

## 香港地區普通話水平測試的創立與發展

鍾秉樞

### 一、普通話水平測試的創立

開宗明義，本人在考評局供職，但本文所討論各點，僅為我的一孔之見，並不能代表局方的立場。

本港以普通話水平測試為名的考試曾有兩個，分別由考評局（前稱考試局）和國家語言文字工作委員會舉辦。

考評局舉辦的普通話水平測試分初級和高級，分別創立於 1988 年和 1990 年。其實，初級試全稱普通話水平測試，高級試則被稱為普通話高級水平測試。本文題目的普通話水平測試，指的是前者。為便於說明，後文就把前者叫初級試。

初級試可以說是本港歷史最悠久的普通話測試。上世紀 80 年代中、後期，由於香港回歸祖國已是不爭的事實，港人與內地關係密不可分，在政治、社會、經濟以及生活環境方面均不斷轉變。當時，用人單位（僱主）對如何甄別普通話人才方面有一定的需求。考評局回應了時代的需求，延聘專家、學者共同研討，終於研製出初級試，並在 1988 年開展了第一屆測試。初級試所設定的語言文化水平為中學三年級。

初級試的設立原意是為操粵語的香港人服務的。在研發的階段，工作小組分成聽力組、譯寫組和口試組，最後由普通話測試小組總其成。各工作小組所設計的測試大綱和樣本試卷，都是以操粵語的本地人作為測試對象的。

2006 年，是初級試發展的重要年頭。以前，初級試屬於綜合考試；到了這一年，經過測試專家們的一番努力，初級試已經演變為既發單卷測試證書，又發綜合測試證書的測試了。下文純粹為了使指稱更明確，把 2006 年以前的初級試叫做舊初級試，把 2006 年以後的初級試叫做新初級試，並一一加以介紹。

### 二、普通話水平測試的等級水平設計

舊初級試分聽力、譯寫和口試三份考卷，各卷並非獨立的測試。考生報考初級試，必須照單全收。換句話說，每一卷都算分，都是測試不可分割的組成部分。

舊初級試的設計比較複雜。考卷是三份，但在成績單上羅列的剖象項目卻為七項，分別是譯寫（15%）、聽力（25%）、口試朗讀（20%）、口試短講（20%）、口試會話（20%）、總成績以及聽說能力（即聽力和口試各卷的分數總和）。

測試等級分優異（得分達到 90% 或以上）、良好（得分在 80% 至 89% 之間）、及格（得分在 60% 至 79% 之間）和不及格（得分在 60% 以下）等四級。每一個剖象項目的成績均如此處理。只要考生在某一剖象項目達到某個等級的得分規定，就可以獲得該等級。

考生要取得測試證書，必須具備兩個條件：一、總成績必須達到 60% 以上；二、聽說能力的得分必須達到 60% 以上。

其中第二項要求完全體現了初級試為操粵語的香港人服務這個特點。按照當時學科專家的考慮，漢語拼音或注音符號是操粵語的港人必須掌握的學習工具，它們可以起到正音、記音的積極作用，是學習過程中不可或缺的柺棍，因此，在測試設計上理應佔有一席之地，可是佔分比重又不能太多（佔 15% 差不多了），因為說到底，要判別考生普通話水平的高低，會聽、會說才是必然條件，這是毫無疑問的。會聽、會說與否與言語交際能力息息相關，而拼音工具掌握得熟練與否（記音能力），與言語交際能力倒沒有必然關係。

此外，初級試在確保水平的穩定性方面做了一些考慮。不少考生能較熟練掌握拼音工具，可是說起普通話來卻開不了口。這是測試專家們所無法認同的，這些「水分」必須擠走，才能還初級水平的真面目。基於這種認識，測試專家們別出心裁，設置了一項障礙（hurdle），要求考生必須在符合第一項要求的大前提下，也符合第二項要求。也就是說，初級試要求考生必須也達到聽說能力的得分（即聽力、朗讀、短講和會話的分數總和，合為 85%）的 60%。這一重障礙給舊初級試水平把最後一關，把聽不好、說不好，卻精於拼音的考生拒諸及格的門外，從而使初級水平的穩定性更有保證。

### 三、普通話水平測試的試卷設計

舊初級試的試卷設計也是緊緊扣着為操粵語的香港人服務這條主線的。

強調聽說能力是初級試首要的特點。這不光表現在上述的證書頒發條件上，更重要的是考核內容的綜合性上。

初級試對聽力有較高的要求。聽力卷考核聽寫詞語、口語詞認辨、理解判斷和廣普詞句對譯。從操粵語的香港人的角度看，他們使用普通話時，首先需要聽明白一些普通話詞語，也需要掌握一定數量的常用普通話口語詞（例如：「紅火」、「拉扯」、「現眼」，等等）。另外，作為母語是粵語的語言使用者，他們也需要多了解一些粵語詞句在普通話裏怎麼說。更重要的是，他們必須能夠在語篇層面上準確地聽取言語信息，因此，聽力的核心是理解判斷。初級試要求他們能通過性質不同的語篇（內容類別不同、形式有敘述和對話的差異、長度有短篇、中篇和長篇的要求，所考核的能力各有重點），克服粵語的干擾，根據卷面所提供的選擇項，獲得所需的言語信息。試卷題型門類齊全，從詞句到語篇的理解都兼而有之，考核層次的高低也比較全面。

初級試對口語能力有能傳意、流利、標準的要求。

口試卷考核朗讀（詞語朗讀、對話朗讀）、說話（依題說話或依圖說話）和會話。從操粵語的語言使用者的角度看，他們使用普通話時，需要具備多方面的能力。首先，能否傳意是個關鍵。能傳意的表徵是用詞準確，造句規範，內容言之有物，信息傳達準確，結構有條理。這些是口語能力的核心，是流利和標準的載體。沒有傳意能力，就談不上流利和標準；欠缺流利和標準，傳意能力也沒有根本。兩者是相輔相成的。初級試要求考生具備初步的朗讀、說話、會話能力，能基本表達所需的信息，能運用一些會話策略，一般能使說普通話的人聽懂所說的話，尚算順利地達到交際目的。我們通過不同的題型，驗證考生在語言形式（語音、詞匯、語法）和言語表達的各方面是否達到有關要求。

流利的表徵比較明顯，如語速正常，停頓合理，沒有明顯的中斷等；如果說話流利，而不符合規範要求，這是不足取的。據觀察，不少考生對此認識有所不足。有些人則只重視言語交際能力，而忽略了語言基本功的訓練，說起話來非常流利，語速很快，可是仔細一聽，流利是夠流行了，就是缺乏「那種味兒」。這種缺乏韻味的「中介語」，在實際語言生活中，除非錯得很離譜，否則並不一定構成甚麼溝通障礙。這些人還可能因為說得流利，而贏得一般人的好評。但非專業人員對標準語的認識和理解，到底和專家的不一樣。

在實際語言運用中的一些有利條件，比如說，身體語言、說話雙方的熟悉程度、是否有某種默契而心照不宣等，在測試中都不存在。考生只能依靠本身的實力，還得克服一些限制，如臨場反應等，去證明自己的普通話能力。這是測試中所表現的能力與實際語言運用的相異之處。

初級口試等級標準是統一的；只要考生符合有關等級的規定，就可得該等級。僅就口試能否取得及格或以上成績而言，普通話說得流利只是基本要求而已，能否說得標準才是關鍵所在。

標準的表徵是考生在語言形式（語音、詞匯、語法）運用比較熟練。

唯語音論是有所偏頗的。在考核考生的過程中，我們經常碰到有些人很注重發音，每個字都咬得很準，吐字歸音也算到位，可是不論朗讀，還是說話，語流（輕重音、停頓等）達不到要求，或逐字蹦出，或經常回讀等，顯然與自然語言有所不同。這就是只見樹木，未見森林，認識有所偏差的緣故。這些人在語言生活中，需要加強普通話言語交際的訓練，可能要吃很多苦頭，才能修成正果。

關於規範不規範的問題，我們需要辯證地對待，初級試考核的不是無可挑剔的普通話。在語音方面，他們需要具備基本的普通話語感，有一定的發音能力、認字正音能力，雖然發音有明顯的方音痕跡，只要不至於影響理解，就可以接受。其次，在詞匯、語法方面，需要具備使用規範語句的一般能力。

強調拼音的工具性和兼容性是初級試的第二個特點。譯寫卷選取較常用的普通話字詞做考題。從操粵語的香港人的角度看，他們學習普通話時，必須注意廣普的語音對應規律，哪些有規律可循的，一學一大片，這是初步的工夫，初級試一般少考，甚至不考。再往深處學，哪些是例外的，逐個記憶；哪些是粵籍考生的普通話發音難點，都得牢牢記住。初級試主要考核的是後

一類字詞的發音。試題注重考核語音認辨能力（約佔 75%），也考核漢語拼音或注音符號的標音能力（其中認辨約佔 8%，實際標音約佔 17%）。

此外，譯寫卷對拼音系統持開放的態度。卷面上漢語拼音和注音符號均並列，供考生使用。而考生答案給漢語拼音和注音符號均可接受。目前，在國際範圍內，漢語拼音處於絕對的強勢地位，注音符號的使用範圍日見狹窄。據粗略的估計，在初級譯寫卷中，使用注音符號的考生估計不到 0.5%，但注音符號的地位仍與漢語拼音平起平坐，不分軒輊。

#### 四、普通話水平測試所面臨的挑戰

自初級試應運而生後，上個世紀 90 年代，不同的普通話測試相繼在本港推出，內地一些普通話測試也進入本地測試市場。為便於後文的討論，現先將各測試的基本介紹表列如下：

測試名稱	主要測試對象	測試目的	測試單位
普通話水平測試 (簡稱「初級試」)	本港操粵語的人	用人單位甄別普通話人才的標準	考評局
普通話高級水平測試 (簡稱「高級試」)	本港操粵語的人	用人單位甄別普通話人才的標準	考評局
普通話水平測試 (簡稱「PSC」)	暫以師範和廣播兩大系統的人為主流。外國人也可報考。	廣播員及教師系統上崗的普通話水平證明	國家語委
普通話水平考試 (簡稱「PSK」)	理大學生和社會人士	學生：普通話水平出關試； 社會人士：水平測試	香港理工大學
漢語水平考試 (簡稱「HSK」)	外國人	進入中文系前的甄別考試	國家對外漢語辦公室
會考普通話	本港中五在校生、自修生	中學課程	考評局
教師語文能力評核 (普通話)	本港在職教師/ 準教師	教師上崗的普通話水平證明	香港特別行政區政府 (考評局承辦)

高級試比初級試晚出兩年（1990 年），至 2008 年正式終止測試，一共舉辦過十九屆測試。高級試與初級試齊頭並進，在本港推廣普通話方面互相促進、相輔相成。高級試所設定的語言文化水平為中學五年級。它有聽力、譯寫、口試三卷，卷名與初級試的相同。其成績評定也與初級試大同小異（剖象項目相同，不過佔分比重有所不同）。高級與初級的差異，主要表現在試題的難度上。兩者既有量的差異（如同名各卷的題量、考核時間），更大的的是質的差異（例如：詞語的難度、語段的難度、所考核能力的側重點，等等）。近年來，在試卷編制、題型設計、試題難度、測試等級等方面，高級試與社會的需要漸漸脫節，考生數目越來越少，最後只能退出測試的舞台。不過，它在本港推廣普通話歷史上的功績，是不容抹煞的。

PSC 暫以內地的廣播員和教師作為測試對象，但也同時在本港各大學設點測試，各界人士均可報考。1996 年，國家語言文字工作委員會主辦的普通話水平測試（PSC）在香港大學設點測試。此後協辦單位越來越多，至今

所有大學都協辦 PSC。PSC 是國家級的普通話測試，它具備權威性，而且以測促學，備有備考教材（包括光碟），有助於提高普通話能力，因此受到本港學習者歡迎。另一方面，PSC 水平等級鮮明，在甄別普通話人才方面給用人單位極大的方便，有極高的認受性。可以說，它對本港社會的普通話水平起到了積極推動的作用。該試的特點是注重語音能力的考核，它所走的路，與面向操粵語的港人、注重綜合能力的初級試和高級試有明顯的差異。

其它普通話測試各有特色。PSK 是香港理工大學的出關試，也同時供各界人士報考。它的三級八等標準與 PSC 的三級六等有一定的可比性，往下還延伸了兩級，對初學者具有激勵作用。教師語文能力評核（普通話）屬於教師的專業測試，非教學專業人員也可以報考。它的前景如何，尚待觀察。HSK 以來華學習的外國人作為測試對象，也在本港設點測試，但它的影響比較小。在芸芸普通話測試中，規模最大的要算會考普通話了（近幾年，每屆約考五千人）。它是供中五考生選考的科目，它的卷一聽力（佔 20%）、卷二譯寫（佔 10%）和卷三口試（佔 40%）沒有指明具體考試內容，但為了與 O-Level 接軌，等量齊觀，它多出了卷四語言知識與應用的考卷（佔 30%），考生需要對考試課程有一定的了解才有可能考得好。這些測試在本港各行其道，遍地開花，使本港的普通話測試呈現一番新氣象。

上述各種測試既是初級試的競爭對手，又是促使初級試求變的鞭策力量。其實，初級試更大的衝擊來自測試本身的變化。

先說測試本身的外部環境轉變。初級試是測試專家們傾注了極大的熱情而創立的心血結晶。在它出現以前，本港在普通話測試還是一片空白，它可以說填補了這片空白。初級試的定位，是為在本港生活操粵語的人（中三文化水平以上）服務，這是明明白白的。前人披荊斬棘，筚路藍縷，開闢出一片普通話測試園地，真的功不可沒。但隨着時間的推移，考生的來源有相當大的變化。雖然，我們沒有考生語言背景的統計數據，但據觀察，越來越多的非粵籍考生參加這個測試。這種客觀環境的轉變，使初級試的測試對象漸漸模糊化了，某些題型的設計初衷至此顯得落後於形勢，缺乏應有的針對性，例如：聽力廣普對譯，內地來港定居的北方人聽粵語詞句都成問題，還怎麼翻譯呢？

測試成績，是應該朝化繁為簡的方向走的。舊初級試成績比較詳盡，但失諸瑣碎，對用人單位不一定合用。從學科的觀點看，舊初級試的特點在於它的綜合性、本土性。可是，就因為初級測試證書是一份綜合能力證書，上面羅列了七個剖象項目（聽力、譯寫、朗讀、短講、會話、總成績、聽說能力），是否符合用人單位的實際需要，很難說。用人單位的主管不是普通話教師，也不一定看過考卷，更不要說能理解其內涵了。他們要知道的，是應聘者的普通話說得怎麼樣，其能力能不能符合崗位的需要，這就夠了。

再說測試本身的內部也需要做一些工作。舊初級試中的不合理成分也需要排除。在聽力卷中，應該儘量減少非普通話因素干擾的題型，例如，聽寫涉及識字及寫正字的能力，廣普對譯涉及對粵語的理解能力，這些都是與普通話核心內容無關的東西，應該予以剔除。

綜上而言，初級試應該尋求轉變，以迎合實際需要。由於近年來政治、社會和經濟等方面的變化急劇，以及市民對普通話的認識有所提高，考評局決定委託專家重新檢視初級試，以期它更能夠配合當前本港的普通話運用情

況和迎合市民的需要。在學科專家的共同努力下，新測試大綱《普通話水平測試測試綱要及能力等級說明》就是在這種情況下編制出來的。到 2006 年，新初級試正式出台了。

## 五、普通話水平測試的發展

新初級試並非為變化而變化，破舊立新都是有所依據的。

從測試外部看，新初級試的特點是合中有分，分中有合。合中有分是甚麼意思呢？新初級試下面有口試、聽力、譯寫三卷，名稱維持不變，這是合。但每卷已經改為獨立測試，供考生自由報考，這是分。新初級試的三項獨立測試，其能力等級水平與得分（百分比）的關係仍維持不變。

那麼，分中有合又怎麼講呢？新初級試的每項測試，均獨立地按考生的得分評定其等級水平。如考生在某項測試中達到及格或以上的等級水平，就可以獲頒發該項測試的等級水平證書。新初級試的每項測試所評定的及格或以上的等級水平，合起來又可以根據相關的換證規定（有效期為二十四個月），再獲頒發應得的全科測試證書。這就是分中有合。

這樣的安排，給考生保留了選擇的空間。考生可以按照個人的需要、學習的進度等，報考一、兩項，以至三項測試。想拿單科證書的，就按需要報單科；想一步到位的，就報考三項獨立測試；希望在兩年內獲得全科證書的，可以及早做好安排，等等。考生可以自行訂立應考策略，掌握了參加測試的主動權。

用人單位也會對這樣的安排表示歡迎。舊初級試綜合性的特點跟不上形勢，前面已經說明了。新初級試的單科證書成績一目了然。用人單位要判斷應聘者的單項能力，可以參考他的單科證書，非常有針對性；要判斷其綜合能力，查一下全科證書就行了，簡便易明。

從測試內部看，新初級試部分試卷在題型、內容的詳略、測試時間、測試形式等方面做了一些調整，以配合時代的需求。工作小組測試專家們認定，新初級試各試卷所測試的水平大致上維持不變。下面按試卷逐一說明。

新、舊聽力試卷的詳情見下表：

	舊初級試聽力	新初級試聽力
測試時間	30 分鐘	40 分鐘（減掉開頭的應考說明之後，約為 35 分鐘。）
測試內容與佔分比重	語音層面： 聽寫約佔 14%	語音層面： 語音辨析約佔 6%
	詞句層面： 口語詞認辨約佔 8.5% 廣普對譯約佔 21%	詞句層面： 詞義理解約佔 10% 語句判斷約佔 10%
	語篇層面： 理解判斷約佔 56.5%	語篇層面： 理解判斷約佔 74%
測試形式	聽寫和廣普對譯，需要用文字回答。 口語詞認辨和理解判斷，是多項選擇題。	全卷均為多項選擇題。

新初級試聽力比舊初級試聽力有幾點改善之處。第一、理解判斷的佔分比重由一半強增加到約四分之三，體現了它作為聽力核心的地位。第二、刻意排除了一些非普通話因素的干擾，直接指向聽力測試的本質。從考核內容來說，聽寫和廣普對譯，需要用文字回答，考生會因識字能力、寫正字能力、粵語的理解能力欠佳，而影響成績。而這些因素其實與普通話能力的高低不一定有關係。另外，從考核形式來說，全卷變成多項選擇題，使每一位考生都立足於同一條起跑線上，不會受到語言文化水平、粵語能力等因素的制約。第三、新試卷仍然從操粵語的香港人的角度設題，但把廣普對譯取消了，使外省考生不會處於先天的不利形勢（聽不懂粵語，其普通話表達能力無從發揮），增加測試的適應面以及公平性。第四、增加了一些語用元素，使測試更具實用性。語篇在短篇（約 150 字）、中篇（約 200 字）、長篇（約 450 字）之外，加上超短篇對話，就此提問與語用有關的問題。

基本上，新初級試譯寫的變動較小，主要是在題量方面（一、略微增加了試題數目；二、增加漢字詞語寫拼音/注音。）詳見下表：

	舊初級試譯寫	新初級試譯寫
測試時間	30 分鐘	30 分鐘
測試內容 與佔分比 重	語音判斷 選擇題約佔 16.5%	語音判斷 選擇題約佔 10%
	拼音/注音寫漢字： 語篇中的詞語約佔 53.5% 拼音/注音詞語約佔 13.5%	拼音/注音寫漢字： 語篇中的詞語約佔 48% 拼音/注音詞語約佔 12%
	漢字寫拼音/注音： 漢字詞語約佔 16.5%	漢字寫拼音/注音： 語篇中的漢字約佔 15% 漢字詞語約佔 15%
測試形式	選擇題約佔 16.5% 填空題約佔 83.5%	選擇題約佔 10% 填空題約佔 90%

## 新、舊初級試口試的詳情見下表：

	舊初級試口試	新初級試口試
測試時間	備試時間 9 分鐘， 測試時間 9 分鐘。	備試時間 8 分鐘， 測試時間 8 分鐘。
測試內容 與佔分比 重	朗讀：約佔 33.3% 音節（30 個） 對話（約 100 個音節）	朗讀：約佔 40% 音節（50 個） 對話（約 100 個音節）
	短講：約佔 33.3% 依題說話或 依圖說話。	短講：約佔 30% 依題說話或 依圖說話。
	會話測驗：約佔 33.3%	會話測驗：約佔 30%
測試形式	由兩位主考負責考核。	由兩位主考負責考核。

新初級試口試有兩項變動。其中之一是測試時間由原來的四分鐘減少一分鐘。據測試專家的看法，測試時間以儘量精簡為原則，兩分鐘就能測出水平的，不必測三分鐘；三分鐘能測出水平的，不必測四分鐘。事實上，兩分鐘太短，四分鐘又太長，基於這種認識，新初級試口試的測試時間定為三分鐘。另一項變動是取消會話，增加情景對話題型。這得從兩方面說。第一、為甚麼要取消會話題呢？第二、為甚麼要增設情景會話題呢？

先回答第一個問題。考生發言的質和量，是會話評分的依據。會話測試時，除非考生不說話，中斷發言幾次，說話結巴而間接使發言的量減少，否則這個量化評分是沒有意義的。會話題之所以需要取消，與三個因素大有關係。

第一個是主考能力因素，可能對考生發言的質有影響。能當主考的本身應具備相當高的素質（普通話水平、普通話教學經驗和測試經驗等）。但主考之間的一些個體差異問題是客觀存在的。例如：對測試的認識、言語表達能力、言語反應能力、評分能力，等等。這諸多的因素，都足以使不同的主考在評分上帶有一定程度的主觀性。舉例說，一位主考說話簡潔到位，善於引導考生發揮自身的言語水平，另一位則說話囉囉嗦嗦，一句話分成幾句說，他們孰優孰劣，不言而喻。再舉一例，一位主考善於層層深入，能就考生的具體情況檢驗其實際水平，問題的難度螺旋上升，另一位則平鋪直敘，停留在同一個難度上發問，誰比誰強，最清楚不過了。由此可見，考生表現很受會話雙方的互動關係影響。

第二個因素是測試前的培訓工作。整體評分的穩定性從個人開始，抓好培訓工作是個關鍵。要做好培訓工作，知易行難，必須看資源分配能否配合，不然，徒有空想，而落不到實處也沒有用。主考是否經常測試，對主考工作的熟練程度也很不一樣。此外，他們必須評分穩定，不然，給分飄忽就難辦了。

第三個因素是評分程序。簡化評分程序有助於提高整體評分的穩定性。主考參考資料寫得再仔細、再清楚，要是在考場內主考要兼顧的事情太多，心神不定，其評分的準確性可能存在疑問。

下面回答第二個問題。環顧各普通話測試的說話題，除了會考普通話和初級試外，都不設情景會話題。情景對話是一種很有特色的題型，考核考生的表達能力（言語反應能力、內容的得體程度、流暢性和條理性等）和對話



言形式運用的熟練程度。備試時，考生只看到簡略的情景。實際測試時，主考根據主考專用試卷的指引，引導考生進入情景對話，然後根據主考試卷所預設的六個問題發問，一直到測試結束為止。

預設的試題經審題委員會精心處理。它們在難度上、考核水平的遞進上都做了充分的考慮，能讓考生充分發揮所見，表現其成段說話的能力。主考用語力求精簡、扼要，以不多佔用考試時間為原則。這比會話的隨意發問顯然有進步。

統一考題。主考拿着主考專用試卷，按考生的應考情況或照本宣科，或靈活處理，以適應臨場需要，這樣可以減少主考因素的影響，另一方面，使用同一份試卷的考生都會被問大同小異的考題，測試的公平性大大提高。

測試程序得以簡化。既然主考專用試卷裏面基本上提供了發問的基礎，主考就可以專心做好核心事務（即評分工作），而不需要多考慮問甚麼，這就比裏裏外外一把抓強多了。

情景對話補救了會話題的不足之處。如果說會話題是全開放式的題型，那麼情景對話實際上是半控制式的。必須指出的是，情景對話與真實生活中的對話還有所不同，但它已經很接近真實的對話了。從測試的有效程度來看，它顯然比會話邁進了一大步。

## 六、結語

上面介紹的新初級試，既修正了舊初級試的一些弊端，也引進了一些新思維、新方法，使新初級試在保留它的綜合性、本土性的特點的同時，也有所創新、有所發揚。可以說，它在本地普通話測試市場獨樹一幟，希望更多的朋友能夠重新審視它的價值，這是本人的心願。

（本文所論的「初級試」（即由考評局所舉辦的「普通話水平測試」）已於2013年舉辦最後一屆考試。）

### 作者簡介

鍾秉樞先生，於香港考試及評核局擔任普通話評核發展經理。長期從事語文教學、語言測試工作。曾先後於中學、政府機構和大專院校任教中文科、粵語和普通話。