

# 海峽兩岸新聞詞彙對比\*

王珊

## 摘要

華語研究是了解華人社會變化與發展的重要途徑，而新聞作為一種具有真實性、廣泛性和多樣性的媒介，提供了研究語言所反映的社會現象的重要數據來源。本研究利用中文十億詞語料庫 Tagged Chinese Gigaword 大陸和台灣的新聞語料，選取華語常用詞「提高」，對比分析它的主謂關係和述賓關係中的搭配詞及其語義類的異同，探究二十世紀九十年代初至二十一世紀初期海峽兩岸的新聞關注的提高問題。共有詞有助於揭示新聞報道中海峽兩岸人民共同關注的提高現象，而獨有詞則可以更直觀地展示兩岸對提高問題的關注差異。考察這些異同點可以幫助我們從語言層面了解兩岸新聞的華語使用情況，窺探海峽兩岸社會的發展面貌。

**關鍵詞：**詞彙 大華語 新聞 提高 大陸 台灣

---

王珊，澳門大學人文學院中國語言文學系，澳門大學協同創新研究院，聯絡電郵：shanwang@um.edu.mo

\* 【基金項目】本研究受澳門大學資助（MYRG2022-00191-FAH, MYRG2019-00013-FAH）。

## 一、引言

「大華語」是指以普通話 / 國語為基礎的全世界華人的共同語（李宇明，2010）。陸儉明（2019）指出，大華語以普通話為基礎，而在語音、詞彙、語法和語用方面有一定彈性與寬容度，是全球華人的共同語。中國在改革開放後，與各個華語區之間的交流日益頻繁，現代漢語和華語的逐步融合成為不可避免的趨勢（周清海，2016）。研究各個華語區的語言狀況，有助於促進不同地區之間的交流，深化對華語的多樣性的理解。

華語在語音、詞彙、語法上存在顯著差異（湯志祥，1995）。詞彙作為表達思想與傳遞信息的基本語言單位，能夠敏銳地反映社會與文化，是日常交際中不可或缺的元素。因此，詞彙研究成為華語研究的主要領域之一（周清海，2017）。許多學者已在華語詞彙方面展開了深入探索。例如，田小琳（2004）提出了「社區詞」的概念，並考察了香港社區詞的特徵；邵敬敏、石定栩（2006）考察了「港式中文」詞彙的發展歷史、性質、變異手段和變異原因；郭熙（2012）比較了「講話類」詞群在不同華語社群中的變異情況；鄧思穎（2018）提出了全球華語詞彙研究的參數分析框架。此外，《全球華語詞典》（李宇明，2010）和《全球華語大詞典》（李宇明，2016）這樣的標志性成果為了解華語區的詞彙使用提供了重要參考。

儘管如此，對於海峽兩岸大陸和台灣的新聞詞彙反映社會現象異同的研究仍然有限。為深入理解這些差異，本文基於中文十億詞語料庫（Tagged Chinese Gigaword）中的中國大陸地區新華社（XIN）和中國台灣地區通訊社（CNA）的新聞報道（下文分別簡稱「大陸新聞」和「台灣新聞」），選擇常用詞「提高」進行分析。通過考察該詞在主謂關係和述賓關係中的搭配詞及其語義類，分析兩岸新聞中「提高」的共有詞和獨有詞的特徵，並探究二十世紀 90 年代初至二十一世紀初期兩岸新聞對於「提高」相關現象的關注點異同。此研究不僅有助於豐富華語詞彙的研究成果，還能夠為兩岸的溝通與理解提供新的視角。

## 二、兩岸新聞語料庫

本研究採用新聞語料，這是因為新聞具有真實性，能夠反映社會的熱點問題；具有廣泛的傳播性，覆蓋各類人群和地區；具有多樣性，涵蓋政治、經濟、文化、科技等各個領域，能夠全面呈現社會的各方面，從而使其成為研究語言所反映的社會現象的重要數據來源。中文十億詞語料庫（Chinese Gigaword Corpus）彙集了來自大陸、台灣和新加坡的新聞文本。該語料庫的第一版於 2005 年發佈<sup>1</sup>，之後在 2009 年推出了包含分詞和詞性標注的第二版（Tagged Chinese Gigaword Corpus）<sup>2</sup>。本文使用了該語料庫中的兩個部分：台灣通訊社（CNA）的新聞和大陸新華社（XIN）的新聞。

詞彙特性速描（Word Sketch）是一個基於大型語料庫的系統，用於描述特定詞彙的語法關係及其搭配。該系統能夠顯示搭配詞的出現頻率和顯著性等特徵（Kilgarriff et al., 2005）。它的中文語料版本，中文詞彙特性速描系統（Chinese Word Sketch Engine，簡稱 CWS）也具有相同的功能（Huang et al., 2005）。

本研究在 CWS 系統中選擇了 Tagged Chinese Gigaword 的 XIN 和 CNA 語料，設置了以下檢索條件：最小頻率設定為「5」，語法關係的最大範圍設置為系統默認的上限「999」，最小顯著性設定為「0.00」。隨後，在這兩個子語料庫中搜索「提高」一詞，並從生成的語法關係列表中選取其常見的搭配關係，即主謂關係和述賓關係。通過人工篩選，提取出正確的搭配，數量如表 1 所示。(1) 主謂關係：搭配組合形式為「~ 提高」，搭配詞「~」作主語（subject），在 XIN 的語料中共有 416 個詞種，篩選出正確搭配詞種 360 個；在 CNA 的語料中共有 401 個詞種，篩選出 369 個正確搭配。(2) 述賓關係：搭配組合形式「提高 ~」，搭配詞「~」作賓語（object），在 XIN 的語料中共有 999 個詞種，篩選出正確搭配詞種 454 個；在 CNA 的語料中共有 999 個詞種，篩選出 641 個正確搭配。總體而言，CNA 語料中正確搭配的

---

1 <https://catalog ldc.upenn.edu/LDC2003T09>

2 <https://catalog ldc.upenn.edu/LDC2005T14>

詞種數量多於 XIN，作主語的多 9 個，作賓語的多 187 個。

表 1：「提高」在 XIN 和 CNA 中搭配詞的詞種數量

分類	XIN 主謂關係		XIN 述賓關係		CNA 主謂關係		CNA 述賓關係	
	原始語料	正確語料	原始語料	正確語料	原始語料	正確語料	原始語料	正確語料
詞種數量	416	360	999	454	401	369	999	641

以圖 1 中 XIN 語料中的述賓關係錯誤搭配「提高全民」為例，正確的搭配應為「觀念」、「水平」、「意識」、「素質」、「認識」等，但 CWS 錯誤地將這些「提高」的對象的所有者「全民」提取為搭配對象。由於語料庫規模龐大，難以對每個搭配錯誤進行人工修改，因此刪除了 CWS 錯誤提取的語料，重點分析在正確搭配的語料中「提高」的主語和賓語的情況。

圖 1：「提高」的錯誤搭配示例

0123	：加強教育，嚴肅法紀，	提高	全民的土地觀念；把開發新
0008	預防各種腸道寄生蟲病知識，	提高	全民的健康水平。</p><p>安樂
0026	說，搞好植樹造林關鍵要	提高	全民綠化意識新華社北京 3 月
0026	群眾對綠化國土的認識，	提高	全民的綠化意識，動員全國
0020	的法律意識和法制觀念，	提高	全民的法律素質。因此，
0021	《鐵路法》的宣傳學習，	提高	全民對《鐵路法》的認識
0021	鐵路法》的宣傳、學習，	提高	全民對實施《鐵路法》重要
0107	將舉行各種活動，進一步	提高	全民、全社會「居安思危
0012	違法犯罪分子作鬥爭，努力	提高	全民的文化素質，創造一
0032	人類共同的利益，一定要	提高	全民的「環境意識」，加強
0068	人類的共同利益，一定要	提高	全民的「環境意識」，使
0086	眾多的人參加體育鍛煉，	提高	全民的身體素質。因此，
0113	為了解放科技生產力，必須	提高	全民的科技意識，要在輿論
0130	為了解放科技生產力，必須	提高	全民的科技意識，要在輿論
0115	的熱潮。社論的發表，對	提高	全民的文化素質和語文水平
0014	。他們提出，當務之急是要	提高	全民的科技意識，使這一
0038	深刻的變革，這裡關鍵要	提高	全民的科技意識，特別是

0074	通知說，加強國防教育，	提高	全民的國防觀念，是我們
0059	群眾性體育活動的蓬勃發展，	提高	全民的體育參與意識，使
0024	生產力」的學習和宣傳，努力	提高	全黨全民的科技意識，這

然後，對 XIN 和 CNA 中與「提高」搭配的主語和賓語分別進行對比，將其歸類為共有詞和獨有詞，並依據《現代漢語分類詞典》（蘇新春，2013）的分類體系進行人工語義標注。該詞典收錄了 8 萬多條現代漢語詞彙，按照意義進行分類，是語義標注的重要參考資源。該詞典採用五級語義層次分類體系：第一至第三層為抽象的類別概念，第四層為抽象與具體之間的過渡層，第五層為具體的詞彙概念，各層級之間在語義上呈現上下位關係。詞典中共有 9 個一級語義類，為抽象事物、性質與狀態、具體物、時空、生物、生物活動、社會活動、輔助詞和運動變化。二級語義類多達 60 餘個。本文對「提高」一詞的搭配詞進行了一級和二級語義類的標注。對於未在該詞典中出現的詞，參考其近義詞在《現代漢語分類詞典》中的語義類別進行標注；對於那些在詞典中出現但其語義類別與詞典不同的詞彙，則根據實際語料情況對其語義類別進行調整。此外，本文還進一步從搭配詞的頻率和顯著性等角度，分析了兩岸詞彙差異及其背後的原因。

### 三、大陸與台灣新聞中「提高」的共有詞

如果與「提高」搭配的詞語同時出現在大陸和台灣的新聞中，這些詞語就是共有詞。對共有詞的分析有助於理解海峽兩岸民衆在「提高」這一問題上的共同關注點。統計結果顯示，在共有詞中，有 159 個詞語作為主語出現，而有 272 個詞語作為賓語出現。

#### （一）主謂關係中的共有詞

在大陸和台灣新聞中，共有 159 個詞語在主謂關係下作為「提高」的主語。根據語義類標注，這些詞分佈在《現代漢語分類詞典》的七個一級語義類中，分別是抽象事物、生物、具體物、社會活動、

生物活動、時空、運動與變化，分佈情況見表 2。這些詞語的語義類別分佈廣泛，其中以抽象事物類詞彙為主，數量為 134 個，佔總數的 84.28%，例如「產值」、「產能」、「產量」、「價值」、「價格」等詞語，這表明新聞報道中對經濟指標或其他抽象概念的關注較多。生物類詞彙次之，有 17 個，佔比 10.69%，如「人員」、「公民」、「農民」、「員工」等，反映出人們對某些現象的關注。表示社會活動、生物活動、時空和運動變化的詞語各只有 1 個。

表 2：主謂關係中大陸與台灣新聞共有詞的一級語義類分佈

一級類	數量	比例	例子
抽象事物	134	84.28%	產值、產能、產量、產銷率、價值、價格
生物	17	10.69%	人員、公民、農民、員工、大家、婦女
具體物	4	2.52%	產品、它、它們、設備
社會活動	1	0.63%	競爭
生物活動	1	0.63%	預算
時空	1	0.63%	年齡
運動與變化	1	0.63%	增速
總量	159	100.00%	/

本文進一步對佔比最高的抽象事物類共有詞進行了二級語義類標注，分佈情況見表 3。這些詞語在二級語義類中涵蓋了數量、屬性、經濟等 9 個類別。其中，數量類詞彙佔了最大比例，達 36.3%，如「產量」、「利率」等，顯示了新聞中對提高量化指標的關注。屬性類詞彙佔 23.7%，如「滿意度」、「可靠性」等，表明對特徵的描述也是重要內容。經濟類詞彙佔 15.56%，如「產值」、「價格」等，反映了對經濟主題的關注。事情類詞彙佔 7.41%，包括「業績」、「效能」等，政治類詞彙佔 5.93%，如「制度」、「政府」等。此外，意識類詞彙佔 4.44%，如「信心」、「思想」等，反映了對思想意識的重視。社會、軍事和科教類詞彙的比例相對較低，分別為 3.7%、1.48% 和 1.48%。

表 3：抽象事物類主謂關係中大陸與台灣新聞共有詞的二級語義類分佈

二級類	數量	比例	例子
數量	49	36.3%	產量、產銷率、利率、使用率、升學率、合格率
屬性	32	23.7%	滿意度、產能、依存度、可能性、可靠性、聲望
經濟	21	15.56%	產值、價值、價格、保險費、關稅
事情	10	7.41%	業績、功效、功能、境界、效能
政治	8	5.93%	制度、國家、政府、機關、機構
意識	6	4.44%	信心、興趣、思想、意識
社會	5	3.7%	企業、單位、國有企業、地位
軍事	2	1.48%	軍隊、部隊
科教	2	1.48%	面積、草案
總量	135	100%	/

儘管海峽兩岸在許多現象上存在共同關注，但這些詞語的顯著性存在差異。表 4 和表 5 分別展示了顯著性 (Kilgarriff et al., 2014) 差異最大的前 10 個詞。

表 4：主謂關係中大陸新聞顯著性高於台灣新聞的詞語

詞語	XIN 主語頻率	XIN 主語顯著性	CNA 主語頻率	CNA 主語顯著性	顯著性差異
水平	3617	56.86	82	28.5	28.36
質量	1794	45.64	24	20.31	25.33
生產率	193	42.5	9	20	22.5
效益	1198	42.54	48	20.39	22.15
檔次	164	37.39	6	15.66	21.73
能力	1201	41.73	119	20.82	20.91
合格率	82	29.4	5	11.88	17.52
覆蓋率	108	30.05	5	13.15	16.9
效率	545	46.39	125	29.62	16.77
成績	261	21.72	12	5.02	16.7

表 5：主謂關係中台灣新聞顯著性高於大陸新聞的詞語

詞語	XIN 主語頻率	XIN 主語顯著性	CNA 主語頻率	CNA 主語頻率	顯著性差異
民衆	7	2.96	469	27.19	24.23
成本	96	20.28	332	38.96	18.68
門檻	5	7.3	49	25.84	18.54
可能性	5	1.44	59	19.37	17.93
生產力	39	13.97	78	31.67	17.7
水準	33	19.28	297	36.97	17.69
金額	10	4.42	102	20.8	16.38
獎金	7	3.71	55	19.99	16.28
居民	14	0.06	47	13.01	12.95
重要性	8	3.88	30	15.63	11.75

通過對比可以看出：(1) 大陸新聞對「水平」、「質量」和「生產率」等方面的提高關注度遠高於台灣新聞。由於大陸在這一時期大力推進改革開放，各領域的發展水平、質量和效率顯著提高。語料庫中「水平」一詞與「生活」、「技術」、「科研」、「產品」、「教育」、「整體」、「農業」、「裝備」等詞共現，覆蓋了從農業到工業，從人民生活到科技、教育、軍事等廣泛領域，反映了這一時期大陸各方面的發展水平普遍甚至顯著提高。「質量」在語料庫中常與「產品」、「生活」、「生產」、「工程」、「教育」等詞共現，這反映出在一時期大陸在發展市場經濟的同時，重視民生發展，人民生活質量、社會生產質量以及教育質量等方面均得到了顯著提升。「生產率」則與「工人」、「勞動」、「土地」、「農業」等詞彙共現，不僅反映了大陸各行各業員工勞動生產率的提高和農業生產率的顯著增長，還涉及到有的國家或地區通過提高企業員工勞動生產率來降低勞動成本的情況。

(2) 台灣新聞對「民衆」和「成本」提高的關注度高於大陸新聞。「民衆」一詞在語料庫中常以「民衆應提高警覺」的形式出現，如提醒台灣民衆增強防火、防詐騙和地震等方面的安全意識。這種高度關注源於台灣的地理和社會環境特點。首先，台灣位於地震帶，地震頻發，提升防震意識對民衆安全至關重要。其次，由於台灣植被繁茂，

春季踏青燒紙活動頻繁，尤其在乾燥季節，火災風險顯著增加，故提升防火警覺成爲新聞熱點。此外，台灣社會還面臨詐騙集團欺詐的問題，進一步凸顯了提高防範意識的重要性。「成本」一詞與經濟政策緊密相關。例如，隨着大陸升息，台商的「貸款成本」增加，這反映了台灣經濟與大陸和美國經濟之間的緊密聯繫。同時，新聞還關注了台灣本地生產和生活成本的上升問題。例如，農產品外銷成本增加，導致台灣在國際市場上的競爭力下降；爲改善水質，供水成本提高；由於紙張價格上漲，導致台灣小學教材成本增加等。這些報道不僅揭示了台灣經濟在全球化背景下的複雜互動，還反映了島內多方面成本上升帶來的經濟和社會壓力，顯示出台灣新聞在這一時期對成本問題的特別關注。

## （二）述賓關係中的共有詞

在述賓關係中，大陸和台灣新聞共有的「提高」的賓語共有 272 個，它們的語義類分佈廣泛，涵蓋了抽象事物、具體物、生物活動、時空、社會活動、生物、性質與狀態這七個語義類，分佈情況詳見表 6。作「提高」賓語的詞彙，抽象事物類佔據了絕對主導地位，共有 257 個詞，佔總數的 94.49%。這些詞如「個人所得稅」、「產出率」、「產能」、「產量」和「產銷率」等，主要涉及經濟、生產和稅收等與社會發展密切相關的抽象概念。這表明，兩岸新聞的共同關注點更多地集中在涉及經濟指標、生產效率等宏觀領域的提高上。相比之下，具體物、生物活動和時空類的詞僅分別佔 1.1%。具體物類詞彙如「嗓門」、「聲音」和「設備」表明在某些具體的物理對象或設備上都有關注。生物活動類詞彙如「品味」、「警惕」和「預測」則反映了對個人行爲等的關注。社會活動、生物以及性質與狀態類的詞彙分別僅佔 0.74%，如「單產」、「評級」、「師資」、「交通安全」等。整體來看，這些數據反映了兩岸新聞在語言使用中的共同關注點主要集中在抽象的經濟、生產等宏觀議題上，而對於具體物理對象、個人活動以及社會現象的關注相對較少。

表 6：述賓關係中大陸與台灣新聞共有詞的一級語義類分佈

一級類	數量	比例	例子
抽象事物	257	94.49%	個人所得稅、產出率、產能、產量、產銷率
具體物	3	1.1%	嗓門、聲音、設備
生物活動	3	1.1%	品味、警惕、預測
時空	3	1.1%	年齡、年限、起點
社會活動	2	0.74%	單產、評級
生物	2	0.74%	師資、自己
性質與狀態	2	0.74%	交通安全、尊嚴
總量	272	100%	/

表 7 展示了述賓關係共有詞中抽象事物類詞彙的二級語義類分佈情況。共有詞主要集中在兩個類別：屬性類和數量類。屬性類詞彙是最主要的類別，佔比最高，共 88 個詞，佔比 34.24%。這些詞彙如「產能」、「體能」、「信任度」、「信譽」和「免疫力」等，反映了兩岸新聞在報道「提高」時關注內在屬性或質量的提升。例如，「免疫力」作為大陸和台灣新聞共同關注的「提高」對象，反映了兩地對提升民眾身體素質的重視。數量類詞彙緊隨其後，有 83 個詞，佔比 32.3%。例如「產出率」、「入學率」、「儲蓄率」、「亮度」和「份額」等，這些詞彙集中在指標的提高上，尤其是在經濟、教育和社會發展領域。比如，兩地都非常關注「入學率」，反映出兩地對教育發展的重視，表明在教育普及和提升教育質量方面，大陸和台灣都投入了大量資源和精力。其他類別如經濟類、事情類、意識類、社會類和科教類的詞彙佔比相對較低。這些類別雖然在「提高」這一述賓關係中的分佈較少，但仍然反映出兩岸新聞對某些具體領域的次要關注點。

表 7：述賓關係中大陸與台灣新聞共有詞抽象事物的二級語義類分佈

二級類	數量	比例	例子
屬性	88	34.24%	產能、體能、信任度、信譽、免疫力
數量	83	32.3%	產出率、入學率、儲蓄率、亮度、份額
經濟	50	19.46%	個人所得稅、價值、價格、保險費、關稅

事情	19	7.39%	作用、功效、功能、境界
意識	9	3.5%	信心、興趣、憂患意識、意識
社會	5	1.95%	地位、士氣、學歷、第二產業、等級
科教	3	1.17%	聲調、知識、百分點
總量	257	100%	/

儘管海峽兩岸新聞在許多現象上存在共同關注，但對這些現象關注的顯著性存在差異。表 8 和表 9 展示了顯著性差異最大的 10 個詞。通過對比分析，可以發現以下幾點：(1) 大陸新聞對提高「產出率」、「素養」和「單產」的關注度遠高於台灣新聞。「產出率」常與「土地」、「資源」等詞共現，反映了大陸對農業生產和資源利用問題的重視。提高產出率不僅是優化農業生產的關鍵，更是應對資源有限性的重要措施。「素養」一詞常與「思想理論」、「科學文化」、「道德」、「政治」、「文明」、「專業」等詞彙搭配，反映出大陸在這一時期對提升全民族思想道德素養和科學文化素養的重視。對素養提高的關注，既是為了提高國民整體素質，也是為了適應現代化建設的需要。「單產」是「單位面積產量」的簡稱，是衡量農業生產水平的重要指標。大陸新聞對「單產」的顯著關注，主要源於其對農業可持續發展的重視，以及對「三農」問題（農業、農村、農民）的持續重視。此外，大陸在這一時期加大了高科技在農業生產中的應用，農業機械化程度不斷提高。提高單產不僅是滿足糧食需求的關鍵措施，也是推動農業現代化和可持續發展的必然要求。這種關注反映了大陸在農業領域的戰略佈局，力求通過科技進步和生產效率的提升，確保農業的長遠發展和糧食安全。(2) 台灣新聞對於「投票率」、「生產力」、「收益」和「額度」的關注度明顯高於大陸新聞，這反映出台灣新聞在政治生活和經濟發展方面的關注。提高「收益」的重點多集中於「漁民收益」和「農民收益」，這與台灣的地理特徵息息相關。作為海島，台灣四面環海，漁業成為當地漁民的重要收入來源，對於「額度」的關注，主要體現在「貸款額度」、「信用額度」和「保險額度」等金融領域。此外，還涉及到「罰金額度」、「獎金額度」等其他方面。這種情況可能與台灣在

90年代初期金融業的開放有關，隨着商業銀行、投資信托公司、財產保險公司的相繼成立，台灣的金融業得以迅速發展。

表 8：述賓關係中大陸新聞中顯著性高於台灣新聞的詞語

詞語	XIN 賓語頻率	XIN 賓語顯著性	CNA 賓語頻率	CNA 賓語顯著性	顯著性差異
產出率	101	37.67	5	16.81	20.86
素養	257	38.88	47	18.78	20.1
單產	433	44.92	14	25.3	19.62
利用率	817	52.71	99	34.1	18.61
含量	979	39.67	77	22.02	17.65
修養	138	29.76	11	12.45	17.31
折舊率	42	33.12	7	18.05	15.07
自覺性	228	33.34	10	18.42	14.92
百分點	3522	67.38	986	52.72	14.66
自信心	55	23.01	5	16.81	14.28

表 9：述賓關係中台灣新聞中顯著性高於大陸新聞的詞語

詞語	XIN 賓語頻率	XIN 賓語顯著性	CNA 賓語頻率	CNA 賓語顯著性	顯著性差異
投票率	6	7.42	145	34.54	27.12
生產力	315	25.71	667	52.43	26.72
收益	40	9.53	246	35.01	25.48
額度	22	10.07	272	32.88	22.81
破案率	21	19.86	93	42.33	22.47
產能	11	9.87	159	31.04	21.17
憂患意識	14	12.63	48	32.33	19.7
費率	5	5.61	147	25.22	19.61
稅率	61	17.1	310	36.35	19.25
效能	191	23.94	319	42.79	18.85

#### 四、大陸與台灣新聞的獨有詞

獨有詞是指僅出現在大陸或台灣新聞中的詞語，這些詞語能直

接反映兩岸對「提高」現象的不同關注點，有助於更直觀地探究兩地新聞對經濟、政治、社會等方面的報道的差異。與 CNA 相比，XIN 在主謂關係中的獨有詞共有 201 個，在述賓關係中的獨有詞共有 182 個；而與 XIN 相比，CNA 在主謂關係中的獨有詞有 210 個，述賓關係中的獨有詞有 369 個。本部分不區分這些獨有詞出現在何種語法關係下。大陸新聞獨有詞分佈在 7 個語義類中，主要集中在抽象事物類，佔比超過八成。詞語如「鄉鎮企業」、「工業」、「主動性」等，通常與「提高」搭配使用，反映了大陸在推動經濟結構調整、工業化進程以及增強主動性等方面的努力。生物類詞彙，如「公眾」、「員工」、「自我」、「自身」、「隊伍」等，也常與「提高」搭配，反映了對集體和個人能力提升的重視。其他類別，具體物，如「中醫藥」、「有機質」、「煤礦」；社會活動，如「宣傳」、「裝備」、「訓練」；生物活動，如「用水」、「實驗」、「應用」；性質與狀態，如「核安全」、「食品安全」、「熱情」；以及時空，如「蘇南」、「地方」等，雖然數量較少，但這些詞彙在與「提高」搭配時，往往指向特定領域的提高。

台灣新聞的獨有詞也分佈在多個語義類別中，其中絕大多數屬於抽象事物類，佔比超過八成。這些詞彙包括「所得」、「薪資」、「能見度」、「保費」、「耐震度」、「高度」等，與「提高」搭配時，通常反映了台灣新聞社會對經濟收入、生活質量、環境標準等方面的關注。在生物類詞彙中，如「僑民」、「台商」，反映了台灣新聞對其海外華人社區和經濟參與者的關注。其他語義類別中的詞彙，如社會活動類的「保障」、「獲利」、「獎勵」，強調社會經濟活動的提升；具體物類詞彙，如「債券」、「抗體」、「籌碼」，展示了在金融投資和健康防護領域的需求；生物活動類詞彙，如「用水」、「品味」、「實驗」、「應用」，涉及生活質量和科學研究；時空類詞彙，如「刑期」、「年資」、「目標區」、「空間」，反映了對司法、職業發展和空間規劃的關注。總之，台灣新聞的獨有詞在與「提高」搭配使用時，集中體現了其在經濟發展、社會保障、科技進步以及環境保護等方面的關注。

本文進一步對大陸和台灣新聞獨有詞中抽象事物類詞彙的二級語

義類分佈情況進行了分析，以深入探討兩岸獨有詞的差異。在大陸新聞中，抽象事物類詞彙的二級語義類分佈在屬性、數量、經濟等 10 個類別中。其中，屬性類的獨有詞最為突出，佔比約為三分之一，如「總水平」、「自由度」、「自覺性」、「準確性」等，反映了大陸社會對質量、標準和行為規範的高度重視。數量類詞彙，如「優質率」、「充足率」、「兌現率」、「入戶率」、「出航率」，展示了對具體統計指標的關注，尤其是在工業生產和資源管理領域。經濟類詞彙，如「助學金」、「造價」、「贊助費」、「烟價」，則體現了對經濟支持、成本控制和市場價格的重視。此外，大陸的抽象事物詞彙還涵蓋了其他類別，包括事情，如「做法」，反映了對行為的關注；社會，如「鄉鎮企業」、「工業」、「支柱產業」，顯示了對基層經濟和產業結構的重視；意識，如「使命感」、「安全感」、「情趣」、「緊迫感」、「自尊心」，強調了心理狀態的重要性；科教，如「教育」、「物質文明」、「科學知識」，突出了對教育和科學發展的關注；政治，如「黨委」、「北京市」、「縣委」、「成員國」、「產油國」、「沙特」，說明了對政治結構和國際關係的關注；文體，如「公開賽」、「女排」、「賽事」反映了對體育文化的關注；軍事，如「警力」反映了對安全力量的提高。

相比之下，台灣新聞的抽象事物類獨有詞主要集中在數量（如「租屋率」、「自製率」、「資本額」、「濁度」）和經濟（如「買價」、「事務費」、「產值」、「代金」、「價碼」、「傭金」）兩個類別，各佔約三分之一。這反映了對經濟活動、市場運作以及財務管理方面的重視。台灣新聞報道中獨有詞的其他語義類別還包括屬性，如「上限」、「下限」、「便利性」，表明對市場動態和服務質量的關注；政治，如「公權力」、「刑罰」、「女權」，反映了對權利和法律的重視；事情，如「所得」、「業績」、「代價」、「優勢」、「措施」、「效力」，顯示了對成果、成本和競爭力的關注；社會，如「位階」、「保險」、「聲勢」、「官等」、「職等」，反映了對社會階層和職業發展的重視；意識，如「樂趣」、「參與感」、「憂慮」、「性欲」、「警覺心」，強調了個體心理和情感狀態的重要性；科教，如「華文」、「語調」、「音調」、「理論」、「網絡」、

「議題」，體現了對語言、教育和科技的重視；速率，如「建廠率」、「年增率」，展示了對經濟增長和產業發展的監控。

## 五、結論

「華語」是全世界華人的共同語，其使用在不同的地區存在一定的差異。而這些差異往往可以通過不同地區的新聞報道中的語言使用得到反映。新聞以其真實、廣泛和多樣的特點，成為研究這些差異所反映的社會現象的重要信息來源。本文基於中文十億詞語料庫（Tagged Chinese Gigaword）中大陸和台灣的新聞語料庫，對「提高」一詞在主謂關係和述賓關係中的搭配詞進行了對比分析，探討二十世紀 90 年代初至二十一世紀初兩岸關注的「提高」議題，從而了解這一段時間兩岸的兩家新聞媒體關注的社會現象。

共有詞研究有助於了解海峽兩岸新聞共同關注的提高現象，經分析發現：(1) 在主謂關係方面，共有 159 個詞語在主謂關係中作為「提高」的主語，這些詞語主要集中在抽象事物類，佔 84.28%，如「產值」、「產能」等，反映了對經濟指標的關注；生物類詞彙佔 10.69%，其他類別較少。佔比最高的抽象事物類詞彙中，數量、屬性和經濟類詞彙佔主導，顯示了新聞對量化指標和經濟主題的重視。此外，在共同關注的現象中，顯著性存在差異，大陸新聞更關注「水平」、「質量」和「生產率」等的提升，反映了改革開放期間各領域的發展；而台灣新聞更關注「民衆」警覺性和「成本」等的提高，反映出其對安全意識和經濟成本的重視，這與台灣的地理環境和經濟背景密切相關。(2) 在述賓關係方面，共有 272 個詞語作為「提高」的賓語，這些詞語主要集中在抽象事物類，佔 94.49%，如「個人所得稅」、「產出率」等，反映了對經濟、生產和稅收等宏觀領域的關注。其他類別如具體物、生物活動和時空類詞彙佔比很少。佔比最高的抽象事物類詞彙中，屬性和數量類詞語佔主導，顯示了新聞對內在屬性提升和量化指標改進的重視。此外，在共同關注的現象中，顯著性存在差異。大陸新聞更關注「產出率」、「素養」和「單產」等的提高，反映了其在農業生

產、思想道德教育和科技應用等領域的發展重點。而台灣新聞則更關注「投票率」、「生產力」和「收益」等的提高，體現了對政治參與、經濟發展和金融領域的重視，這與台灣的地理環境和經濟密切相關。

此外，對獨有詞的研究進一步揭示了兩岸在「提高」現象上的不同關注點。大陸新聞的獨有詞主要集中在抽象事物類，佔比超過八成，反映了對經濟結構調整、工業化和個體素質提升的重視，如「鄉鎮企業」、「工業」、「主動性」等。相比之下，台灣新聞的獨有詞也多集中在抽象事物類，但主要關注經濟收入、生活質量和環境標準，如「所得」、「薪資」、「能見度」等。進一步分析顯示，大陸新聞的獨有詞在二級語義類中以屬性、數量和經濟為主，反映了對質量標準、統計指標和經濟支持的重視；而台灣新聞的獨有詞則集中在數量和經濟類，突出對經濟活動、市場運作和財務管理的提高的關注。兩地新聞在政治、社會、意識、科教等領域也表現出不同的側重點，分別體現了各自社會背景下的關切和發展方向。

隨着華語地區和全球範圍內的政治、經濟、文化交流日益頻繁，華語詞彙體系也在不斷發展。這種動態變化為詞彙研究提供了豐富的素材。深入分析和比較不同地區的華語詞彙，可以更全面地了解大華語的語言面貌，有助於推動華語研究的發展，進一步增進各地區的相互了解與互動。

## 參考文獻

- 鄧思穎 (2018)：全球華語詞彙研究的參數分析——以「點算」為例，《漢語學報》，4，2-10+95。
- 郭熙 (2012)：華語視角下的「講話」類詞群考察，《語言文字應用》，4，2-9。
- 李宇明 (2010)：《全球華語詞典》，北京：商務印書館。
- 李宇明 (2016)：《全球華語大詞典》，北京：商務印書館。
- 陸儉明 (2019)：樹立並確認「大華語」概念，《世界華文教學》，1，3-5。
- 邵敬敏、石定栩 (2006)：「港式中文」與語言變體，《華東師範大學學報（哲學社會科學版）》，2，84-90。
- 蘇新春 (2013)：《現代漢語分類詞典》，北京：商務印書館。
- 湯志祥 (1995)：中國大陸、台灣、香港、新加坡漢語詞彙方面若干差異舉例，《徐州師範學院學報》，01，103-109。
- 田小琳 (2004)：香港社區詞研究，《語言科學》，3，29-37。
- 周清海 (2016)：「大華語」的研究和發展趨勢，《漢語學報》，1，13-19+95。
- 周清海 (2017)：「大華語」與語言研究，《漢語學報》，2，61-66+96。
- Huang, C.-R., Kilgarriff, A., Wu, Y., Chiu, C.-M., Smith, S., Rychly, P., Bai, M.-H., & Chen, K.-J. (2005). Chinese Sketch Engine and the Extraction of Grammatical Collocations. *In Proceedings of the Fourth SIGHAN Workshop on Chinese Language Processing* (pp. 48-55). Jeju Island, Korea.
- Kilgarriff, A., Baisa, V., Bušta, J., Jakubíček, M., Kovář, V., Michelfeit, J., Rychlý, P., & Suchomel, V. (2014). The Sketch Engine: ten years on. *Lexicography*, 1(1), 7-36.
- Kilgarriff, A., & Huang, C.-R., & Rychlý, P., & Smith, S., & Tugwell, D. (2005). Chinese Word Sketches. *ASIALEX 2005: Words in Asian Cultural Context*. Singapore.

# A Comparative Analysis of Vocabulary in Cross-Strait News

WANG, Shan\*

## Abstract

The study of Chinese language is an important way to understand the changes and developments of Chinese society, and news, as an authentic, widely accessible, and diverse medium, provides essential data for examining the social phenomena reflected in language. This study utilized the Chinese Gigaword corpus, specifically news from the Chinese Mainland and Taiwan, to analyze the commonly used word 提高 *tígāo* ‘improve’, conducting a comparative analysis of its collocations in subject-verb and verb-object relationships. It examined the phenomenon of improvement as reported by representative news media on both sides of the Taiwan Strait from the early 1990s to the early 21st century. Studying common words helps understand the shared concerns of news agencies across the Strait, while the analysis of unique words reveals the differences in focus between the two sides more directly. Investigating these similarities and differences enhances our understanding of the use of Chinese in cross-Strait news reporting and provides insight into societal developments on both sides of the Taiwan Strait.

*Keywords:* vocabulary, global Chinese, newspaper, improve, the Chinese mainland, Taiwan

---

\* WANG, Shan, Department of Chinese Language and Literature, Faculty of Arts and Humanities, University of Macau, Macao SAR, China. Institute of Collaborative Innovation, University of Macau, Macao SAR, China.