

以 TOCFL 連結 CEFR 與 ACTFL：標準設定程序的應用*

藍珮君¹、陳柏熹²、許嘉凌¹

¹ 國家華語測驗推動工作委員會、² 國立臺灣師範大學教育心理與輔導學系

緒論

歐洲共同語文參考架構(Common European Framework of Reference for Languages; CEFR)與美國外語教學委員會能力指標(ACTFL Proficiency Guidelines ; ACTFL 指標)為國際上兩大主要的語言能力架構。CEFR 係歐洲理事會(Council of Europe)於 2001 年所公布的指標，除了在歐洲地區廣泛使用，也翻譯為多種語言版本，如：日語、韓語、俄語等。ACTFL 指標則是由美國外語教學委員會(American Council on the Teaching of Foreign Languages ; ACTFL)於 1986 年提出的指標，此項指標曾在 1999 年以及 2001 年分別針對口語和寫作能力進行調整，目前最新的版本為 2012 年版，並在最高等級之上，再增加了優異級(Distinguished)，ACTFL 指標目前主要通用於美國地區。

由於不同的編制歷史與文化脈落，兩大語言架構的分級方式也不盡相同，CEFR 將語言能力分為三等六級(如圖 1)，ACTFL 指標則細分為五等十一級(如圖 2)，然而，由於能力指標的使用目的皆為測量語言能力的表現，在各個等級也都描述了語言學習者能夠展現的語言能力或能達成的溝通任務，因此在一些等級的能力描述中可觀察到二者間有著許多共通的要素。是故，若能進一步建立起 CEFR 與 ACTFL 指標兩大語言架構的對應關係，無論是對語言教學、學習或能力的評量都將有相當大的助益，也能讓使用兩大架構的美國、歐洲或其他地區產生更多的交流或合作。

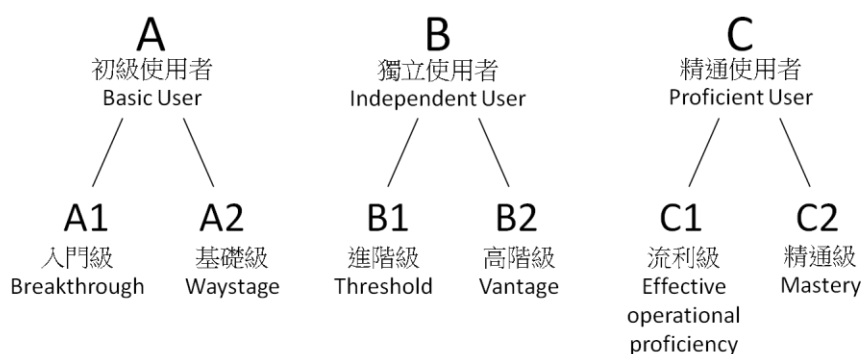


圖 1 CEFR 架構

* 本研究由華測會研究發展組林玲英、曾文璇、孫雪芹、張瑋庭共同完成，特此致謝。

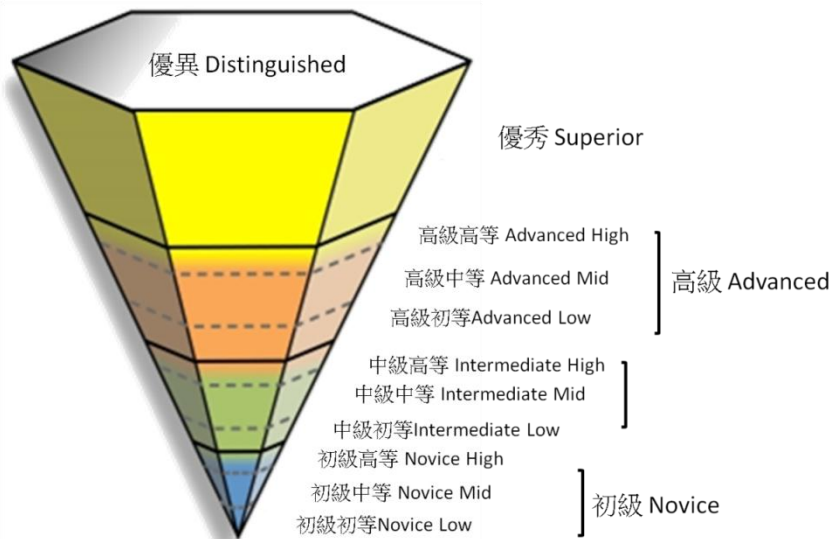


圖 2 ACTFL 指標架構

目前已有一些為數不多的研究探討 CEFR 與 ACTFL 指標等級間的關係，但僅侷限於口語或寫作能力，尚未有關於聽力或閱讀能力的研究發表(Tschirner, 2005; Buitrago, 2006; Martínez, 2008; Slagter, Surface & Watson, 2010)。此外，歐洲理事會與 ACTFL 自 2010 年起已連續四年舉辦座談會，共同邀請美國與歐洲地區的語言專家學者，針對 CEFR 與 ACTFL 指標的內涵及對應關係進行交流，不過至今仍未發表較為明確的結果。

國家華語測驗推動工作委員會(華測會)，長期針對母語非華語之學習者的語言檢定需求研發華語文能力測驗(Test Of Chinese as Foreign Language, TOCFL)的，同樣意識到此一議題的重要性，因此華測會先於 2008 年起參考 CEFR 內容對 TOCFL 測驗題型及架構進行調整，並在 2012 至 2013 年進行標準設定研究，制定出 TOCFL 聽力測驗與閱讀測驗入門級、基礎級、進階級、高階級、流利級與精通級門檻，完整地將測驗等級對應至 CEFR 的 A1 至 C2 六個等級(藍珮君、陳柏熹、林玲英、李嘉真、孫雪芹、曾文璇, 2013)。接著，為了讓使用 ACTFL 指標的美國地區學習者也能夠有所參照，華測會再於 2014 年進行 TOCFL 聽力測驗、閱讀測驗對應至 ACTFL 指標的標準設定研究，目的在制定 TOCFL 聽力與閱讀測驗對應至 ACTFL 指標各級的通過門檻，同時也希望藉由 TOCFL 測驗，為 CEFR 與 ACTFL 指標搭起橋樑，完成兩大語言架構等級間的對應。因篇幅有限，本研究僅先針對 TOCFL 聽力測驗進行說明。

研究方法

本研究程序為採用 Bookmark 法(Lewis, Miztel, & Green, 1996)進行標準設定研究，制定 TOCFL 聽力測驗對應至 ACTFL 指標各級的通過門檻能力值，再與已完成的 CEFR 各級通過門檻能力值進行對照。進行標準設定的首要步驟，是確定通過等級的個數並制定各等級的表現水準描述(performance level description,)

PLD)，PLD 通常是針對最低能力受測者(minimally competent examinee)的能力進行描述，若測驗結果要區分為數個等級，則需制定數個 PLD。先前已提到 ACTFL 指標分級較細，共涵蓋十一個等級，且在初級(Novice)、中級(Intermediate)和高級(Advanced)三等中又再細分出初等(Low)、中等(Mid)以及高等(High)三個次級，考量 ACTFL 指標十一個等級描述的內容，在實際操作上不易再制訂出最低能力受測者的描述，故研究者與測驗專家討論後，在這次研究中調整 Bookmark 法的操作方式。

Bookmark 法原始的進行方式是請判斷成員從一份將試題難度由易至難排列完成的測驗卷中(ordered item booklet, OIB)，依照最低能力描述，標記出符合該能力描述的學習者，答對機率為 67%(RP67)的試題，再依據已知的試題難度參數計算出能力值，並以該能力值做為通過門檻。在本研究中，將 ACTFL 指標的各級能力描述視為典型能力描述，成員改為依照各等級的典型能力描述標記出 RP67 的試題，再依據難度參數計算出典型學習者之能力值，而各等級的通過門檻設定為相鄰二等級典型學習者能力值的中間數。如：以優秀級(Superior)和優異級(Distinguished)的典型能力描述分別標記出 RP67 的試題，計算出能力值後，其中間值即為優異級的通過門檻。

本研究於 2014 年 6 月邀請華語教學相關領域學者與資深華語教師共 10 人參與為期二天的研究，第一天首先介紹標準設定流程與 Bookmark 法的操作方式，說明 ACTFL 指標聽力部分的能力描述，以及簡介 TOCFL 聽力測驗的題型，讓與會成員先行熟悉標準設定作法以及了解能力描述內容。接著，發下聽力測驗題本，請與會成員依據 ACTFL 指標的描述，獨自標記初級初等(Novice Low)至中級高等(Intermediate High)六個較低等級答對機率為 67%的六道試題，並寫下判斷理由。標準設定題本的組成以題庫難度分布為依歸，以 0.5 logit 為區間進行挑題，每個區間大約各 10 題，共 179 題。在成員完成第一回合判斷後，研究者統整並說明判斷結果，說明中提供的資料包括各等級每位成員標記的試題編號與對應之能力值、各等級能力值的平均數、中位數、最大值與最小值，以及依照 2013 年應試者能力值，得出各等級暫定門檻的累積通過比率。研究者帶領成員討論並分享每一等級的判斷依據，待與會成員皆陳述意見後，再各自進行第二回合的判斷，並填寫會議問卷，回收相關文件資料，結束第一天的議程。

第二天研究者先告知成員前一天第二回合六個等級的判斷結果，再發下題本與 ACTFL 指標能力描述文件，繼續進行高級初等(Advanced Low)至優異(Distinguished)五個較高等級的題號標記，程序與第一天相同，共進行二個回合。完成 TOCFL 聽力測驗與 ACTFL 指標等級的對應後，研究者彙整 TOCFL 聽力測驗在 CEFR 各等級通過門檻的能力值，比對出 CEFR 與 ACTFL 指標等級之間的關係。

研究結果

(一)TOCFL 對應至 ACTFL 指標的效度證據

Kane(1994)提出可從程序性效度(procedural validity)、內部效度(internal validity)與外部效度(external validity)三方面檢核標準設定研究的結果，本研究 TOCFL 對應至 ACTFL 指標的標準設定過程，均依照既定的步驟進行，判斷後成員之間亦進行充分的討論與分享，交換判斷理由與依據，回收問卷調查結果顯示成員了解標準設定流程與 Bookmark 法進行方式，且了解 ACTFL 指標各等級能力描述的意涵等研究進行所要求的相關資訊，以上皆顯示標準設定結果具有程序性效度。而觀察第一回合與第二回合的各等級判斷結果後，可發現十一個等級在第二回合的標準差均小於第一回合，表示經過討論後，成員間的判斷趨於一致，可做為標準設定結果的內部效度證據。

(二)CEFR 與 ACTFL 指標的對應關係

研究者整理 TOCFL 分別與 CEFR 和 ACTFL 指標各等級對應的能力值，連結兩者等級間的關係，結果如圖 3 所示，數值部分因涉及測驗機密，故僅提供相對應關係之示意圖；而初級初等(Novice Low)由於 OIB 題本中通過門檻能力值附近的試題過少，研究者認為無法對此一等級作準確的判斷，因此初級初等的結果不納入對照。由圖 3 可知，CEFR 的精通級(C2)略低於 ACTFL 指標的優異級(Distinguished)，而流利級(C1)介於優秀級(Superior)和高級高等(Advanced- High)之間，高階級(B2)在高級高等(Advanced- High)和高級中等(Advanced- Mid)之間，進階級(B1)與基礎級(A2)分別相當於中級高等(Intermediate- High)、中級中等(Intermediate- Mid)，而入門級(A1)介於初級高等(Novice-High)與初級中等(Novice- Mid)之間。

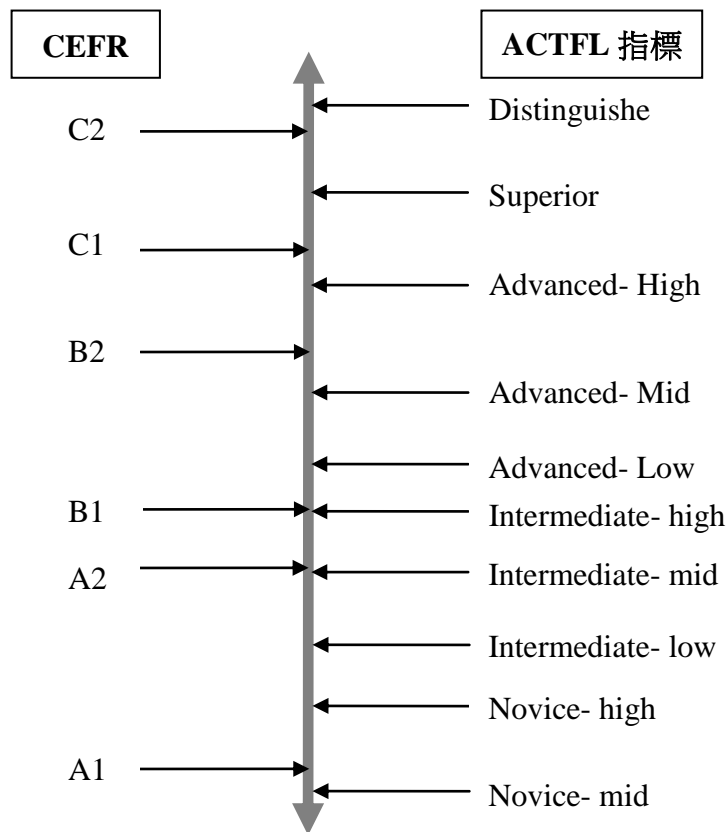


圖 3 CEFR 與 ACTFL 指標對應結果

討論與建議

本研究先以標準設定方式，採用 Bookmark 法將 TOCFL 聽力測驗對應至 ACTFL 指標，制定出各等級通過門檻，並提出程序性效度與內部效度證據，以支持此一對應結果的有效性。接著，再與 2013 年完成的 TOCFL 聽力測驗與 CEFR 等級的對應結果做比較，可說是透過二次 TOCFL 聽力測驗的標準設定程序，建立起國際上二大語言架構的對應關係。過去針對兩大架構的對應研究主要多為比較考生測驗成績與自評問卷的結果，或是就指標描述內容進行對應，本研究首創利用標準設定程序進行語言架構關係的連結，此一方法運用了實徵的測驗資料以及成員的專業判斷，所獲得的結果較為客觀可靠，可提供未來相關研究做為參考。此外，本研究結果提供外界 CEFR 與 ACTFL 指標在聽力部份的等級對照參考，可讓參加 TOCFL 的應試者，除了得知自己 CEFR 等級外，也能了解自身語言程度相當於 ACTFL 指標的等級，讓測驗結果發揮更多的用途，為本研究結果的重要貢獻。

關鍵字：ACTFL、CEFR、TOCFL、標準設定

參考文獻

- Baztán, Alfonso Martínez (2008). *La evaluación oral: una equivalencia entre las guidelines de ACTFL y algunas escalas del MCER* (doctoral thesis). Universidad de Granada. ISBN 978-84-338-4961-8.
- Council of Europe. (2001). Common European Framework of Reference for Languages: learning, teaching, assessment (chap.1 & chap.4). Retrieved January 17, 2007, from http://www.coe.int/t/dg4/linguistic/Source/Framework_EN.pdf
- Council of Europe. (n. d.). *Language versions*. Retrieved August 13, 2007, from http://www.coe.int/t/dg4/linguistic/Source/List_Cadre_traduc.doc
- Kane, M. (1994). Validating the performance standards associated with passing scores. *Review of Educational Research*, 64, 425–461.
- Lewis, D. M., Mitzel, H.C., & Green, D. R. (1996). *Standard setting: A bookmark approach*. Paper presented at the Council of Chief State School Officers National Conference on Large Scale Assessment, Boulder, CO.
- 藍珮君、陳柏熹、林玲英、李嘉真、孫雪芹、曾文璇(2013)。Yes/No Angoff 法在華語文聽力測驗與閱讀測驗的應用，中國測驗學會 2013 年會暨心理與教育測驗學術研討會，臺灣台中。

聯絡方式：martinalan@sc-top.org.tw