

中小學學童生活口語理解評量工具 建構與效度研究

陳美芳

臺灣師範大學特教系教授

吳怡潔

台北市東區特殊教育資源中心
專業支援教師

本研究主要在建構評量中小學學童生活口語理解能力的測驗並進行效度分析。本測驗除可用於評量 3 至 9 年級學童生活口語理解能力外，並期望能協助診斷有閱讀困難者的優勢能力。測量的成份包含表面理解（對話或敘述中提供明顯線索）及推論（需由對話或敘述推論而得）。測驗分為國小中年級（G3-4）、高年級（G5-6）及國中（G7-9）三版本。本研究在台灣北、中、南三區選取 1290 位 3 至 6 年級國小學童、905 位 7 至 9 年級學童建立全國樣本；由於考量東部地區的特殊性，本研究特別在東部地區選取學童建立東部地區常模。在信度方面，本測驗各年級的內部一致性 α 係數大致在 .70 以上，各年級相隔兩週的再測信度在 .66-.80 之間。在效度方面，本測驗與閱讀理解的相關大致在 .40 至 .69 之間，與識字量的相關在 .34 至 .58 之間，與聲韻覺識的相關在 .25 至 .42 之間。大致與閱讀理論及過去相關研究結果一致。本測驗評量的是非正式語言理解能力，以小學三年級及五年級學生為樣本的研究顯示，本測驗與正式語言理解能力的相關分別為 .49 及 .69，顯示兩者雖有關但不同，兩者在聽覺理解評量方面或可提供不同的訊息。最後研究者並由閱讀簡單觀點模式，檢驗本測驗對閱讀理解的解釋力與調節功能。

關鍵詞：口語理解、聽覺理解、閱讀理解、閱讀困難

文獻探討

一、聽覺理解在閱讀理解模式中的角色

由兒童閱讀的發展來看，兒童在閱讀之前先經歷聽話的階段，口語能力對多數學童來說，是學習閱讀的先備條件。McCormick (1995)認為聽覺理解是估計閱讀潛能的有效指標。有些學者更主張要估計閱讀能力，宜採用聽覺理解測驗，而非智力測驗（如：Aaron, 1991；Stanovich, 1991）。美國最常用於閱讀診斷的測驗（如：Woodcock Reading Master Test-Revised, Woodcock Diagnostic Reading Battery, Comprehensive Inventory of Basic Skills-Revised），均包含聽覺或口語理解分測驗（洪儷瑜，2005）

過去數十年來，許多研究者及教學者提出各種閱讀歷程有關的理論及模式，其中較為常見有下列三種類型：由上而下模式（top-down model）、由下而上模式（bottom-up model）、互動模式（interactive model）。前述三項模式之共同點在於將閱讀理解之認知運作歷程大致分為高低兩種層次處理：高層次能力意指讀者運用背景知識經驗，針對閱讀主題進行假設、預測、理解；而低層次能力則強調文本為主體，讀者由處理最小語言單位（音素）開始，直到處理較大單位（音節、詞彙、片語和句子等）的過程。除了由上而下、由下而上模式分別認為高或低層次在閱讀歷程中佔有主導地位之外；互動模式即是指高低兩層次在閱讀歷程中相互協助，甚至相互補償的情形。不論模式如何，基本上都包含認字與理解兩部份（柯華葳、李俊仁，1999），有些學者進一步發展閱讀模式，嘗試更明確的描述閱讀相關成分與閱讀理解的關係。

Gough 和 Tunmer (1986) 首次發表「閱

讀簡單觀點模式」(the Simple View of Reading)。Gough 曾在 1983 年指出在正常閱讀情況下，越有效率的讀者對於文本的依賴越少；另外，他認為即便在閱讀歷程中識字解碼與語言理解（聽覺理解）的發生不見得有固定的順序性，識字解碼與語言理解仍可以明確分割開來，而針對這兩種成分之間關係進行分析。之後 Gough 及其同事便整合過去學者研究結果，開始發展「閱讀簡單觀點模式」並試圖詮釋閱讀現象，將複雜的歷程簡化為易懂的方程式以進行運算分析及預測。

「閱讀簡單觀點模式」主張複雜的閱讀認知歷程可以簡單的區分成兩大成分：識字解碼（以 D 代表）和語言理解（以 L 代表）；若將閱讀理解（以 R 代表）、識字解碼（D）以及語言理解（L）看成三個從 0（nullity，無）到 1（perfection，完美）的變項，則三者關係可表示為易於計算之方程式 $R=D \times L$ 。識字解碼（D）和語言理解（L）在公式中相乘的關係顯示了兩者缺一不可的關係，這兩項不得為零閱讀理解（R）才可能存在，也就是讀者必須同時具備兩項能力才有辦法從書面文字獲得意義達到有效的閱讀理解。後續多項研究檢驗「閱讀簡單觀點模式」或由其衍伸而來之其他模式（Carver, 1993; Chen & Vellutino, 1997; Hoover & Gough, 1990; Joshi & Aaron, 2000），結果多數支持該模式確實找到構成閱讀理解之兩項重要成分；Hoover 等人的研究、及 Chen 和 Vellutino 進行的複製研究探究兩成分之間的交互關係，研究結果支持識字解碼和語言理解兩大成分會隨閱讀發展階段而有相互影響，並導致影響閱讀理解能力表現。

Hoover 和 Gough (1990) 將簡單模式中的「語言理解」定義為：「一種能接收詞

彙訊息，並取得句子和篇章解釋的能力。」，可使用聽覺管道接收口語或書面文章訊息之後回答問題的方式測量。由測量方式看，Hoover 和 Gough (1990) 對「語言理解」的定義與其他研究中的「聽覺理解」或「口語理解」相當。

由於過去學者對於語言理解等相關用詞及定義略有不同，本文原則採用各家學者理論或測驗的原始用法。本文中所述的「語文理解」包含書面及聽覺形式的理解；「語言理解」與「聽覺理解」意義相似，皆指對於聽覺管道輸入訊息之理解能力；許多研究也並未區分「聽覺理解」和「口語理解」，因此過去研究或評量工具中語言理解、聽覺理解和口語理解的內涵可能相通，但在本研究中發展的「口語理解」評量工具特別針對聽覺理解中較為非正式、日常生活訊息的理解能力。

二、聽覺理解的成分

Wren (2000) 將閱讀理解的認知基礎成分建構成相當精細的模式圖，他認為閱讀理解包含語言理解及識字兩部份，其中語言理解的測量不應涉及書面閱讀，從指導語到測驗題的呈現均應以口語形式呈現，亦即應以口語形式評量語言理解。孩童受限於識字能力，常無法在閱讀理解測驗中表現理解能力，如果仔細比較孩童的閱讀理解能力、語言理解能力和識字能力，就可清楚孩童閱讀理解的困難是出自理解能力的問題或是識字能力的缺陷。

聽覺或口語理解可由不同向度分析，Wren (2000) 分析語言理解的結構時，將語言理解分為正式與非正式語言兩類，正式語言指較為抽象或較不需考慮上下文的語言，在班級教學中的語言學習多屬此類；非正式語言則是平常日常生活的對話，較依賴

上下文脈絡，所談及的訊息多半是與情境相關或是較具體的事情，朋友與家人間的對話多屬此類。過去已有些研究探究學童在閱讀與聽覺理解的關連，並嘗試發現學童理解的亞型，這些研究中的聽覺理解測量多屬正式語言，但生活中的口語其實多屬非正式語言，或許評量學童生活口語理解能力，有助於發現學童在閱讀中更精細的優勢能力。研究者因此嘗試發展以學童生活對話情境為主的口語理解測驗。

我們也可從語言理解的層次將語言理解分為「表面語意理解」與「推論」，表面理解是指對話或敘述中已提供明顯線索，聽者或讀者掌握這些線索即可理解意義；推論則是較精緻的理解，敘述中未提供完整的線索，聽者或讀者必須自己推敲上下文才可理解。柯華葳 (1999) 將閱讀理解分為部份處理及文本處理，其中部份處理是指處理少量文本，建構初級意義單位，如：字義搜尋；文本處理是針對較長文本的理解，包含「文義理解」與「推論」。柯華葳並以此構念為命題結構，編製閱讀理解測驗。本研究在口語理解層次的構念與柯華葳 (1999) 「文本處理」的兩層次相通。

三、國內聽覺理解的評量工具

國內早期對學童語文困難評量的工具設計，即已將聽覺理解相關能力納入，例如：吳武典、張正芬 (1984) 編製之「國語文能力測驗」，早年是學障資源班鑑定、語文低成就學童篩檢、語文學習困難分析使用的主要工具 (洪麗瑜, 2005)；這項測驗包含八項分測驗，其中「聽覺記憶」及「聽覺理解」兩分測驗即屬口語理解能力的測量。近年以語文困難學童為可能施測對象而編製的語文綜合測驗，也有包含聽力評量者，如：洪碧霞、邱上真 (1997) 編製之「國語

文成就測驗」，聽力評量是其中重要部份，測量的成分包含字音辨識、聽覺記憶和短文

理解。近年國內開始有學者將聽覺理解列為獨立的測驗，茲整理如表一。

表一 國內聽覺理解能力評量工具

作者	測驗名稱	測驗內容	適用對象	施測方式
陳美芳 (1997)	聽覺理解測驗	聽句子選圖片	小二、小五	個別施測
錡寶香 (1998)	國小學童語意與語法測驗	語意測驗：詞彙誤用、詞彙定義；語法判斷修正、造句；口語記憶；理解口語篇章大意並推論	小二、小五	個別施測
陳美芳 (1997、1999)	聽覺記憶測驗	語句記憶複述	小二到國一	個別施測
林寶貴、錡寶香 (2000)	兒童口語理解測驗	聽覺記憶、語法理解、語意判斷、短文理解	小一到小六	個別施測
陳美芳 (2000、2003)	語文理解能力測驗 (聽覺理解部份)	對十篇短文的聽覺理解、閱讀理解	小二到小六	團體施測
楊志堅、蘇啟明、沈文娟 (2006)	學童早期口語理解能力檢測	聽覺詞彙 (選出近義詞)、語意/語法判斷、語句理解、聽覺語詞記憶	小一	團體施測

國內現有聽覺理解的測驗工具的適用對象幾乎僅限於國小階段學童，除楊志堅、蘇啟明、沈文娟 (2006) 的測驗測量口語詞彙及語句外，其他聽覺理解測驗測量的都是正式語言。

四、聽覺理解在閱讀障礙診斷的意義

(一) 聽覺理解與閱讀理解有關

由廣義觀點定義閱讀理解的學者主張探究閱讀理解應同時考慮聽覺理解，如：Aaron 和 Joshi (1992) 將理解分為聽覺理解和閱讀理解，Wren (2000) 主張閱讀理解包含兩大成分：識字解碼 (decoding) 與語言理解 (language comprehension)，並特別強調語言理解的測量應不涉及書面理解，以免孩童因識字能力的限制，無法表現理解能力。

過去學者針對識字解碼、聽覺理解兩大成分對於閱讀理解的重要性方面已進行多項研究，結果皆顯示：識字解碼、聽覺理解

及閱讀理解能力皆隨年齡成長而提升；三項能力間有顯著相關，識字解碼及聽覺理解各與閱讀理解能力有相關 (邱上真，洪碧霞，1997；Catts & Kamhi, 1999; Chen & Vellutino, 1997; Hoover & Gough, 1990; Joshi & Aaron, 2000)。就三項能力在不同閱讀發展階段關係演變部份，多位學者認為初學讀者 (國小中低年級學童) 的識字解碼占閱讀理解能力較為重要的比重；而隨著學童識字解碼能力提升至精熟程度，聽覺理解能力將會漸漸成為影響閱讀理解能力較大的因素 (Perfetti, 1985；Stanovich, 1991；Vellutino, Scanlon, & Tanzman, 1994)。

台灣地區的研究也顯示聽覺理解與閱讀理解間有正相關 (如：劉信雄和曾世杰，1996)，陳美芳 (1997、2003) 的研究發現，在國小階段，聽覺理解或聽覺記憶測驗與閱讀理解測驗的相關大致在 .4 至 .6 之間。陳美芳 (1997、1998、1999) 針對國語文低能力組學童進行三年追蹤研究，發現聽覺理解與

記憶低落的兒童在日後閱讀理解能力低落的比例頗高。

(二) 聽覺理解能力對學習的影響

由溝通的觀點分類語言，聽和讀都是屬於接收性語言(receptive language)，接收性語言是表達性語言的基礎，就學習而言，接收性語言具關鍵影響力。無論聽覺理解或閱讀理解有缺陷，對學習皆會造成重要負面影響。比較學生在聽覺理解與閱讀理解的表現，可發現學生優勢管道，並作為補救教學的重要參考。

陳美芳(2000)曾編製一套語文理解測驗，其中包含題型與難度相當的閱讀理解與聽覺理解兩部份，以常模資料，由差異標準誤推估學生閱讀理解與聽覺理解的得分差異，在.05的顯著水準下，各年級優聽型(聽覺理解顯著優於閱讀理解)的檢出率在7.12%至15.80%之間，優讀型(閱讀理解顯著優於聽覺理解)的檢出率在0.72%至4.49%之間。陳明仁(2002)以該測驗為評量工具，發現無論國語文或自然學科的低成就學童中都有約兩成以上的學童是優勢管道偏好型，其中多為優聽型。陳明仁的研究也發現針對優聽型學童提供利於其優勢管道的評量調整，學生會得到顯著的幫助。Spear-Swerling(2005)主張在三年級以前的閱讀教學應同時考慮促進學童口語能力的發展、對年紀更長的閱讀障礙學童應善用口語理解活動以增進其閱讀理解能力。這些研究均顯示聽覺或口語理解對學習閱讀的重要性。

(三) 閱讀障礙亞型的研究發現

有些研究者關切閱讀障礙的不同類型，其中口語困難的學習障礙亞型逐漸受到重視(洪儷瑜, 2005)。許多探討閱讀困難的研究，從閱讀理解與聽覺理解的比較著手，嘗試發現閱讀障礙亞型，這些學者認為

這樣的比較有助於瞭解閱讀理解困難的根源。邱上真及洪碧霞(1997)在臺灣地區的研究發現，如果以聽覺和書面閱讀能力為向度區分閱讀困難學童，大致可區分為：均衡型(聽覺與書面表現均差，此類型人數最多)、聽覺型(聽覺優於書面，人數次多)、書面型(書面優於聽覺，此型人數最少)。

五、文獻綜合分析

由文獻分析顯示，有些持廣義觀點定義閱讀理解的理論，強調閱讀理解包含聽覺理解，「閱讀簡單觀點模式」更將複雜的閱讀認知歷程區分成識字解碼與語言理解(聽覺理解)。實徵研究證據也支持聽覺理解與閱讀理解有值得注意的相關。聽覺理解的評量對診斷閱讀障礙亞型有重要幫助、閱讀理解能力對學習有影響，若能因應學生優勢學習管道調整教學或評量，學生學習效果可提升。整體而言，研究證據支持聽覺理解的評量對瞭解學生閱讀困難、診斷閱讀障礙學生有重要意義。學生口語能力是未來閱讀的基礎，口語能力的學習是由生活中與人互動開始，屬於非正式語言；但綜觀國內現有聽覺理解的評量工具，缺少對非正式生活口語的評量工具，已發展可流通使用測驗的適用對象幾乎均限於國小階段學生。本研究希望針對現有工具的不足，發展適用於國小中年級至國中階段學生的口語理解(非正式語言或生活口語理解)評量工具，希望對研究者瞭解閱讀理解、實務工作者診斷閱讀障礙有些貢獻。

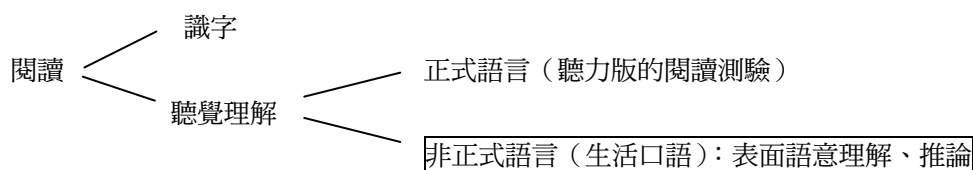
研究方法

一、本測驗編製之構念

綜觀國內聽覺理解的評量工具，可發現近年來雖陸續有些學者發展測驗以評量學

童聽覺理解能力，但有些測驗是為學術研究而發展，並未正式發行或建立常模；有些測驗缺乏和閱讀理解之間關連性的效度探討；少數可用於閱讀困難診斷的聽覺理解評量工具又僅能用於小學階段。基於以上考量，本測驗的主要目的在發展適用於小學至國中階段，可用於評量學童口語理解能力、診斷有閱讀困難者優勢能力的測驗。聽覺理解能力又可區分為正式語言理解與非正式生活口語理解，目前已發展的評量工具多在評量正式語言理解，本測驗則在評量學童非正式的生活口語能力，如圖一所示。

本研究採取「閱讀簡單觀點模式」，將閱讀認知歷程區分成識字解碼與語言理解。根據此觀點，在閱讀困難診斷時應考量「識字解碼」與「聽覺理解」，由語言的類型分析，語言理解又可分為「正式語言」及「非正式語言」，這兩類又各可包含「表面語意理解」及「推論」兩層次。依據以上理論及語言類型與層次，本測驗希望本測驗能協助形成關於閱讀困難者優勢學習管道的假設，評量者可利用系列閱讀相關測驗，形成如表二的推論，並進一步結合其他評量工具加以檢驗。



圖一 本測驗在閱讀困難診斷的定位

表二 本測驗期望提供的診斷用途：分析閱讀困難者的可能優勢管道

閱讀理解	識字	聽覺理解		診斷目的（形成假設並進一步檢驗）
		正式	非正式	
×	○	×	×	識字無困難、閱讀與聽覺理解都有困難、單純理解的困難？
×	×	○	○	聽覺理解無困難、學習的優勢管道在聽？
×	×	×	○	生活化的口語理解無困難、借助口語引導有助學習？
×	×	×	×	智能障礙、其他多面性的困難？

註：表中「○」表能力正常或超常，「×」表能力落後

二、編製過程

（一）測驗題目編製

本測驗邀請十位教師及大學部高年級學生（國文系輔修特教、特教系輔修國文）參與命題及審查。在命題說明會中詳細說明測驗希望測量的能力、形式與命題原則：

1. 測量能力

表面語意理解（對話或敘述中有提供明顯線索）、推論（需由對話或敘述推論而得）。

2. 測驗版本

測驗分為國小中年級（G3-4）、高年級（G5-6）及國中（G7-9）三版本。

3. 命題原則

（1）測口語理解能力，題目或題幹的語

言力求生活化、口語化。

(2)題目的情境脈絡盡量接近學生的生活經驗。

(3)盡量避免因專業領域的困難而影響閱讀或造成對某些人的優勢。

(4)題目與選項使用的語言盡量不是完全相同。

4. 題目情境

包含學校學習與活動、居家生活及社區生活等一般中小學學童熟悉的生活情境。

5. 審題

題目編製完成後，並進行三階段之審查，第一階段由命題人員互審並修改，第二階段由兩位作者逐題討論修改，第三階段請六位國中小學在語文或特教領域已有相當資歷的教師審題，確認題目辭彙及情境適合各階段學生之生活經驗。

(二) 施測與例題

本測驗的問題情境及題目皆由語音播放，無書面文字；題目選項由語音播放，並提供書面文字。每一問題情境都是一段對話，為清楚區隔對話者，測驗之對話錄音分別由一位女性及一位男性擔任。問題情境播放後，再播放兩個問題，一題測量表面語意理解，一題測量推論。以下為本測驗之例題，其中問題一測「推論」，問題二測「表面語意理解」。

【例題】

(女):「王小明，請你唸下一段課文。」

(男):「老師，他沒有帶課本」

(女):「王小明，為什麼上課前不先告訴老師？你和隔壁同學一起看。」

問題一：「『老師他沒有帶課本』，這句話是誰說的？」

1. 王小明
2. 王小明的同學
3. 老師

問題二：「老師要王小明做什麼事情？」

1. 告訴老師有人沒有帶課本
2. 和隔壁同學一起看課本
3. 看隔壁同學

(三) 預試

1. 第一次預試

本測驗編製完成後，先於台北市古亭國小進行個別施測，建立施測時間與流程，並對錄音結果作初步修正。第一次團體預試於台北縣土城國小進行，選取三至六年級各兩班學童實施測驗，因考慮預試時間太長可能造成的負面影響，各版本均分為 A、B 兩式。本次預試的目的在初步估計測驗題目的難度與鑑別度，並瞭解本測驗在團體施測情境的狀況。經統計分析後，各年級平均通過率約在 70-88%之間，編製者針對部份題意不清或錄音品質不佳之題目進行修正或重錄。

2. 第二次預試

第二次預試樣本取自台北縣、高雄縣及台東縣，各版本均分為 A、B 兩式進行預試，取樣校數及取樣學生人數如表三及表四。

表三 預試樣本校數與人數統計 (A 版)

		年級			
		三	五	七	八
台北縣	錦和國中			37	35
	崇林國中			34	38
	鷺江國中			37	36
	莒光國小	66	65		
	頂溪國小	28	34		
高雄縣	忠義國小	36	31		
	成洲國小	31	32		
	林園國中			39	41
	龍肚國中			32	44
	湖內國中			39	36
台東縣	林園國小	36	34		
	中潭國小	32	30		
	大湖國小	30	33		
合計	知本國中			30	34
	知本國小	26	25		
合計		285	284	248	264

表四 預試樣本校數與人數統計 (B 版)

		年級			
		三	五	七	八
台北縣	錦和國中			72	66
	崇林國中			39	36
	鷺江國中			34	39
	莒光國小	32	23		
	頂溪國小	52	62		
	忠義國小	35	26		
	成洲國小	32	30		
高雄縣	林園國中			35	40
	龍肚國中			33	31
	湖內國中			36	38
	林園國小	35	34		
	中潭國小	33	29		
	大湖國小	31	32		
	合計	275	263	278	281

預試結果題目太容易，因此選題原則：

1. 刪除同一題組中兩題通過率皆超過 .90 者，但有少數題目因鑑別力佳而予保留 (G3-4、G7-9 版各一題，G5-6 版 2 題)。
2. 選鑑別力較佳的題目，盡量選與總分相關在 .2 以上者。各版本選題結果，形成單一版本的正式測驗，測驗時間為 30 分鐘。

三、建立常模

以隨機取樣抽取常模樣本時，因東西部學生人數的差距懸殊，東部地區選到的學生數往往很有限，甚至完全未取到東部學生，加上東部地區學生學習的特殊性與整體落後狀況明顯，例如：94 學年國中學測，台東縣國中三年級的孩子，有 57% 總分在 100 分以下，是全國各縣市的最後一名 (曾世杰，2008) 東部地區使用測驗，對照常模時可能面臨不知如何意義化的困難。在東部地區使用測驗若僅採用全國常模，可能有限

制，例如：未包含該地區學童的常模可能不適用於解釋該地區學童的表現狀況；為特教診斷與安置目的使用測驗時，全國常模可能有過高的檢出率。基於對東部地區測驗特殊需求的關心，本測驗特別分開編製全國常模與東部地區常模。

(一) 取樣方式

全國常模取樣的原則如下：

1. 亂數值的產生方法，皆採取均等分配的方法。
2. 資料來源為教育部統計處 92 學年度國民小學、國民中學校別資料，依縣市排序 (即依照教育部的學校代碼排序) 抽樣，並事先刪除離島地區。
3. 若取到私立學校或班級人數小於 9 人之學校，以相鄰的下一個學校代替。

(二) 常模樣本取樣學校與基本資料

全國取樣的學校包括北中南三區，國小 10 所，國中 9 所。東部的常模是取自花蓮、台東兩縣，包括國中小各 8 所，詳見表五至表八。

全國及花東地區常模樣本在本測驗得分之基本統計資料如表九及表十所示。

表五 全國國小常模樣本人數資料

地區	學校名稱	年級				各區人數
		三	四	五	六	
北區	龍安國小	31	31	30	34	485
	信義國小	32	29	32	28	
	興南國小	32	28	29	31	
	茄苳國小	29	26	30	33	
中區	橋孝國小	34	33	35	32	401
	新平國小	34	33	31	35	
	公誠國小	34	33	32	35	
南區	新民國小	38	39	35	35	404
	佳里國小	34	35	30	32	
	潮和國小	37	33	28	28	
合計		335	320	312	323	1290

表六 全國國中常模樣本人數資料

地區	學校名稱	年級			各區 人數
		七	八	九	
北區	懷生國中	36	27	30	308
	重慶國中	39	37	36	
	中和國中	34	38	31	
中區	北新國中	40	42	43	344
	新光國中	37	38	36	
	斗六國中	38	38	32	
南區	左營國中	37	29		252
	佳里國中	38	39	38	
	潮州國中	35	36		
合計		334	324	246	904

表七 花東地區國小常模樣本人數資料

地區	學校名稱	年級				各區 人數
		三	四	五	六	
花蓮	忠孝國小	33	34	34	35	476
	宜昌國小	30	30	32	30	
	玉里國小	30	27	30	26	
	瑞穗國小	33	24	23	25	
台東	豐里國小	22	25	30	34	429
	建和國小	20	23	22	21	
	馬蘭國小	33	35	31	36	
	福園國小	25	23	25	24	
合計		226	221	227	231	905

表八 花東地區國中常模樣本人數資料

地區	學校名稱	年級			各區 人數
		七	八	九	
花蓮區	國風國中	34	37	38	397
	宜昌國中	28	34	33	
	玉里國中	33	31	31	
	瑞穗國中	34	33	31	
台東區	東海國中	38	36	35	387
	知本國中	28	30	34	
	寶桑國中	33	31	28	
	池上國中	30	33	31	
合計		258	265	261	784

表九 全國常模樣本各年級聽覺理解測驗之基本統計資料

測驗名稱	年級	人數	最小 值	最大 值	平均數	標準差
聽覺理解 測驗 G34	小三	335	2	24	19.77	3.221
	小四	320	10	24	20.86	2.406
聽覺理解 測驗 G56	小五	312	1	26	21.77	3.069
	小六	288	1	26	21.89	3.436
聽覺理解 測驗 G79	國一	334	2	30	24.54	3.476
	國二	324	11	30	25.35	3.543
	國三	246	9	30	25.63	3.719

表十 花東地區常模樣本各年級聽覺理解測驗之基本統計資料

測驗名稱	年級	人數	最小 值	最大 值	平均數	標準差
聽覺理解 測驗 G34	小三	226	5	24	18.45	3.794
	小四	221	3	24	19.53	3.688
聽覺理解 測驗 G56	小五	227	3	26	20.56	3.765
	小六	231	10	26	20.97	3.599
聽覺理解 測驗 G79	國一	258	6	30	23.29	3.974
	國二	265	10	30	24.02	3.764
	國三	261	11	30	24.35	3.430

信效度研究

一、信度

(一) 內部一致性

以全國常模樣本進行各版本分年級之測驗內部一致性分析， α 係數如表十一。本測驗內部一致性係數稍低，可能是生活口語理解測驗受各種生活經驗與情境脈絡影響較大，影響了題目的同質性。有些同時包含識字、閱讀及聽力成分的語文能力測驗，似也有聽力部份內部一致性較低的現象，如：邱上真、洪碧霞（1997）發展的國語文能力測驗，在三、六年級也都有聽覺測驗內部一致性最低的現象。

表十一 各年級測驗內部一致性分析

	年級	人數	α 係數
G34 (24 題)	三年級	341	0.740**
	四年級	319	0.601**
G56 (26 題)	五年級	306	0.713**
	六年級	250	0.761**
G78 (30 題)	七年級	332	0.653**
	八年級	310	0.755**
	九年級	234	0.815**

** $p < .001$

(二) 再測信度

本測驗選取台北市某國小三至六年級學生、某國中七至九年級學生，進行相隔兩週之重測，各年級學生人數及再測信度如表十二。

表十二 各年級重測信度 (相隔兩週)

年級	N	重測信度
三年級	26	0.663**
四年級	35	0.676**
五年級	31	0.437*
六年級	30	0.580**
七年級	34	0.723**
八年級	33	0.797**
九年級	40	0.706**

** $p < .001$

表十三 聽覺理解測驗之年級差異分析

	年級	N	M	SD	F 值	組間比較
G34 (24 題)	三年級	335	19.77	3.22	19.808**	四 > 三
	四年級	321	20.80	2.67		
G56 (26 題)	五年級	312	21.77	3.07	1.443	五、六無差異
	六年級	288	21.89	3.44		
G78 (30 題)	七年級	258	23.29	3.97	7.557*	八、九 > 七
	八年級	265	24.02	3.76		
	九年級	261	24.35	3.43		

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

二、效度

(一) 年級差異

本測驗以全國常模樣本進行各版本的年級平均數差異分析，結果如表十三。各版本在較高年級之平均數均比較低年級之平均數稍高，但其中五、六年級學童之平均數差異未達顯著水準，顯示本測驗 G56 版本在鑑別這兩年級學童的差異方面有其限制。

(二) 與閱讀成分之相關分析

本測驗以全國常模樣本，探討本測驗與各相關閱讀成分能力間的相關，採取的效標工具見表十四之說明。本測驗與閱讀成分間的相關分析見表十五。本測驗與閱讀理解的相關大致在.40 至.69 之間，與識字量的相關在.34 至.58 之間，在小學階段，本測驗與識字量評估測驗的相關不低於和閱讀理解的相關，可能由於識字量測驗包含注音與造詞兩部份，造詞能力其實已牽涉理解，但到國中階段，本測驗與閱讀理解的相關在.621 到.692，高於本測驗和識字量的相關 (.341 到.476)；與聲韻覺識的相關在.25 至.42 之間。

表十四 本研究閱讀成分效標評量工具

作者	測驗名稱	測驗內容	適用對象	施測
柯華葳、詹益綾 (2007a)	國民小學閱讀理解篩選測驗 (G2 到 G6 版)	G2 版測量學童命題組合、句子理解和短文理解能力，其餘版本增加多義字題。	G2 到 G6 分別適用於二至六年級學童	團測
柯華葳、詹益綾 (2007b)	國民中學閱讀推理測驗	測量學生直述 (共變、比較/對照、描述) 及類括 (分類、序列) 題型的閱讀推理能力。	八到九年級學生	團測
洪儷瑜、王瓊珠、張郁雯、陳秀芬、陳慶順 (2007)	識字量評估測驗 (A12、A39 版)	測量學生注音及造詞能力。每一題皆為一國字，受試者需在空格中注音，並造一個詞，注音造詞皆對才給分。	版本 A12 適用於一、二年級學生、版本 A39 適用三到九年級學生	團測
曾世杰、陳淑麗、謝燕嬌 (2007)	聲韻覺識篩選測驗	由 CD 逐題播放一個音，受試者在四個選項中勾選正確的注音。	一至九年級	團測

表十五 聽覺理解測驗與閱讀相關成分之關連分析

年級	人數	閱讀	識字量	聲韻覺識
三年級	335	.447**	.430**	.369**
四年級	320	.398**	.282**	.260**
五年級	312	.478**	.420**	.251**
六年級	288	.470**	.441**	.287**
七年級	334	.621**	.476**	.383**
八年級	324	.639**	.450**	.369**
九年級	246	.692**	.341**	.424**

** $p < .001$

(三) 與正式語言聽覺理解及閱讀理解之相關分析

本測驗測量之聽覺理解屬於非正式生活口語理解，過去聽覺理解評量工具測量重點則在短文或句子的正式語言聽覺理解 (如：陳美芳，2000)。研究者主張生活口語的理解與正式語言聽覺理解間應有區隔，兩者可在學童語言理解及優勢管道診斷方面提供不同功能。為檢驗以上主張，本研究在台北縣後埔國小抽取三年級及五年級學童各兩班，實施本測驗及陳美芳 (2000)

發展之語言理解能力測驗 (內容見表一)。兩測驗在三年級及五年級之相關請見表十六。由表十六顯示，本測驗和正式語言聽覺理解的相關均略高於非正式語言閱讀理解的相關均略高於非正式語言閱讀理解；在五年級的相關數值略高於三年級階段。

表十六 本測驗所測口語理解與正式語言聽覺理解與閱讀理解之相關

	正式語言聽覺理解	正式語言閱讀理解
三年級 (n=69)	.485**	.366**
五年級 (n=67)	.689**	.601**

** $p < .01$

(四) 聽覺理解在閱讀簡單模式中的角色

吳怡潔 (2007) 採用常見字流暢性測驗 (洪儷瑜，王瓊珠，張郁雯，陳秀芬，2007)、國民小學二至六年級閱讀篩選測驗 (柯華葳，2006) 及本測驗，分析台灣學童在識字解碼、語言理解和閱讀理解等表現以檢驗「閱讀簡單觀點模式」。結果顯示識字解碼、語言理解及兩成分組合而成之方程式皆可以顯著解釋閱讀理解的變異情形；其中尤

以方程式與閱讀理解之相關情形最為顯著。此外，由表十七可進一步發現識字解碼與方程式皆達顯著相關，然而年級間差異不大；但是語言理解則隨著年級提升與各方程式相關漸增。此結果顯示語言理解影響閱讀理解的重要性會隨著學童年級提升。

表十七 各年級學童識字解碼、語言理解與方程式 $R=D \times L$ 相關情形

	年級	識字解碼 D	語言理解 L
閱讀簡單觀點模式方程式 $R=D \times L$	三年級 N=149	.967**	.560**
	四年級 N=163	.924**	.577**
	五年級 N=170	.948**	.650**
	六年級 N=143	.914**	.783**
	七年級 N=158	.929**	.772**
	八年級 N=148	.973**	.671**

** $p < .01$

吳怡潔（2007）並使用階層式「調節性迴歸」（moderated regression）程序進行調節作用的檢驗。該檢驗首先需建立迴歸公

式，效標變項為閱讀理解（R）；預測變項為 1.識字解碼（D）、2.調節變項語言理解（R）、3.前兩者交互作用項（ $D \times L$ ），迴歸公式為：

$$R = \alpha + \beta \cdot D + \beta \cdot L + \beta \cdot (D \times L)$$

不同變項依序投入迴歸處理。第一層先進入迴歸的是識字解碼（D）、第二層再進入的是語言理解（L）、第三層再加入識字解碼與語言理解兩變項的乘積（ $D \times L$ ），也就是識字解碼與語言理解的交互作用項。本文以五年級學童為例，以階層式「調節性迴歸」（moderated regression）檢定識字解碼（D）與語言理解（R）兩變項之交互作用項是否對閱讀理解變異的預測有顯著貢獻。由表十八可知第三層識字解碼與語言理解之乘積項達到顯著水準，顯示兩變項的交互作用效果成立。以相同方式針對其他年級進行檢驗，結果發現除了五年級之外，尚有六年級、八年級有相同情形，其第三層乘積項之 β 值及顯著水準分別為：六年級 β 值 0.834， $t=2.306$ ， $p=.023$ ；八年級 β 值 1.054， $t=2.344$ ， $p=.020$ 。在二、三、四和七年級部份，交互作用項對閱讀理解的預測力則未達顯著水準。整體而言，中低年級學童的語言理解和識字解碼能力對閱讀理解能力的預

表十八 五年級學童以識字解碼及語言理解預測閱讀理解之變項交互作用分析

	投入變項	未標準化係數 B 值	標準化係數 Beta 值	t 值	R ² 改變量	F 值 改變量
第一層	常數項	5.576		4.195**		
	識字解碼 D	.316	.570	8.984**	.325	80.705**
第二層	常數項	-3.937		-1.643**		
	識字解碼 D	.249	.449	6.884**		
	語言理解 L	.551	.304	4.659**	.078	21.702**
第三層	常數項	14.985		2.102*		
	識字解碼 D	-.430	-.776	-1.762*		
	語言理解 L	-.329	-.181	-.985		
	乘積項 $D \times L$.031	1.495	2.810**	.027	7.898**

* $p < .05$ ，** $p < .01$

測可能較獨立，相互間的影響不大；但較高年級學童的語言理解和識字解碼兩項能力對閱讀理解的影響可能是非獨立的，二者會互相調節彼此對於閱讀理解變異的解釋力。

對本測驗的評述與應用建議

一、研究結論與評述

本研究嘗試發展適用於小學至國中階段，可用於評量學童聽覺理解能力、診斷有閱讀困難者優勢能力的生活口語能力測驗，並進行相關信效度檢驗。在信度方面，本研究進行了各年級的內部一致性及再測信度檢驗，結果尚可接受。在效度檢驗方面，本研究主要提供了以下三方面的支持證據或值得注意與繼續研究的議題：

1. 口語理解能力與閱讀理解、識字能力間有顯著相關，中學階段與閱讀理解的相關比小學階段高

研究者探究本測驗和不同閱讀成分之間的關係，發現三至六年級階段，學生在本測驗展現的口語理解和閱讀理解之間的相關介於.398至.478之間，和識字量之間的相關接近(.282至.441)，由於識字量測驗需造詞，可能牽涉詞彙理解能力，研究者推測此可能是本測驗和識字量測驗相關偏高的原因。在效度數據，值得注意的是本測驗與閱讀理解測驗在國小階段相關係數介於.398至.478之間，但到國中階段，相關提高到.621至.692。過去多項研究發現識字解碼、聽覺理解及閱讀理解三項能力間有顯著相關（如：邱上真，洪碧霞，1997；Catts & Kamhi, 1999；Joshi & Aaron, 2000）。許多研究進一步發現這三項能力間的關係會隨發展階段而有些改變，初學讀者（國小中低年級學童）的識字解碼佔閱讀理解能力較為重要的比重；而隨著學童識字解碼能力提升

至精熟程度，語言理解能力將會漸漸成為影響閱讀理解能力較大的因素（Perfetti, 1985；Stanovich, 1991；Vellutino, Scanlon, & Tanzman, 1994）。本研究發展的聽覺理解測驗是過去研究或評量工具較少採用的生活口語能力測驗，結果仍然呼應了過去聽覺理解的研究發現。

2. 非正式語言（生活口語）理解與正式語文理解有中度正相關，與正式語言聽覺理解的相關略高於正式語言閱讀理解

研究者主張生活口語的理解與正式聽覺理解間應有區隔，兩者可在學童語言理解及優勢管道診斷方面提供不同功能。為檢驗以上主張，本研究選取三及五年級學生進行非正式語言（生活口語）理解與正式語言理解的評量，並進行相關分析。結果顯示在三年級階段本測驗所測知生活口語能力與正式語言聽覺理解的相關為.485，與正式語言閱讀理解的相關為.366；五年級階段則分別為.689、.601。研究顯示正式語言與非正式語言間有中度相關，且與正式語言聽覺理解的相關略高於正式語言閱讀理解，或可支持口語的理解與正式聽覺理解間有區隔，兩者可在學童語言理解及優勢管道診斷方面提供不同功能。另一方面，這部份的統計結果仍發現非正式語言（生活口語）理解與正式語言理解的相關在高年級階段高於較低年級階段。

3. 口語理解及識字測驗對閱讀理解具解釋力與調節功能

吳怡潔（2007）以本測驗為語言理解評量工具，分析台灣學童在識字解碼、語言理解和閱讀理解等表現以檢驗「閱讀簡單觀點模式」。顯示識字解碼、語言理解及兩成分組合而成之方程式皆可以顯著解釋閱讀理解的變異情形；採階層式「調節性迴歸」（moderated regression）程序進行調節作用

的檢驗，結果顯示在五、六、八三個年級語言理解對於識字解碼與閱讀理解能力的關係具有調節效果，即語言理解（L）調節了識字解碼（D）與閱讀理解（R）的關係、識字解碼（D）調節了語言理解（L）與閱讀理解（R）的關係。

在較高年級出現語言理解和識字解碼對閱讀理解調節作用，可能由於高年級學童兩大成分間相關較高，因此其中一項成分的進步會帶動另一項成分的表現，作用於閱讀理解能力上的反應就會越強，而產生了互相加乘的影響。或許我們可以進一步推測，當一名學童接受教育年份漸增，具備基礎閱讀理解能力後，其識字解碼能力及語言理解能力需要同時進步，才能在閱讀理解能力方面獲得更顯著的進步，達到進階閱讀理解水準；甚或能以閱讀能力為工具進行其他領域的學習。

二、本測驗應用與限制

本測驗主要在測量三至九年級學童生活口語理解能力，並協助診斷有閱讀困難者的優勢能力。在閱讀困難者的優勢能力評量方面，使用者可參考前述表一的建議，利用其他閱讀診斷系列測驗，形成關於學童優勢能力的假設，並進一步以觀察、評量及教學活動驗證。本測驗也可用於其他特殊需求學童的口語理解能力評量，如：對於亞斯柏格症候群學童，本測驗可配合其他正式語文能力評量，以協助全面瞭解學童語言能力。但本測驗可能有以下限制，使用者應注意：

1. 本測驗評量的是生活口語理解能力，使用者可參考研究者由閱讀理論與語言類型分析所建立的閱讀困難診斷假設構念（見表二），以日常觀察搭配使用其他正式語言的聽覺理解測驗，對學童聽覺理解能力進行全面瞭解。

2. 本測驗通過百分比偏高，比較適合用在特殊需求學童聽覺理解能力的評量。

3. 本測驗採播放 CD 或錄音帶施測，易受施測環境、播音品質、學童專注力、聽力等因素的影響。在這些條件不利的情况下，可能會低估受試者能力。

4. 本測驗初步編製完成，雖已有常模及初步效度資料，但有些編製構念仍待蒐集更強的效度證據。使用者可利用其他工具進行多元評量，以彌補本測驗之不足。

參考文獻

- 吳怡潔（2007）：「閱讀的簡單觀點模式」在中文一般及閱讀理解困難學童之驗證研究。國立臺灣師範大學特殊教育學系碩士論文（未出版）。
- 吳武典、張正芬（1984）：國語文能力測驗之編製及其相關研究。*測驗年刊*，31，37-52。
- 林寶貴、錡寶香（2000）：兒童口語理解測驗之編製。*特殊教育研究學刊*，19，105-125。
- 邱上真、洪碧霞（1997）：國語文低成就學生閱讀表現之追蹤研究（II）—國民小學國語文低成就學童篩選工具系列發展之研究（II）。行政院國家科學委員會專題研究計畫成果報告（NSC86-2413-H-017-002-F5）。
- 柯華蕙（1999）：閱讀困難篩檢測驗施測說明。台北：行政院國家科學委員會、教育部特殊教育工作小組印行。
- 柯華蕙（2006）：國民小學二至六年級閱讀篩選測驗。台北：教育部。
- 柯華蕙、李俊仁（1999）：閱讀困難的理論架構與驗證。載於柯華蕙、洪麗軒編輯：學童閱讀困難的鑑定與診斷研討會論文集。台北：教育部。
- 柯華蕙、詹益綾（2007a）：國民小學閱讀理解篩選測驗。台北：國立臺灣師範大學特殊

- 教育中心。
- 柯華葳、詹益綾 (2007b)：國民中學閱讀推理測驗。台北：國立臺灣師範大學特殊教育中心。
- 洪碧霞、邱上真 (1997)：國民小學國語文低成就學童篩選工具系列發展研究。**特殊教育研究學刊**，15，83-107。
- 洪儷瑜 (2005)：由語文學習困難的評量工具談其概念與運用。載於洪儷瑜、王瓊珠、陳長益主編：**突破學習困難—評量與因應知探討** (2-28 頁)。台北：心理。
- 洪儷瑜、王瓊珠、張郁雯、陳秀芬 (2007)：常見字流暢性測驗。台北：國立臺灣師範大學特殊教育中心。
- 洪儷瑜、王瓊珠、張郁雯、陳秀芬、陳慶順 (2007)：識字量評估測驗 (A12、A39 版)。台北：國立臺灣師範大學特殊教育中心。
- 陳明仁 (2002)：語音唸題對紙筆測驗結果之影響—以國小四年級自然科為例。國立臺灣師範大學科學教育研究所碩士論文 (未出版)。
- 陳美芳 (1997)：國小學童聽覺理解與聽覺記憶能力之研究：不同國語文程度學生的比較。**特殊教育研究學刊**，15，293-305。
- 陳美芳 (1998)：國小學童口語語言理解與閱讀理解能力之關係。**特殊教育研究學刊**，16，171-184。
- 陳美芳 (1999)：國語文低成就學童口語理解能力的發展。**特殊教育研究學刊**，17，189-204。
- 陳美芳 (2000)：語文理解能力測驗之編製。載於柯華葳、陳美芳撰：**中文閱讀歷程評量及相關測驗編製報告** (3-24 頁)。台北：國立臺灣師範大學心理與教育測驗研究發展中心。
- 陳美芳 (2003)：語文理解能力測驗之發展與效度分析。**特殊教育研究學刊**，24，1-14。
- 曾世杰 (2008)：閱讀補救教學研究—臺東縣的經驗分享。4 月 22 日於國立臺灣師範大學特殊教育系特殊教育研究專題發表之投影片資料。
- 曾世杰、陳淑麗、謝燕嬌 (2007)：聲韻覺識篩選測驗。台北：教育部。
- 楊志堅、蘇啟明、沈文娟 (2006)：學童早期口語理解能力之檢測。**師大學報：教育類**，51(1)，213-232。
- 劉信雄、曾世杰 (1996)：閱讀低成就學童及一般學童的閱讀成分分析研究。行政院國家科學委員會專題研究計畫成果報告 (NSC83-0301-H-024-009)。
- 錡寶香 (1998)：國小學童語意與語法能力之研究。**國教學報**，10，165-196。
- Aaron, P.G. (1991). Can reading disabilities be diagnosed without using intelligence tests? *Journal of Learning Disabilities*, 24, 178-186.
- Aaron, P. G., & Joshi, R. M. (1992). *Reading problems: Consultation and remediation*. New York: The Guilford Press.
- Carver, R. P. (1993). Merging the simple view of reading with rauding theory. *Journal of Reading Behavior*, 25(4), 439-455.
- Catts, H. W., & Kamhi, A. G. (1999). Causes of reading disabilities. In H. W. Catts & A. G. Kamhi (Eds.), *Language and reading disabilities* (pp. 95-127). Boston, MA: Allyn & Bacon.
- Chen, R., & Vellutino, F. R. (1997). Prediction of reading ability: A cross-validation study of the simple view of reading. *Journal of Literacy Research*, 29, 1-24.
- Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading and reading disability. *Remedial and*

- Special Education*, 7(1), 6-10.
- Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and Writing: An Interdisciplinary Journal*, 2, 127-160.
- Joshi, R. M., & Aaron, P. G. (2000). The component model of reading: Simple view of reading made a little more complex. *Reading Psychology*, 21, 85-97.
- McCormick, S. (1995). *Instructing students who have literacy problems*. Englewood Cliff, NJ: Merrill.
- Perfetti, C. A. (1985). *Reading Ability*. New York: Oxford University Press.
- Spear-Swerling, L. (2005). *Components of effective reading instruction*. LD Online. Retrieved Aug. 1, 2005, from <https://www.ldonline.org/article.php?id=1557&loc=111>.
- Stanovich, K. E. (1991). Discrepancy Definitions of Reading Disability: Has Intelligence Led Us Astray? *Reading Research Quarterly*, 26(1), 7-29.
- Vellutino, F. R., Scanlon, D., & Tanzman, M. S. (1994). Components of reading ability: Issues and problems in operationalizing word identification, phonological coding, and orthographic coding. In G. R. Lyon (Ed.), *Frameworks of references for assessment of learning disabilities*. Baltimore, MD: Paul H. Brooks Publishing Co.
- Wren, S. (2000). The cognitive foundations of learning to read: A framework. Retrieved May 10, 2008, from <http://www.sedl.org/reading/framework>.

收稿日期：2008.05.22

接受日期：2008.09.17

The Development and Validity Analysis of a Chinese Spoken Language Comprehension Test

Chen Mei-Fang

Professor, Dept. of Special Education,
National Taiwan Normal University

Wu Yi-Chieh

Teacher, Taipei City East Special Education
Resource Center

ABSTRACT

The purpose of this study was to develop a test for assessing elementary and junior high school students' abilities to understand spoken Chinese. The test which was designed had 3 versions for G3-4, G5-6 and G7-9 students, and the test items were heard by test-takers on a CD player. 1290 G3-6 students and 905 G7-9 students were sampled from three major geographical regions in Taiwan, by way of establishing the norm; a norm for Taiwan's eastern region was established in order to account for the specific assessment needs of that region. The procedures for item construction and for establishing norms, reliability and validity are reported here. The newly-designed test was found to have an acceptable degree of reliability, with internal consistency coefficients above .7. The test-retest reliability coefficients ranged between .66 and .80, and a good fit was found between the validity evidence, past research findings and the Simple View of Reading. Comparing this spoken language test with formal language comprehension tests, the correlation coefficients were within .49 to .69. Thus it was found that these two kinds of language comprehension tests were directly correlated, but that they could offer different information when used for Chinese spoken language or reading assessments.

Keywords: Chinese, spoken language comprehension, reading comprehension, listening comprehension, reading difficulties