

香港普通話口語的評估與測試

周小兵

一、測量、測試與評估

測量 (measurement) 指按一定的規則給事物的屬性指派數字或符號。它包含 3 個要素：(1) 測量的物件及目標，如事物、行為、能力等。能力一般只能通過外顯行為或特徵來推論。(2) 指派數字或符號來代表事物或屬性的量。如桌子 75 厘米高；成績 90 分。(3) 測量依據的標準和方法，如尺子，秤。心理測量有主觀性，如測量口語發音的準繩是標準度還是流利度。

測試 (test) 是指對行為樣本所做的客觀的標準化的測量。它包含 3 個要素：(1) 行為樣本。有代表性，是能力的外在表現。測試語言能力，要挑一部分說的話，寫的句子，對題目的各種反應，等等。(2) 符合實際的測量標準。要考慮題目的難易度、區分度；結果的可靠性程度 (信度)；結果的有效度 (效度) 等。(3) 標準化的測量。題目編制、測量實施以及對分數的解釋方面應有一套嚴密的系統的程式。這樣才能具有統一標準，對不同人的測量結果具有可比性。

評估 (evaluation) 指為作出某種決策而收集、分析資料，並作出解釋的系統過程。與測量、測試相比，評估的含義更廣、綜合性更強。它也包含 3 個要素：(1) 某種決策，即評估的目的，如升級、畢業、入學、給獎學金、錄用、等等。(2) 所收集的信息。這是決策的基礎。一般來說，消息的質量越高，越可靠，相關性就越強，作出正確決策的可能性越大。(3) 對資料的分析、解釋。這涉及決策者本人的能力。資料可靠，如果分析、解釋不正確，也會作出錯誤的決策。

二、測試的類別

測試的分類可根據不同的角度。角度不同，類別不同。

根據學習階段，可分為入學編班測試，隨堂測試，期中、期末測試。

依據用途，可分為成績測試，診斷測試，水平測試，潛能測試等。成績測試的內容一般是剛學過的課程，學什麼測什麼，主要關注學習成功的部分。而診斷測試跟成績測試相反，關注失敗的程度，並借此找出學和教兩方面的補救方法。水平測試考某一方面的整體水平，不是具體某一門課程的測試。國家語委的普通話水平測試，香港考試局的普通話測試，英語的 TOEFL, EPT 等，都是水平測試。潛能測試用於預示學生學習某種語言的潛在能力和天賦。

根據方式，可以分為分離測試和綜合式測試。前者如單項的語音、詞彙、語法測試，後者如聽寫、完形填空、翻譯、作文等。

根據對分數的解釋，可分為常模參照測試和標準參照測試。常模指一群人在一類考試中的成績；為顯示能力的差異，結合其他考生分數來反映某一個考生的分數，如 TOEFL。標準測試只參照一個規定好的標準，不管考生之間的分數差異，如普通話水平測試。

根據閱卷方式，可分為主觀測試和客觀測試。前者根據主觀

判卷，如口語、作文測試。後者答案唯一，如多項選擇題，可機器評卷。

還有一種交際性測試。試題涉及語法和語用，各佔 50%。都正確給滿分；語法正確語用不正確，或相反，給 50 分；都不正確，0 分。

三、香港普通話口語能力的評估與測試

1. 總的原則

是語言評估，而不是文化評估。不要過多涉及內地文化與香港的差異，避免把語言評估變為文化評估。

是準第二語言評估，而不是母語評估。口語方面，廣州話跟普通話的相似度不高：音素 60%（包括聲母、韻母），聲調 33%，音節 25%；辭彙 60%；語法 80%。但書面語方面，相似度很高（超過 90%）。因此，從港人語言學習的角度來說，普通話是準第二語言。

是語交際技能評估，而不是語言藝術表達的評估，更不是語言知識評估。一般不需要測試聲母的發音方法和部位，哪些聲韻母可以結合等；也不需要測試考生朗誦、演講的表現。只需評估用標準或比較標準的普通話進行交際能力。

2. 口語測試的模式

（1）直接型口試

面對面的說話。面試（問答）常有 3 種形式：A. 一對一，考官身兼面試員和評判員；B. 二對一，一個面試員，一個評判員；C. 二對二或三，面試員對二或三名考生輪流提問，並引導考生互相提問或一起完成考試任務。直接型口試的方法主要有 5 種：

朗 讀：音素、單詞、句子、段落。考官注意發音、連讀變詞、輕重格式、語調、停頓與連接等。此法不能完全反映口頭表達能力。補救方法是，讀完再問若干問題讓考生回答，以全面考查其口語能力。

問 答：問題已經備好，多是日常生活話題，如個人、家庭、學習、愛好、假日等。每個話題問 10 個左右的問題。如「體育活動」：

你參加體育活動嗎？你最喜歡什麼體育活動？多長時間進行一次體育活動？喜歡看體育活動嗎？最喜歡看什麼體育活動？你游泳嗎？怎麼學會游泳的？常去哪裏游泳？你建議我進行什麼體育活動？為什麼？什麼運動最難？為什麼？

看圖說話：給考生一張圖畫，讓考生準備幾分鐘，然後根據圖片提問。單幅適宜測試描述能力，多幅適宜測試敘述能力。適合充分展現口頭語言表達能力和想象力。

扮演角色：給出幾種不同的交際場景，讓學生扮演不同角色，並作出語言反映。說話要正確、流利、得體，符合身份、話題的要求。接近於語用測試，適用於中等以上水平的測試。如：

你正在體育場看足球賽。在最緊張的時刻，一個高個子擋住了你的視線，你對他說什麼？

你生病在家，一個朋友來看你，你很感激，該對他說什麼？

討論：適合高水準考生。給幾個題目，讓其任選一個，準備幾分鐘，然後就所選題目講 2–3 分鐘的話。

(2) 半直接型口試

將一些問題錄在磁帶上，讓考生回答。或給出一些圖畫，讓考生敘述。考生的話也錄在磁帶上。優點：省時，省力，效率高；避免考官水平不一致的影響；使考生不易緊張。缺點：交互能力得不到考查。

(3) 間接型口試

通過筆試方法來測量考生口語水平。如通過辨音、完成對話來檢查考生的口頭表達能力。但很難測量學生的真實口語能力。

3. 口語測試的評分

口語測試是主觀測試，評分標準難把握。一般有兩種方法：

(1) 分析法

口試表現分為若干要點：內容，語音，語調，辭彙，語法，流利程度，準確性等。各要點得分總和等於總分。

(2) 綜合法

部分若干項，只根據總體印象給分。

也有兩種方法結合的。

參考文獻

1. 桂詩春 (1988)《應用語言學》，湖南教育出版社。
2. 劉潤清、韓寶成 (2000)《語言測試和它的方法》，外語教學與研究出版社。
3. 呂必松 (1966)《對外漢語教學概論 (講義)》，國家教委對外漢語教師資格審查委員會辦公室。
4. R.M.利伯特 等[美]著 (劉範 等譯) (1983)《發展心理學》，人民教育出版社。
5. 盛炎 (1990)《語言教學原理》，重慶出版社。
6. 周小兵 (1996)《第二語言教學論》，河北教育出版社。

作者簡介

周小兵先生，中山大學教授，博士生導師，對外漢語教學中心主任，普通話教育研究中心主任。兼任：世界漢語教學學會理事；中國對外漢語教學學會副會長，華南分會會長；全國中學語文教育研究會華文教育研究中心副理事長。專業特長：現代漢語，應用語言學（對外漢語教學、普通話研究）。發表論文上百篇。出版專著：《怎樣教你的孩子學說話》（合作），《句法·語義·篇章》，《第二語言教學論》，《現代漢語介詞研究》（合作），《對外漢語教學中的副詞研究》（合作）。主編教材：《粵語區人學習普通話教程》，《中級漢語閱讀教程》。主編論文集《東西方文教的橋樑》。發表散文、小說近百篇。