally useful performance assessment be based in the curriculum? *Exceptional Children*, 61(1), 15–24. <https://doi.org/10.1177/001440299406100103>
- Fuchs, L. S., Fuchs, D., Hamlett, C. L., Walz, L., & Germann, G. (1993). Formative evaluation of academic progress: How much growth can we expect? *School Psychology Review*, 22(1), 27–48.
- Graney, S. B., Missall, K. N., Martínez, R. S., & Bergstrom, M. (2009). A preliminary investigation of within-year growth patterns in reading and mathematics curriculum-based measures. *Journal of School Psychology*, 47(2), 121–142. <https://doi.org/10.1016/j.jsp.2008.12.001>
- Hintze, J. M., & Christ, T. J. (2004). An examination of variability as a function of passage variance in CBM progress monitoring. *School Psychology Review*, 33(2), 204–217. <https://doi.org/10.1080/02796015.2004.12086243>
- Hintze, J. M., & Shapiro, E. S. (1997). Curriculum-based measurement and literature-based reading: Is curriculum-based measurement meeting the needs of changing reading curricula? *Journal of School Psychology*, 35(4), 351–375. [https://doi.org/10.1016/S0022-4405\(97\)00014-9](https://doi.org/10.1016/S0022-4405(97)00014-9)
- Hintze, J. M., Shapiro, E. S., & Lutz, J. G. (1994). The effects of curriculum on the sensitivity of curriculum-based measurement in reading. *Journal of Special Education*, 28(2), 188–202. <https://doi.org/10.1177/002246699402800205>
- Hudson, R. F., Pullen, P. C., Lane, H. B., & Torgesen, J. K. (2009). The complex nature of reading fluency: A multidimensional view. *Reading and Writing Quarterly*, 25(1), 4–32. <https://doi.org/10.1080/10573560802491208>
- Jenkins, J. R., Graff, J. J., & Miglioretti, D. L. (2009). Estimating reading growth using intermittent CBM progress monitoring. *Exceptional Children*, 75(2), 151–163. <https://doi.org/10.1177/001440290907500202>
- Kim, J. (2004). Summer reading and the ethnic achievement gap. *Journal of Education for Students Placed at Risk*, 9(2), 169–188. https://doi.org/10.1207/s15327671espr0902_5
- Kim, Y. S., Petscher, Y., Foorman, B. R., & Zhou, C. (2010). The contributions of phonological awareness and letter-name knowledge to letter-sound acquisition—a cross-classified multilevel model approach. *Journal of Educational Psychology*, 102(2), 313–326. <https://doi.org/10.1037/a0018449>
- Kuhn, M. R., Schwanenflugel, P. J., & Meisinger, E. B. (2010). Aligning theory and assessment of reading fluency: Automaticity, prosody, and definitions of fluency. *Reading Research Quarterly*, 45(2), 232–251. <https://doi.org/10.1598/RRQ.45.2.4>
- LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information process in reading. *Cognitive Psychology*, 6(2), 293–323. [https://doi.org/10.1016/0010-0285\(74\)90015-2](https://doi.org/10.1016/0010-0285(74)90015-2)
- National Reading Panel. (2000). *Report of the National Reading Panel: Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications*

- for reading instruction: Reports of the subgroups*. National Institute for Literacy. <https://doi.org/10.1598/RRQ.36.3.5>
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). SAGE Publications. <https://doi.org/10.1086/230447>
- Reed, D. K., Zimmermann, L. M., Reeger, A. J., & Aloe, A. M. (2019). The effects of varied practice on the oral reading fluency of fourth-grade students. *Journal of School Psychology, 77*, 24–35. <https://doi.org/10.1016/j.jsp.2019.10.003>
- Reschly, A. L., Busch, T. W., Betts, J., Deno, S. L., & Long, J. D. (2009). Curriculum-based measurement oral reading as an indicator of reading achievement: A meta-analysis of the correlational evidence. *Journal of School Psychology, 47*(6), 427–469. <https://doi.org/10.1016/j.jsp.2009.07.001>
- Santi, K. L., Barr, C., Khalaf, S., & Francis, D. J. (2016). Different approaches to equating oral reading fluency passages. In K. D. Cummings & Y. Petscher (Eds.), *The fluency construct curriculum-based measurement concepts and applications* (pp. 223–265). Springer Publishing Company. https://doi.org/10.1007/978-1-4939-2803-3_9
- Shin, J., & McMaster, K. (2019). Relations between CBM (oral reading and maze) and reading comprehension on state achievement tests: A meta-analysis. *Journal of School Psychology, 73*, 131–149. <https://doi.org/10.1016/j.jsp.2019.03.005>
- Stanovich, K. E. (1980). Toward an interactive-compensatory model of individual differences in the development of reading fluency. *Reading Research Quarterly, 16*(1), 32–71. <https://doi.org/10.2307/747348>
- Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21*, 360–407. <https://doi.org/10.1177/0022057409189001-204>
- Walberg, H. J., Strykowski, B. F., Rovai, E., & Hung, S. S. (1984). Exceptional performance. *Review of Educational Research, 54*(1), 87–112. <https://doi.org/10.3102/00346543054001087>
- Wayman, M. M., Wallace, T., Wiley, H. I., Tichá, R., & Espin, C. A. (2007). Literature synthesis on curriculum-based measurement in reading. *Journal of Special Education, 41*(2), 85–120. <https://doi.org/10.1177/00224669070410020401>
- Zvoch, K. (2016). The use of piecewise growth models to estimate learning trajectories and RTI instructional effects in a comparative interrupted time-series design. *The Elementary School Journal, 116*(4), 699–720. <https://doi.org/10.1086/686304>

收稿日期：2020年07月20日

一稿修訂日期：2020年07月21日

二稿修訂日期：2020年10月25日

三稿修訂日期：2021年01月07日

四稿修訂日期：2021年02月25日

五稿修訂日期：2021年03月10日

六稿修訂日期：2021年03月29日

接受刊登日期：2021年03月29日

Bulletin of Educational Psychology, 2021, 53(2), 463–480
National Taiwan Normal University, Taipei, Taiwan, R. O. C.

Growth Model of Oral Reading Fluency Among Second and Third Graders

Yuwen Chang

Department of Education,
National Taipei University of Education

Fluent reading is an indication of proficient decoding and comprehension. Studies have demonstrated that oral reading fluency (ORF) correlates positively with reading comprehension. The ORF assessment comprises a set of individually administered 1-minute reading probes. These tests are designed to monitor learning progress toward predetermined instructional goals and identify children who may require additional instructional support.

Although many studies have explored Mandarin ORF performance levels at each grade, the reading probes used are generally adopted from textbooks. However, results have indicated that the curriculum-based ORF is 50 characters per minute more than that measured by the standard probe assessment. This suggests that reading fluency levels are affected by the equivalence of the reading probes. This study is the first to use standard probes to investigate the developmental and expected growth of Mandarin ORF in the second and third grades of elementary schools.

The number of correct words read per minute (the level) and the growth rate of student reading fluency are two key indicators of reading progress. However, few studies have investigated the growth rate of Mandarin ORF. Early studies tended to assert that within-year ORF growth is linear rather than quadratic, but recent research has indicated that growth rates differ within a school year. This study examined the growth model of reading fluency and estimated its growth rate.

Students with a large vocabulary tend to have a higher ORF level. According to the Matthew effect, students with higher ORF levels have a greater growth rate. Previous Mandarin ORF research does not support this inference. Furthermore, the correlation between initial ORF levels and growth rates has not been reported. However, this issue is important; if students with different ORF levels have different growth rates, then appropriate learning growth goals should be set based on their initial abilities when planning interventions. This study tracked the development of ORF skills over the course of one academic year to investigate the following research questions: (1) What are the growth models of ORF for second- and third-grade students in elementary schools? (2) What is the growth rate of ORF over the semesters for second- and third-grade students? (3) Is the growth rate for reading fluency related to the ORF level at the beginning of the semester?

The participants comprised 57 second-grade and 51 third-grade students and 53 girls and 55 boys. The students underwent an ORF test (Chang, 2018). For each grade, 19 reading probes, with a text length of 350–400 characters, were used. The alternate-form reliability was between .80 and .89. The generalizability coefficients for the second- and third-grade conditions were approximately .84 and .85, respectively, when one reading probe was used to estimate ORF performance. When three reading probes were used, the generalizability coefficients were .94 and .95, respectively. The correlation coefficients between ORF score and reading comprehension test score were .60 and .67, respectively.

The ORF tests were administered at 3-week intervals during each semester in the 2012 academic year. Each student took ORF tests at 13 time points, with three reading probes administered at the beginning of the school year (pretest), at the end of the first semester (midterm test), and at the end of the second semester (posttest). The average score across the three reading probes was calculated to obtain reliable growth rates. At each of the other time points, a single ORF score was collected.

Eight examiners were trained in the administration and scoring of the ORF tests in a 4-hour workshop, during which the assessment and scoring procedures were introduced and the examiners practiced the procedures in pairs to ensure they could conduct the procedure reliably. Overall interrater reliability was .99 for this study.

A hierarchical linear model was used to assess the goodness of fit between the data and five models: the null model (M0), random intercept model (M1), random coefficient model (M2), piecewise growth model (M3), and quadratic growth model (M4). The intraclass correlation (ICC) of the M0 model for the second grade was .73, meaning that 73% of the difference in reading fluency scores originated in differences between students (level 2) and 27% of the variance originated from differences within individuals (level 1). The ICC for the third grade was .81. The results suggested high heterogeneity between students and the necessity of a hierarchical model.

Among five models, a piecewise growth model had the lowest deviance value and thus the best fit with the data. Thus, the children had different growth rates in the two semesters within the same academic year. On average, the number of characters the second and third graders could read correctly per minute was 109 and 137 at the beginning of the academic year, respectively. At the end of the third grade, the mean ORF was 165 characters. In the fall, the second-grade ORF growth averaged 13 characters per semester, whereas the average was approximately 18 characters per semester in the spring. The third graders' mean growth rates were 20 and 6 characters in the first and second semesters, respectively. The second-grade growth rates in the two semesters were similar, but ORF gains were more modest during the second half of the academic year for the third grade.

The correlation between ORF and the growth rate for the second-grade students was not significant. However, in the first semester of the third grade, the correlation between ORF and growth rate was significant and had a correlation coefficient of .32. In the second semester, the correlation between ORF and growth rate was -.43. In different grades and semesters, the relationships between initial ORF and growth rate varied. This finding reminds teachers that they should consider the students' grades and initial ORF when setting teaching goals.

The results of this study established the expected ORF level and growth rate for second- and third-grade students, enabling teachers to identify high-risk students that are falling behind in reading to determine the effects of intervention by examining their change in growth rate and adjust their instruction accordingly. Knowing student's ORF scores may help teachers plan reading instruction that is more responsive to the individual. By collecting multiple assessment scores, educators and researchers no longer must assume linear developmental trends because nonlinear functions can be fit to the time-series data. The nonlinear growth pattern suggests that instructional goals derived from a linear change model may require revision and that student growth expectations may require modification at different points in the academic year.

The results only partially support the Matthew effect. At the beginning of the semester, the growth rate of the second-year students with strong ORF was not significantly greater than that of those with weaker ORF. However, in the first semester of the third grade, the level of ORF and the growth rate was significantly and positively correlated. Participants with strong reading fluency grew faster in their reading ability during the fall semester than those with weaker ORF. Therefore, the results may enable educators to recognize the need for more intensive instructional support during periods when rapid growth is expected, which is the second semester of the second grade through the first semester of third grade.

Keywords: oral reading fluency, piecewise growth model, growth rate, standard passages

