

從 LIWC 到 C-LIWC：電腦化中文字詞分析的潛力

林瑋芳

國立臺灣大學
心理學系

黃金蘭**

國立臺灣科技大學
人文社會學科

林以正

國立臺灣大學
心理學系

情緒書寫已經被證實對身心健康產生有正向的效果。Pennebaker 研究室發展「語文探索與字詞計算」(簡稱 LIWC) 的電腦分析程式以探索情緒書寫產生正向效果的心理機制。隨著後續的相關研究發展與需求，LIWC 經幾次改版，目前的 LIWC2007 也已被證實是一項可以探討語言使用的心理特性的工具，具有穩定的信效度，也被廣泛應用在各研究領域中。本文整理 LIWC 與中文版 LIWC (C-LIWC) 的發展及相關研究發現與應用。儘管 LIWC 是一項以字詞為單位的量化分析工具，研究者在使用時，仍然需在理論導向在特定主題脈絡的引導下進行分析與闡釋，絕非完全去脈絡化的變項套用。C-LIWC 發展完成，應可作為一項研究華人心理特性的研究利器，成為質化與量化研究間的溝通橋樑。

關鍵詞：字詞分析、文本分析、LIWC

緒論

一個清晰的自我往往是心理適應的重要基礎。可是，個體如何在每天的活動中，不斷的去理解自我，去建構自我，並且據以與世界互動，甚至提供一個成長的可能性？一百多年前，William James (1890) 在其心理學原理中就提出，人具有兩個基本的我，一個是客體我 (me)，另一個則為主體我 (I)。Hermans、Rijk 與 Kempen (1993) 延伸這個想法提出了「對話自我」(dialogical self)。他們主張，每一個「我」都是一個發聲者，自我的歷程則是在不斷對話中連結出意義的歷程。而要產生良好的對話，一個念頭、想法、或判斷就要能夠被具體的捕捉，進而成為發聲者，並且與其他面向的自我發聲互相的對應。由這些論述看來，個體的內在聲音，如果能夠形成一個良好的自我對話，就越能夠形成一個清晰的自我。

自我的對話具體的呈現方式之一就是書寫，其中，最為一般大眾所熟知的就是日記書寫。在日記書寫中一個人重新感受與回溯經驗，在與自我的對話中去與一個過去式的自我產生對話，連結這些不同時間、位置的自我之想法、心情、感受、甚至更後設性的省思自我。在這樣的自我對話中，過去事件被重整與再建構；被重新詮釋後產生了一個新的經驗，甚至產生一個新的自我面貌。有鑑於日記書寫似乎具有自我對話與自我整合的可能效果，心理學家 Pennebaker (1997) 推論，良好的日記書寫應當也可以有效改善個體的身心健康狀態，特別是當個體可以用文字去捕捉相當具有流動性的情緒經驗，得以將這些看似飄忽不定的情緒經驗具體的用文字來發聲，在表露與抒發之外，也可能提供反思，進而產生療癒的效果。Pennebaker 發展的情緒書寫派典 (Pennebaker's expressive writing paradigm, Pennebaker & Beall, 1986) 就發揮了這樣的功能。在他的研究中，參與者被隨機分成創傷事件組與控制組。書寫創傷事件者要回想一件自己日常生活中所發生過的真實創傷事件，並嘗試以最深刻的想法與感受來書寫該事件內容；控制組的參與者則是書寫中性生活事件。如此連續書寫四天，也在書寫六個月後追蹤其健康狀態。在他的研究中獲得了支持情緒書寫療癒效果的發現，同時也開展了情緒書寫的一系列研究。然而後續對於情緒書寫的研究卻無法有效的提出中介變項或調節變項，來對情緒書寫之

**本篇論文通訊作者：黃金蘭，通訊方式：clhuang@mail.ntust.edu.tw

正向效果的發生機制做更加清楚的解釋 (Burton & King, 2004; Smyth, 1998)。

為了釐清究竟書寫的過程發生了哪些心理歷程使得正向效果得以可能？或許書寫文字內容就提供了最直接的線索，同時也直接表徵了書寫者的內在聲音(金樹人, 2010)。然而文本是質性的資料，並不能直接進行量化的分析。因此，如何擷取其中的關鍵因素，並且建立能夠充分指涉心理歷程的指標，就形成一個相當嚴酷的挑戰。比較簡單直覺的方法是，請不同的評分者來閱讀書寫的內容，並且就研究者有興趣的向度進行評量。然而，這樣的過程不但耗費人力，而且往往評分者之間的一致性也常常偏低 (Pennebaker & Francis, 1996)，更有甚者，這些評量出的分數對於結果變項的預測力也不盡理想 (Graybeal, Sexton, & Pennebaker, 2002)。

面對此一挑戰，Pennebaker 及其同僚於 1990 年初開始著手發展得以進行語文特性分析的電腦工具。Pennebaker 在語文分析的策略上，決定回到書寫中最基本的單位，亦即字與詞的使用。或許是基於最簡單的事實就是最直接的證據的理念，他們早期發展之語文字詞分析程式也相當簡單，就是針對文本中所使用的詞類來計算其出現的頻率，諸如代名詞、情緒詞及認知詞等的相對使用百分比。他們將此工具命名為「語文探索與字詞計算」(Linguistic Inquiry and Word Count, 簡稱 LIWC, 發音如 look)。隨後歷經數次的修訂，從 LIWC (Francis & Pennebaker, 1993)、LIWC2001 (Pennebaker, Francis, & Booth, 2001)、至目前的 LIWC2007 (Pennebaker, Booth, & Francis, 2007; Pennebaker, Chung, Ireland, Gonzales, & Booth, 2007)，此一分析工具已經相當豐富而且具有良好的信效度。

語言探索與字詞計算工具的運作

LIWC 其實包含兩個主要的部分，其中之一是電腦軟體部分，負責比對分析與運算；另一個更為核心的部分則是詞典。在文本分析時，當抓取到一個詞之後，要判斷這個詞究竟具有怎樣的特性，就必須倚靠程式中的詞典來定義。例如：「鼓勵」，同時就具有社會詞、情緒詞、正向情緒詞、認知等性質；又如「懊惱」，也應同時隸屬情緒詞、負向情緒詞、悲傷詞、認知詞、差距詞等類別。這些類別的歸屬就必須倚靠其中的詞典來定義。LIWC 運作的基本原理是，每當打開一個文本檔案，程式就以空白字元為區隔單位來擷取每個詞，並將每個詞與詞典進行比對。若在詞典的某個類別下找到相對應的詞，就在該類別的頻率上累加次數。若同一個詞同時屬於不同的類別，那麼所有包含該詞的類別都會同時增加計數。依照這個方式逐詞的比對到文本中所有詞窮盡為止。最後以該文的總字數為分母，計算出各個類別累計頻次佔總詞數的百分比。

就英文版的 LIWC2007 來說，目前包含了 64 個定義類別。這些類別包括基本的語文向度類別，例如：22 個語文特性類別（如代名詞、冠詞等百分比）、3 個副語言學 (paralinguistic) 類別（如應和詞、停頓詞等）。除了已經具有明確定義的語文類別之外，LIWC 也另外包含了 32 個心理特性類別（如情感詞、認知詞等）、7 個個人化類別（如工作、休閒活動等）。再加上程式自動計算出的 6 個一般描述性類別（總詞數、每句詞數等）以及 12 個標點符號類別（如句點、逗點等），總計可輸出 80 個字詞類別。

值得特別一提的是，由於 LIWC 詞典建立的主要目的在於偵測語文使用者之心理特性，其目的並非要涵蓋所有日常生活中所可能使用的語彙，因此建立詞典的目的主要並不在於其涵蓋率。倘若就語言的特性大類來分，一般而言，可以將人們所使用的語言分為內容詞（或稱實詞）與功能詞（或稱虛詞）。內容詞通常會因為溝通雙方所談論的主題與情境有關，很可能主要包含的是動詞、名詞與形容詞，其變異性大，包含的範圍也太廣，並不見得能夠被充分的涵蓋。但是另一方面，所謂的功能詞（諸如代名詞、連接詞、助詞等），由於其本身並不具有內容的意涵，其功用往往是黏著內容詞來達到完整的表達。這些詞過去往往被研究者所忽略，然而近年來，Pennebaker 的研究團隊發現這些功能詞對於各種重要的指標也具有很穩定的預測力，甚至倘若要指標一個人的語言風格，這些被用來組織內容詞的功能詞，反而更具有表徵的效果 (Chung & Pennebaker, 2007; Ireland & Pennebaker, 2010; Tausczik & Pennebaker, 2010)。就目前 LIWC2007 的英文詞典

來看，其涵蓋的詞數目僅有約 4500 個詞，但是竟然就已經可以偵測到一般書寫文字的百分之八十五以上的內容，其中功能詞的使用比率大約在百分之五十左右 (Pennebaker, Chung et al., 2007)。這除了顯示 LIWC 有相當不錯的偵測率之外，也可看出功能詞在日常語文使用上所扮演的角色之重要性。

LIWC 的應用

由於語言是人類思考、表達、溝通都不可或缺的工具，因此也提供了 LIWC 在各方面得以應用的可能性。從此一工具被發展出來之後，目前已經累計了上百篇的相關研究 (整理見 Tausczik & Pennebaker, 2010)。LIWC 的發展初始關切的焦點在於情緒書寫如何透過具體的文字抒發，產生療癒的可能，主張透過文字的使用頻率反映內在的歷程變化。後續相關研究也將 LIWC 應用到不同的層面，探討人格特質、心理健康、教育學習乃至人際關係等不同的領域，顯示 LIWC 作為字詞分析的研究工具，具有極大的應用潛力，以下將就 LIWC 的應用層面做更詳細的說明與介紹。

語言如同一個人的認知、情緒、與態度，具有顯著的個別差異。而究竟 LIWC 所偵測的指標是否能夠反映這些個別差異則是考驗 LIWC 效度的第一個指標。在 LIWC 發展的早期，Pennebaker 與 King (1999) 就發現，外向性高者會使用較多的正向情緒詞，而神經質高者則傾向於使用較多的負向情緒詞。後續以經驗取樣的方式分析日常語言的使用所做的研究也支持大五人格 (big five personality) 與 LIWC 的各項指標具有穩定的相關 (Mehl, Gosling, & Pennebaker, 2006)。其他以不同的材料來源，也大都得到類似的結果。例如，Hirsh 與 Peterson (2009) 分析自我敘事的文本，或 Yarkoni (2010) 分析網路部落格的書寫者，都得到 LIWC 類別頻率與大五性格之間的相關。此外，在性別差異方面，女性使用較多的第一人稱代名詞與較多的認知詞，此一模式與低社經地位，或較低年齡者具有相同的模式，很可能顯示情境因素所影響的權力結構也會反映在語言的使用上 (Pennebaker, 2011)。換言之，LIWC 可以捕捉穩定的個別差異。

對心理健康等的預測力方面，Stirman 與 Pennebaker (2001) 發現自殺詩人傾向於在作品中使用較多的第一人稱單數代名詞與死亡詞。Handelman 與 Lester (2007) 進一步比較自殺身亡與自殺未遂者的筆記，卻發現相對於自殺未遂者，自殺身亡者反而有較多的正向情緒、較多的社會相關詞、以及較多的未來式。這很可能也顯示，儘管同樣具有自殺傾向，由於自殺未遂者比較會釋放出較多的相關訊息，因此得以獲得注意，但是自殺身亡者則很可能較未被注意到而產生悲劇。當然，此一結果除了顯示，LIWC 對於不同自殺狀況的人具有預測力之外，如何解釋這些差異，以及如何運用這些差異的偵測來提供更有效的協助與防止，則還有待更多的研究釐清。

除了一般可以很直覺的預測情緒詞的使用可以區辨是否具有憂鬱傾向之外，更有趣的是代名詞與憂鬱傾向也具有穩定的關連性。尤其第一人稱單數代名詞 (我、我的、自己等) 與負向情緒也具有穩定的關連性 (Chung & Pennebaker, 2007; Rude, Gortner, & Pennebaker, 2004)。例如，Rude 等人 (2004) 請大一的新生寫下他們進入大學的經驗，就發現具有憂鬱傾向於使用較多第一人稱單數代名詞，較多的負向情緒詞。而即便是已經從憂鬱復原但是先前曾經有過憂鬱病史的學生也比從來沒有憂鬱病史的學生使用較多的第一人稱代名詞。這些研究顯示，在語言使用中，一般研究者甚或是敘說者本身都不會覺察的一些線索，卻可能提供很重要的訊息，而 LIWC 則提供了一個窗口用以偵測這類的線索。

回到日記書寫的療效分析上，透過 LIWC 的字詞分析，Pennebaker (2011) 也發現具有療癒效果的情緒書寫傾向於具有下列的特色：在情緒詞方面，雖然是書寫負向事件但是負向情緒詞的使用在中等程度，同時也能夠有較多的正向情緒詞者有較好的療癒效果；在認知詞方面，當認知詞隨著書寫經驗而逐漸增加者，也會比整體認知詞使用較少者或一開始書寫就出現大量認知詞者會得到較佳的效果；在代名詞方面，雖然是書寫自己的經驗，但是倘若能夠在不同的書寫中具有較多代名詞使用的變異性者，也會有

較佳的療癒效果，這些代名詞的轉換很可能代表了書寫的觀點轉換。在字詞分析整體的效度上，Pennebaker 相關的研究也指出，透過字詞分析甚至可以比由諮商工作者更能有效的區辨不同情緒書寫文本的療癒效果。

此外，應用 LIWC 探討不同的心理疾患或身心適應問題的研究也相當的廣泛。例如，LIWC 曾被應用於區辨厭食症 (Anorexia Nervosa) 與貪食症 (Bulimia Nervosa) 兩種不同飲食疾患之患者，在語言使用上的差異 (Davies, Swan, Schmidt, & Tchanturia, 2012)。由於厭食症通常伴隨著情緒壓抑的特徵，Davies 等人的研究結果也相對應的發現，相較於一般正常人，厭食症患者有較少的書寫總字數和較少的正向情緒詞；反之，貪食症患者與正常人在語言使用特性上則沒有明顯差異。Holtgraves、McNamara、Cappaert 與 Durso (2010) 則是透過 LIWC 分析帕金森氏症的患者在語言使用基本特徵。帕金森氏症在初期會出現不對稱的半邊肢體於靜止時不停顫抖或動作困難等障礙，也會伴隨著一些語言缺陷和說話困難。Holtgraves 等人的研究結果發現，帕金森氏患者中左半邊運動困難的患者，由於其主司語用功能的右半腦受損較為嚴重，因此，患者會出現使用較少的動詞、功能詞及句子較短等的語言使用特徵。除此之外，相關研究也涉及以 LIWC 探討情緒疾患 (Vanheule, Meganck, & Desmet, 2011)、人格疾患 (Arntz, Hawke, Bamelis, Spinhoven, & Molendijk, 2012) 等，在在顯示 LIWC 在臨床的應用上深具潛力。

近來 LIWC 在教育相關領域的相關領域也有成長。例如：Robinson、Navea 與 Ickes (2013) 發現大學生在學期前所撰寫的自傳中的第一人稱代名詞、標點符號的運用、簡單字、以及其他指標等都對於學期成績具有預測力。他們推論，學生的自我焦點對於學業成就表現具有負向的影響力。顯示語言對於成績的表現還是有相當的預測力。

除了個人層次的分析之外，語言分析也可以擴展到關係的層次。例如，兩個人之間所使用的語言型態也可能具有一定程度的相似性，而此相似程度對於其關係也可能有所影響。為了考驗此一可能性，近年來，Pennebaker 及其同僚嘗試建立一個綜合性語言風格匹配度 (language style matching, 簡稱 LSM, Ireland & Pennebaker, 2010) 的指標。由於功能詞主要是擔負著黏著內容詞的角色，因此個人風格的展現最有可能是透過功能詞來展現。基於此一理念，他們計算 LIWC 中功能詞類別的使用一致性，計算出兩篇文本 (或對話) 在這些功能詞使用上的相似程度，並以此來指標語言風格的匹配度。在他們新近的研究中，以快速約會中的對話作為研究材料，即便是短短五分鐘左右的談話，由這麼短的對話中所計算出的匹配程度，就可以有效的預測雙方的好感度：若以交往中的情侶為研究對象，同樣的也可以預測情侶在三個月之後是否持續交往。此一指標甚至比主觀自陳的感受對於行為具有更好的預測力 (Ireland et al., 2011)。將此一指標運用在小團體溝通的相關研究時，也發現成員間的 LSM 與團體凝聚力及合作表現有顯著的關聯性 (Gonzales, Hancock, & Pennebaker, 2010)。

綜合上述相關文獻討論，顯示 LIWC 已經廣泛地應用於臨床、教育及親密關係等不同的領域，為開放式文本分析提供了一種有效的分析策略，亦即語言字詞的使用頻率本身，即可成為一個穩定且富有意義的預測指標。然而，LIWC 發展之濫觴是以英文為主要語言，上述相關研究也多來自英文文本，而中文作為全球使用人口最多的一種語言，是否也存在著以 LIWC 進行分析的可能性呢？

中文版 LIWC 的建立與初步研究成果

由於 LIWC 這套軟體具備了因應研究者需求自建詞典的特性，使得各國語文得以透過自行建立不同語言詞典的方式來使用 LIWC 程式進行文本分析。目前已經有德文、西班牙文、韓文、義大利文等多國語文版本的 LIWC 詞典建立完成。中文版的 LIWC 詞典 (簡稱 C-LIWC) 也在 Pennebaker 教授的授權下，於 2012 建立並正式發表 (黃金蘭等人, 2012)。C-LIWC 是以 LIWC2007 詞典作為藍本，先刪去不適用於中文特性的類別 (例如，冠詞類、各種動詞時態類別)、對保留下來的所有類別詞逐一進行翻譯及同義詞的增添；之後再經過多次的 3~6 人的研究小組討論，逐字確認類別；同時參考國內各

詞庫與語詞分類系統等，經由小組討論共同決定加入一些中文特有類別(如數量單位詞、語尾助詞、時態指稱詞等)，加入字詞及逐字確認其所屬類別；最後再進行斷詞確認、類別從屬關係確認以及最後的整體再確認等步驟。研究團隊歷時半年多經由上述步驟，以及信效度檢驗，完成了中文版 LIWC 詞典，其中包含了 29 個語文、32 個心理、7 個個人化等類別。目前 C-LIWC 對於一般的書寫文本大約已有八成以上的偵測率，有相當不錯的表現。

中文版 LIWC 的建立，提供華語文使用者心理特性研究的一項重要工具，提供了一個嶄新的研究取徑。發展至今，研究團隊除了重覆驗證一些國外研究結果，顯示 C-LIWC 同樣具有穩定的效度和預測力之外，也對 C-LIWC 的相關應用進行廣泛的探索。以下介紹研究團隊一些初步的發現與研究成果。

過去研究發現，經驗消費與物質消費不僅在消費策略上具有明顯差異，在提取這兩種不同型態的消費經驗時，也會產生不同的情緒與適應狀態。唐思柔(2011)透過字詞使用頻率的分析方式，直接比較經驗消費與物質消費在文字使用上的差異。她以受試者內設計，使用 C-LIWC 分析比較 91 位部落客就其美食或旅遊經驗以及購買某項 3C 產品所書寫的分享文各一篇。結果發現，相對於物質消費，經驗消費的書寫使用較多第一人稱複數代名詞、現在式時態標記、社會性字詞、休閒詞與生理相關字詞；物質消費則出現較多的第一人稱單數代名詞、因果詞、視覺、聽覺詞與工作詞。上述的書寫特性差異也相當程度呼應了過去對於此兩種消費類型帶來心理特性的差異(Carter & Gilovich, 2010; Van Boven, 2005)。

另一方面，以功能詞的相似性作為比較基礎的語言風格匹配性(LSM)的相關探索，也是我們檢驗 C-LIWC 效度的一個重要方向。竇郁文(2012)把 LSM 概念應用在廣告說服效果上。她編寫了一份部落客的商品特色介紹文，請研究參與者看完介紹文之後，書寫對該商品的看法，並且評量對該部落客的喜歡程度以及對該商品的購買意願。透過 LSM 的計算，結果發現研究參與者的書寫文本風格與部落客介紹文的書寫風格越相近，參與者越喜歡部落客，購買意願越高。進一步分析發現，這個喜好度中介了 LSM 對研究參與者購買意願的預測效果。換言之，竇郁文的研究發現 LSM 的效果是透過提昇研究參與者對部落客的喜好度來影響其購買意願，顯示使用 C-LIWC 的文本比對所計算得到的 LSM，確實反映了兩文本的風格相似性，可以有效預測對對方的喜好度，也呼應了過去文獻指出 LSM 能夠預測關係雙方好感度的研究結果。

然而，語言風格的相似性(LSM)未必是一個絕對正向的指標。張硯評(2012)探討感恩的操弄效果，他將參與研究的夫妻檔，隨機分成兩組，一組為感恩組，另一組則為抱怨組。夫妻雙方要在每週一和週四，互相寫 e-mail 給對方，持續三週。其中，感恩組書寫的內容是對於配偶過去三至四天的作為表達感謝；抱怨組的書寫內容則是告訴對方生活中的擾人之處。結果發現在憂鬱情緒的改變量上，實驗操作與 LSM 有顯著的交互作用。亦即，在相互感恩的情況下，LSM 越高，越有正面效果；然而，在抱怨訴說的情況下，LSM 越高反而有負面的影響。對照 LSM 應用在不同研究上的結果，顯示語言風格相似性的影響力不存在絕對的方向性，而是在不同的情境脈絡中呈現不同的效果。

C-LIWC 的研究團隊從 2010 年開始著手建構中文 C-LIWC 字典，到 2012 年正式發表，目前雖然索取此研究工具的人數相當多，但是大多數還無法在這麼短暫的時間內完成正式的研究報告。若就上述的幾個初期的研究結果來看，此工具確實可以在中文的環境中重複驗證其他語言的典型發現，然而一個工具的效度仍是需要更多與更長期資料的累積與辯證，這還有賴於相關領域的研究者來共同努力與發揮創意。作者們以為，雖然目前 C-LIWC 的效度考驗仍然不足，但是這是可以透過未來不斷的修正來改善，而且更重要的是，此一工具對於中文環境的研究者來說，提供了另一個研究的視角，除了可以與其他研究策略進行對話與比對，同時也可以開拓新的研究可能性。

自建詞典檔特性

C-LIWC 詞典檔（黃金蘭等人，2012）在編製的過程中，除了一些語文特性差異的增訂刪除之外，原則是複製了 LIWC2007 的英文版詞典。此一做法的優點是在比較中英文使用者的語言特性時，彼此是建立在共同的基礎之上的，此特性有助於心理特性的跨文化研究進展。但相對的，這也使得 C-LIWC 可能缺乏一些更符合華語文使用者心理特性的語詞類別。這一兩年來，我們的研究團隊除了致力於 C-LIWC 的各種信效度檢驗之外，更將文化對語言使用的影響納入考量，修訂符合華人文化、更能偵測華人心理特性的詞類，使其更能貼近本土文化意涵。最近甫完成雙向度情緒類別（黃金蘭，2013）。以下就雙向度情緒類別為例，說明其建立之基本程序。

LIWC2007 詞典的原始類別，主要是將情緒分為正向情緒與負向情緒兩類，在負向情緒下則進一步區分出焦慮詞、生氣詞與悲傷詞等三個子類。然而，在情緒相關的文獻很早就指出除了情緒價的區分向度外，激發程度也是一個重要的情緒分類向度（Russell, 1980, 2003）。跨文化比較論述也提出相對於歐美文化，亞洲文化強調低激發正向情緒的追求，視其為理想情緒（Tsai, Knutson, & Fung, 2006）。Lee、Lin、Huang 與 Fredrickson（2013）更直接聚焦低激發正向情緒，提出安適幸福感（peace of mind）的構念與測量。因此，研究者認為如果能夠在原有的情緒詞類上進一步區分高低激發程度，不僅能夠使 C-LIWC 詞典更貼近華人本土文化內涵，也有助提升其作為偵測心理特性之語言分析工具的重要性。

情緒詞類在修訂編製的過程中，除了 C-LIWC 情緒類別中原有的情緒詞之外，還參考了國內相關的情緒詞資料庫（卓淑玲，2011；鄭昭明，2011），也在團隊討論的過程中腦力激盪，聯想其他情緒相關詞來進行補充。最後，綜合以上不同來源所蒐集到的詞彙作為研究材料。考量研究參與者的填答容受度，研究者將這些研究材料切分成數個版本，分別請研究參與者對每個詞彙分別就情緒正負向程度與激發程度進行評量。將資料彙整後，並針對與卓淑玲（2011）情緒庫有所重疊的詞彙，進一步比對分析，發現兩者的相關非常高，確立了評量的可信度。最後則依據情緒價與激發性兩向度進行詞類的切分，得到四類情緒詞，分別為高激發正向、低激發正向、高激發負向及低激發負向。

在雙向度情緒類別建立完成後，研究者進一步進行效度檢驗。研究者在網路部落格與宗教電子報蒐集了佛教徒與基督教徒的學習感想或見證文章各 50 篇，進行文本內情緒詞類的使用率比較。結果如預期發現，雖然兩類文本都穩定看到正向情緒詞的使用率高於負向情緒詞，但是佛教徒更強調低激發的正向情緒，相對的基督教信徒則更看重高激發的正向情緒。此結果確立了情緒雙向度詞類具有穩定的預測效度（黃金蘭，2013）。

從上述例子可以看到，LIWC 軟體的使用並不局限於現有詞典使用頻率的計算，而是可以依據研究者的研究需求，建立起符合特定研究目的之詞典。這樣的彈性，也為開放資源或合作研究提供了一個便利的管道。為了提供一個資訊交流的園地，研究團隊也建置了一個 C-LIWC 專屬網站（<http://cliwc.weebly.com/>），鼓勵相關研究者未來能在此網站分享其自定的詞典檔，互相交流合作，以推廣相關的應用。

LIWC 標記的語言風格

以 LIWC 作為工具的研究取徑即在強調個體對於自然語言的使用方式，形成了個人特殊的語言風格，也顯露了其心理特性，因此語言使用頻率本身即具有特定意義。換句話說，就語言風格而言，重要的不是「說了什麼」，而是「如何說」。「說了什麼」通常會隨情境目標而有很大的不同，「如何說」則相對穩定，可形成個人的特色。那麼，究竟何謂個人的語言風格？以下是修改自 Chung 與 Pennebaker（2007）在介紹語言風格時所舉的例子。有三個人對冰淇淋感受的表達分別是：

A：我得說我喜歡冰淇淋。

B：吃一球冰淇淋的經驗當然是令人相當滿意。

C：讚！好物啊！

這三個人都表達了對冰淇淋的喜愛，但是他們所表達的方式卻展現了不同的個人特色。A 透露出語帶保留的特性；B 使用非常正式而拘謹的表達方式；C 就展現較為隨性不受拘束的特色。事實上此三人都對冰淇淋具有正向評價，而真正使人對 A、B 及 C 三人產生不同印象與感受的關鍵，乃在於不涉及實質內容意義的功能詞使用。以 LIWC 為研究工具的相關文獻，尤其重視此功能詞的語言標記，強調功能詞的運用頻率不僅反應了重要的心理意涵，而且具有實徵上的預測能力 (Tausczik & Pennebaker, 2010)。

上述冰淇淋的例子，不只是功能詞的使用不同，其實在內容詞的使用上也略有差異。以下，作者們再舉兩個例子來說明單純功能詞的使用也會清楚的展現出不同的語言風格。試著比較以下三個句子：

句子 1：那是一個杯子。

句子 2：這是一個杯子。

句子 3：這是我的杯子。

以上三個句子中，真正有意義的內容詞只有「杯子」，但是不同的功能詞卻傳達了杯子和描述者之間不同的關係變化。仔細比較三個句子，杯子與描述者的心理距離在句子 1 中較遠而客觀，到句子 3 中則顯得近而主觀。顯示功能詞的使用雖然不涉及任何實質的內容意義，卻反應了描述者內在的思維想法。

或如另一個例子，想像在超市閒逛時，遇到試吃攤販熱情地介紹商品，他可能是這樣說的：

A：這家公司的產品.....

B：我們公司的產品.....

C：我們家的產品.....

哪一個人的說法會讓你最想靠近試吃，停下來聽他介紹呢？A 試圖從第三者的客觀角度切入，自己和公司是兩個獨立的單位；B 則在所有格（我們）上反映了自己和公司的聯結；C 甚至更進一步，把公司視為自己的家，顯示出對公司或產品更高的認同感。姑且不論銷售員如何舌粲蓮花地介紹商品，功能詞僅僅是幾個字的不同，就反映出了銷售員與公司之間不同的關係，也帶給人不同的距離感受。

以詞為分析層次的心理意涵

本節將著重於回應 LIWC 這套研究工具在使用上經常為其他研究者質疑之處，亦即由於在指標計算過程中完全地去脈絡化，會不會因為過度的簡化而無法提供有意義的分析？作者們認為這是一個對工具使用相當大的誤解。例如，如上所述，功能詞往往被傳統的內容分析所忽略，但卻具有重要心理意涵，即可說明並非納入更多的內容就一定會提昇效度。除此之外，作者們也認為，任何量化的研究都不可避免的必須簡化資料，而一個新工具的使用應該是提供對於研究脈絡更豐富的觀察角度而不是限制，但究竟是提昇或限制了研究的豐富性，則有賴於研究者對於研究問題與研究脈絡的覺察與掌握。以下將對以 LIWC 為研究工具的基本立場進行詳細的說明。

LIWC 的字詞分析方式，旨在計算個體使用文字或語言的使用頻率，透過特定類別字詞使用頻率的高低，反映個體內在的思維想法、態度價值及情緒狀態等。有別於過去心理學研究較常針對語言和文字本身所傳達的內容進行分析，LIWC 單純計算語言文字的使用頻率作為語言標記，對前後文呈現的內容意涵忽略不計的方式常常受到不同的質疑和批評。對此，作者們認為不同的研究取向和研究方法，自有其優點與限制，真正重要的是各種研究方法背後是否具有正確的邏輯基礎，及其對於研究指標的定位與解釋方法是否合理。

以第一人稱代名詞「我」為例，目前以 LIWC 為研究方法的文獻普遍認為「我」的使用頻率反映了個體自我涉入程度，使用越多的「我」，代表個體自我涉入程度越

高。試著比較以下兩個內容意義完全大異其趣的句子：「我喜歡你」和「我討厭你」。兩者雖然在語意上確實是南轅北轍，但共同的地方是它們都反映了個體是以「我」為出發點，關注的焦點在自己身上，反映了自我聚焦的思維特性。而這也是 LIWC 研究方法立論的基礎，以電腦程式計算個人不同詞彙的使用頻率，固然無法將內容脈絡一併納入分析，卻依然得以突顯語言使用頻率本身的意義。且字詞使用頻率作為一種日常生活中自然使用的語言，是使用者本身不自覺的行為表現，反而更突顯其不可經由意識操弄影響，能夠更直接反應書寫者或說話者當下的思緒標記（Pennebaker, Mehl, & Niederhoffer, 2003），更補足了目前自陳式量表測量容易受到社會讚許性所影響的缺點。

納入情境脈絡的詮釋

既然 LIWC 是一種只計算字詞使用頻率的工具，是否代表其完全去除情境脈絡的意義呢？答案是否定的！事實上同樣的字詞在不同的書寫主題下，可能反映了截然不同的意涵。LIWC 此類以字詞使用頻率為指標的研究，一再強調語言的使用是作為一個標記（marker），反映的是個體在使用語言當下的心理特性。因此，使用此類工具的研究者，不在宣揚某一指標必然代表某種心理機制，或必然會帶來正面或負面的影響等完全去脈絡化的思維，更重要的是研究者如何使用字詞分析這樣的工具，來貼近自己的研究主題與脈絡。

以「我」的相關研究為例，目前學界普遍地共識是以「我」的使用頻率代表個體的自我涉入程度，但是應用在不同的主題之中，卻產生完全不同的解釋意義和預測效果。在 Newman、Pennebaker、Berry 與 Richards（2003）的研究探討「我」的使用頻率與說謊之間的關連，主張說實話的時候，其自我涉入程度較高，因而使用較多的我。他們請研究參與者針對「墮胎」這個主題，分別發表與自己想法相符與相悖的論點並說明理由。發表與自己想法相符的論點時，視為實話文本；發表與自己想法相悖的論點時，視為謊話文本。研究結果發現在實話文本中，個體相對使用較多的「我」。換句話說，當應用在比較個人說實話或說謊話的研究主題時，「我」的使用頻率呈現顯著差異。但此一結果絕非代表任何時候或情境，只要個體使用了越少的「我」，就代表個體在說謊。試著回想自己日常生活中的經驗，在自我介紹和替兩個互不相識的朋友介紹彼此的兩種情境中，前者通常使用較多的「我」，而後者通常使用較多的「他／她」。這樣的差別不能也不應該解釋成個體在後者的情境中傾向說謊。由此更突顯了 LIWC 作為研究工具時，研究者如何詮釋解讀語言標記的結果，將主題脈絡以及理論為基礎進行推論尤其關鍵。

除了應用在說謊的研究主題上，「我」的使用頻率也被應用到不同的研究中，展現了不同的指標意義與影響效果。例如在憂鬱相關的研究中，Rude 等人（2004）請研究參與者以書寫自己對於大學生活的想法和感受。結果發現正在憂鬱狀態、過去曾有憂鬱病史及沒有憂鬱病史的三組大學生，在書寫中使用「我」的頻率具有顯著差異，正在憂鬱狀態的個體書寫最多的「我」。另一方面，Simmons、Gordon 與 Chambless（2005）將「我」的使用頻率應用到親密關係中，邀請強迫症患者及其配偶到實驗室來，討論他們認為影響關係的重要議題，並以「我」的使用頻率反應個體在討論時的自我揭露程度（self-disclosure），結果卻似乎相反地發現當個體使用越多的「我」，其婚姻滿意度越高。

單單是「我」的使用頻率，應用到不同的研究中，即展現不同指標效果。由此可知，字詞使用頻率並非固定地代表某些特定意涵，而是作為一項研究工具，隨著研究主題、研究方法不同，用以標記研究者所感興趣之特定心理狀態或是思維歷程。對比 Rude 等人（2004）與 Simmons 等人（2005）的研究，Rude 等人的研究發現「我」的使用頻率為一負向指標，使用越多「我」的個體越憂鬱；Simmons 等人的研究則發現「我」的使用頻率為一正向指標，個體使用越多的我，其婚姻滿意度越高。由此更突顯了「我」的使用頻率沒有絕對的好或壞，完全取決於研究者依照研究主題脈絡、書寫主題與研究方法的不同，進行不同的推論而定。以語詞為單位的研究取徑，絕對非撒網捕魚式的由下而上指標探索，而是更需要有嚴謹理論推導為本的由上而下的指標檢驗。

文化之影響及跨文化研究的引申

另一方面，人稱代名詞在跨文化議題的相關應用上也佔有一席之地。近來不少應用促發派典 (priming paradigm) 所進行的文化研究，是以「我」促發個體化主義或獨立我；「我們」促發集體化主義或相依我，皆有穩定的效果 (Brewer & Gardner, 1996; Gardner, Gabriel, & Lee, 1999)。透過人稱代名詞不僅促發了不同的自我建構，更可能進一步影響個體的價值或行為傾向。更有趣的是，這樣的反應聯結並非單向性的，人稱代名詞固然可以用來促發不同的自我建構或文化價值，反過來促發不同的自我建構或文化價值也會影響個人使用人稱代名詞的傾向。Na 與 Choi (2009) 的研究發現，相較於促發相依我的個體，促發獨立我的個體傾向使用較少的「我們」，較多的「我」。他們進一步蒐集跨文化樣本，比較美國人與韓國人在使用人稱代名詞上的差異，結果同樣指出美國人 (高獨立我) 傾向使用較多的「我」，而韓國人 (高相依我) 傾向使用較多的「我們」。綜合上述的研究，第一人稱單／複數代名詞與特定的自我建構和文化之間的對應是相互性的，亦即人稱代名詞可以用來促發不同的自我建構與文化特性，另一方面不同的自我建構與文化也會影響個體對人稱代名詞的使用傾向。

近來，借重 LIWC 能夠快速有效率地分析文本資料的優勢，研究者得以追縱長時間累積的大量文字資料，以捕捉文化隨著時間變動的樣貌。Twenge、Campbell 與 Gentile (2012) 以 Google 電子書資料庫為來源，分析美國自 1960 年到 2008 年所出版的書籍中，人稱代名詞的使用頻率變化，發現「我」的使用頻率逐年上升，相對地，「我們」的使用頻率則逐年下降。此一結果也呼應了相關學者認為美國近年來日益重視個人獨特性，強化個體化自我的看法。文化的核心意涵不斷地隨著時間產生微妙地變化，不只影響了個人的思維信念與行為反應，也反映在語言、文字、書籍、音樂等不同的文化產物上。如此，使得人稱代名詞的使用頻率成為一個重要的指標，得以敏銳地反映文化變動的軌跡。

LIWC 透過電腦程式計算個體在語言上的使用頻率，在指標計算的過程中，固然沒有斟酌語意內涵與前後文脈絡，卻不影響其所得之語言使用頻率的指標意義與預測效果，反而更突顯功能詞等無關實質內容的詞彙，如何深刻地反應個體的語言風格。另一方面，LIWC 並非完全揚棄情境脈絡對個人語言使用特性的影響，作者們主張，當納入研究情境作為思考脈絡時，更得以透過對於使用時機與使用情境的掌握而豐富這些指標不同的詮釋意義。更精確地說，LIWC 只是一項以語言使用頻率為基礎的研究工具，有別於自陳式量表的得分高低具有明確的解釋意義，語言使用頻率的高低沒有所謂「絕對」的指標意涵，也展現了更大的應用彈性，端看研究者如何將其應用到自己的研究主題中，並對其賦予合理的指標意義。

結合以上討論，LIWC 在計算分析的過程中，雖然捨棄了前後脈絡和內容意義，但其計算語言使用頻率之指標，也具有獨特且合理的指標意義，不僅能夠反映個體內在的心理歷程，也能夠展現了穩定的預測力，對心理學的文本分析提供了不同的研究取向和分析觀點，近來也成為新興研究工具，廣泛地應用到不同的研究之中 (Tausczik & Pennebaker, 2010)。

質化與量化研究的橋樑

C-LIWC 的建立，不僅為開放式中文文本資料提供了嶄新的分析方式，對於過去本土心理學在質化與量化上各自為政的局面，也提供了一座可能的溝通橋樑。過去三十多年，本土心理學的發展已經奠定了豐富的理論基礎與實徵研究。在理論方面，不論是人格、社會心理學、自我、家庭等相關領域的研究者，都曾提出各種納入本土思維的理論概念；然而，在研究方法上，量化與質化研究雖然各自都展現了豐碩的成果，可惜由於彼此在研究的基本預設上還是有相當大的落差，因此，即使過去也有研究者曾經嘗試融

合兩種研究典範（李美枝、蕭人輔，1997），但實際上質化與量化協同合作的例子還是非常鮮見。

研究者以為，除了不同研究典範之間的基本假設不同之外，缺乏適當的工具也可能是質化與量化之間未能頻繁合作的重要原因之一。以諮商領域的研究來說，諮商者面對真實世界中個案的困頓，不僅是第一線的接觸現象，同時也是觀察現象最為深刻的一個領域，更提供了最具有本土性的材料。事實上，目前諮商領域的研究也確實累積了為數甚豐的質化訪談文本資料，雖然大多已經進行了一定程度的質化分析，但是，倘若這些由質化分析所得到的概念也能夠獲得量化分析的支持對照，那麼資料與概念或可更為豐富，同時也有助於質化與量化研究之間的對話。是以，研究者認為中文版 LIWC 詞典的建立，將可提供質化研究者更進一步地將他們的質化資料，轉化成為可供統計分析的量化資料，用來佐證各種具有文化特色的理論。透過這類研究的推展，不僅更深耕本土心理學的基礎，對於質化研究與量化研究之間互相對話，也提供了積極的貢獻。

舉例而言，金樹人（2005，2010）在日記書寫的方法上引進敘事治療的觀點，試圖讓受訪者從敘說者的視角，逐步跳脫成旁觀者的視角，藉此將個體與作為行動者的當下經驗拉出一段距離。金樹人稱此心理距離的移動為「心理位移」（psychological displacement），並提出了「心理位移日記書寫法」。隨後，他採用「紮根理論分析法」與「共識質化研究法」來分析日記書寫文本與焦點團體的資料中，並論證了「我、你、他」三種不同心理位置的結構特性與改變的機制。但是，金樹人（2010）所提出對於每一個書寫階段所發生的變化歷程，由於太過細膩，不僅難以由評分者評分，就算是研究參與者本身都不見得有足夠的覺察來主動的報告這些歷程是否真實的發生。為了對這些歷程提供量化的檢證，黃金蘭、張仁和、程威銓與林以正（付印中）就利用 C-LIWC 針對心理位移書寫法中的三種書寫位格（我位格、你位格及他位格）進行文本分析。他們根據過去 LIWC 的研究推論各項心理歷程所對應的文字特性進行推論，衍生具體可驗證的假設，其結果大致支持金樹人（2005，2010）對心理位移的位格特性之主張。例如，金樹人認為第一人稱書寫中具有視野狹小、情緒高張且主觀性較強的特性，在字詞分析中反映在顯著較高的聯語，副詞與確定詞的使用率，而交代細膩的特性則與諸如數量、數字、知覺、生理、相對詞有穩定的相關。這顯示，C-LIWC 不僅可以捕捉細微的心理歷程，同時也具有相當的效度，足以考驗質化研究所提出相當複雜的心理歷程。更重要的是，透過字詞分析工具，對於質性研究中研究參與者細膩而深入的觀察，提供了具體的支持證據，以實徵研究的分析結果，支持質性研究所歸納分析的觀點，如此使得質化與量化兩種不同的研究派典，產生對話與互動的可能。未來研究也可奠基於此，繼續深入書寫療癒的練習與應用。例如 Slatcher 與 Pennebaker（2006）的研究即指出情緒書寫會改變了個體的溝通方式，進而影響關係的穩定性。他們發現經過交往中的情侶經過情緒書寫的練習後，個體在對話中將會呈現更多的正向情緒詞，有助於維持關係的穩定性。足見 C-LIWC 不僅能夠捕捉情緒書寫的特徵，也能由此探討書寫療癒潛在的運作機制，提供實徵研究的檢驗證據。

LIWC 的能與不能

LIWC 以字詞使用頻率為分析之單位，為文本分析提供了一項非常重要的研究取徑。然而這個電腦化的量化研究工具當然並非萬能，它與任何研究工具一樣都有其能與不能之限制。本節將論述 LIWC 的優缺點。

綜合本文之論述，LIWC 的優點可以分四個方向加以說明。首先，當然就是它的電腦化帶來的優點，除了省時省力之外，更重要的是它大大提昇了文本分析的客觀性。第二，LIWC 程式搭配的 LIWC 詞典，提供了多項語文類別與心理特性類別，供研究者使用。因此，研究者無需自己建立分析與編碼架構，使用門檻相當低，非常容易上手。且以數千詞的詞典就能夠捕捉到一般文本近 85% 的偵測率，可見 LIWC 詞典已相當高比率涵蓋了一般人的語言使用詞彙。第三，LIWC 自建類別的特性，也大大提昇了使用彈性。

為研究者因應個別研究特性的不同需求，提供了方便的計算工具。最後且是作者們認為最重要的，LIWC 可以透過各類功能詞的使用狀況，捕捉到諸如語言使用風格等重要的心理特性指標。過去，功能詞常被稱為「垃圾詞」，最容易被一般人視而不見，即使專業質性分析研究者也多半把分析聚焦在內容詞上，而忽略功能詞的使用狀況。透過 LIWC 的相關研究卻發現，這類被人忽略的功能詞卻是反映心理特性與狀態的重要途徑(Chung & Pennebaker, 2007)。LIWC 為研究者提供了研究者意識所不容易偵測到的功能詞使用率，甚至捕捉到個體的語言使用風格。

然而，在諸多優點之外，LIWC 當然有它的缺點與限制之所在。LIWC 最主要的缺點就在於 LIWC 以字詞為分析單位此特性及其衍生的議題。首先，LIWC 就無法如同以語句甚或段落為單位的文本分析所捕捉到的深刻幽微的心理意涵。例如，作者曾在日本海嘯相關報導中，看到一段文字『一個祖母，看到海嘯即將撲來，將手中的嬰兒(孫子)高高舉起，交給鄰居，說了一句：「拜託你了。」就被海水沖走。嬰兒活下來。活著，是愛的力量。』短短六七十字，帶來相當令人動容的震撼。但是單看 LIWC 分析，恐怕只能得到一組破碎的變項值，無法看出文本傳達的意義。其次，循上例我們也不難發現，以字詞為分析單位的研究策略，通常需要透過不同類型且多文本的統計比較檢定，才能發覺某特定類型的文本特性，單一文本的 LIWC 分析，除非透過與「常模」或「基準率」的比較，否則其絕對意涵是不易顯露出來的。最後，以字詞為分析單位，自然就大大折損了前後文所提供的脈絡訊息。此脈絡訊息斷裂的缺點，也使得研究者在使用 LIWC 的變項時，要格外小心，尤其在意涵解釋上。以情緒詞類為例，倘若文本中有一個句子：「我今天不開心。」這個句意表達的是一個偏向負面的情緒(至少不是正向情緒)，但 LIWC 分析標記的是一個「否定詞」和一個「正向情緒詞」的使用次，而不是標記成一個「負向情緒詞」的使用次。因此如果將各類情緒詞的使用率當成個體的各種情緒狀態，就可能造成誤差。倘若我們仔細分析，「不開心」表達出的是個體用自己「沒有」處在一個「正向情緒」而非自己處在一個「負向情緒」來表達自己的狀態。換言之，在 LIWC，正向情緒詞的使用率不宜作為正向情緒感受的頻率或強度，而是個體把焦點放在「正向情緒」的頻率或程度。嚴格而言，LIWC 的各類心理特性變項反映的是個體聚焦在該變項的程度。固然，針對上面的例子，作者們提醒 LIWC 使用者需謹慎解釋變項意涵。不可諱言，目前 LIWC2007 或 C-LIWC 確實無法處理這類透過否定詞加某類情緒來表達對立面情緒的議題。研究者若需要特別處理此類議題，則需自建詞典或者是利用其他的文本分析軟體，甚或自行撰寫程式的方式加以處理。

結語

LIWC 以電腦軟體計算字詞使用頻率的研究派典，為語言文本的資料分析畫下一頁嶄新的可能，C-LIWC 的建立更開啟了一扇心靈之窗，幫助研究者得以透過語言使用特性，窺探廣大中文使用者內在的心理歷程與思維奧秘。未來在輔導與諮商相關領域的應用上，亦可充分利用 C-LIWC 來做為輔助分析的工具。例如諮商員或督導可以分析個案晤談逐字稿，或分析個案的自傳式記憶，搭配質性分析架構，以更了解其內在的心理歷程與治療過程的變化。在臨床上，治療師也可以依據對於個案的各種書寫材料進行語文分析，作為治療的參考與評估，甚至於直接作為與個案會談的材料。相信透過 LIWC 這樣全新的研究工具派典的引進，必然有助於華心理學在過去研究基礎中更上層樓！

參考文獻

- 李美枝、蕭人輔(1997)。量化與質化方法在本土心理學研究融合的思維脈絡與研究實例。「第四屆華人心理與行為科學學術研討會」發表之論文，台北。
- 卓淑玲(2011)。情緒詞資料庫。台灣地區華人情緒與相關心理生理資料庫。取自 <http://ssnre.psy.ntu.edu.tw/ssnredb/member.php?action=lostpasswd>

- 金樹人 (2005)。心理位移辯證效果之敘事分析。行政院國家科學委員會專題研究成果報告 (編號: NSC93-2413-H-003-001)。
- 金樹人 (2010)。心理位移之結構特性及其辯證現象之分析: 自我多重面向的敘寫與敘說。《中華輔導與諮商學報》，28，191-233。
- 唐思柔 (2011)。物質消費與經驗消費的對比: 字詞分析的探索與應用。行政院國家科學委員會大專生專題研究成果報告。
- 張硯評 (2012)。感恩表達與配偶之生活適應 (碩士論文)。國立臺灣大學。(系統編號 100NTU05071015)
- 黃金蘭 (2013)。台灣版語文探索與字詞計算字典檔之修訂與信效度檢驗。行政院國家科學委員會專題研究成果報告 (編號: NSC100-2410-H-011-012)。
- 黃金蘭、Chung, C. K.、Hui, N.、林以正、謝亦泰、程威銓、Lam, B.、Bond, M., Pennebaker, J. W. (2012)。中文版語文探索與字詞計算字典之建立。《中華心理學刊》，54，185-201。
- 黃金蘭、張仁和、程威銓、林以正 (付印中)。我你他的轉變: 以字詞分析探討大學生心理位移書寫文本之位格特性。《中華輔導與諮商學報》。
- 鄭昭明 (2011)。情緒詞的加性樹狀結構。「第七屆華人心理學家國際研討會」發表之論文，台北。
- 竇郁文 (2012)。言語變色龍: 說服與語言型態匹配度 (碩士論文)。國立臺灣大學。(系統編號 100NTU05071016)
- Arntz, A., Hawke, L. D., Bamelis, L., Spinhoven, P., & Molendijk, M. L. (2012). Changes in natural language use as an indicator of psychotherapeutic change in personality disorders. *Behaviour Research and Therapy*, 50(3), 191-202.
- Brewer, M. B., & Gardner, W. (1996). Who is this "We"? Levels of collective identity and self representations. *Journal of Personality and Social Psychology*, 71(1), 83.
- Burton, C. M., & King, L. A. (2004). The health benefits of writing about intensely positive experiences. *Journal of Research in Personality*, 38(2), 150-163.
- Carter, T. J., & Gilovich, T. (2010). The relative relativity of material and experiential purchases. *Journal of Personality and Social Psychology*, 98(1), 146.
- Chung, C. K., & Pennebaker, J. W. (2007). The psychological function of function words. In K. Fiedler (Ed.), *Social communication: Frontiers of social psychology* (pp. 343-359). New York, NY: Psychology Press.
- Davies, H., Swan, N., Schmidt, U., & Tchanturia, K. (2012). An experimental investigation of verbal expression of emotion in anorexia and bulimia nervosa. *European Eating Disorders Review*, 20(6), 476-483.
- Francis, M. E., & Pennebaker, J. W. (1993). LIWC: Linguistic inquiry and word count. Dallas, T.: Southern Methodist University.
- Gardner, W. L., Gabriel, S., & Lee, A. Y. (1999). "I" value freedom, but "we" value relationships: Self-construal priming mirrors cultural differences in judgment. *Psychological Science*, 10(4), 321-326.
- Gonzales, A. L., Hancock, J. T., & Pennebaker, J. W. (2010). Language style matching as a predictor of social dynamics in small groups. *Communication Research*, 37(1), 3-19.
- Graybeal, A., Sexton, J. D., & Pennebaker, J. W. (2002). The role of story-making in disclosure writing: The psychometrics of narrative. *Psychology & Health*, 17(5), 571-581.
- Handelman, L. D., & Lester, D. (2007). The content of suicide notes from attempters and completers. *Crisis: The Journal of Crisis Intervention and Suicide Prevention*, 28(2), 102-104.
- Hermans, H. J. M., Rijks, T. I., & Kempen, H. J. G. (1993). Imaginal dialogues in the self: Theory and method. *Journal of Personality*, 61(2), 207-236.
- Hirsh, J. B., & Peterson, J. B. (2009). Personality and language use in self-narratives. *Journal of Research in Personality*, 43(3), 524-527.
- Holtgraves, T., McNamara, P., Cappaert, K., & Durso, R. (2010). Linguistic correlates of asymmetric motor symptom severity in Parkinson's disease. *Brain and cognition*, 72(2),

- 189-196.
- Ireland, M. E., & Pennebaker, J. W. (2010). Language Style Matching in Writing: Synchrony in Essays, Correspondence, and Poetry. *Journal of Personality and Social Psychology*, *99*(3), 549-571.
- Ireland, M. E., Slatcher, R. B., Eastwick, P. W., Scissors, L. E., Finkel, E. J., & Pennebaker, J. W. (2011). Language style matching predicts relationship initiation and stability. *Psychological Science*, *22*(1), 39-44.
- James, W. (1890). *The principles of psychology*: New York, NY: Dover Publications.
- Lee, Y. C., Lin, Y. C., Huang, C. L., & Fredrickson, B. L. (2013). The construct and measurement of peace of mind. *Journal of Happiness Studies*, *14*(2), 571-590.
- Mehl, M. R., Gosling, S. D., & Pennebaker, J. W. (2006). Personality in its natural habitat: manifestations and implicit folk theories of personality in daily life. *J Pers Soc Psychol*, *90*(5), 862-877.
- Na, J., & Choi, I. (2009). Culture and first-person pronouns. *Pers Soc Psychol Bull*, *35*(11), 1492-1499.
- Newman, M. L., Pennebaker, J. W., Berry, D. S., & Richards, J. M. (2003). Lying words: Predicting deception from linguistic styles. *Personality and Social Psychology Bulletin*, *29*(5), 665.
- Pennebaker, J. W. (1997). Writing about emotional experiences as a therapeutic process. *Psychological Science*, *8*(3), 162-166.
- Pennebaker, J. W. (2011). *The secret life of pronouns: What our words say about us*. New York, NY: Bloomsbury Press.
- Pennebaker, J. W., & Beall, S. (1986). Confronting a traumatic event: Toward an understanding of inhibition and disease. *Journal of Abnormal Psychology*, *95*(3), 274-281.
- Pennebaker, J. W., Booth, R. J., & Francis, M. E. (2007). LIWC2007: Linguistic inquiry and word count. *Austin, Texas: liwc. net*.
- Pennebaker, J. W., Chung, C. K., Ireland, M., Gonzales, A., & Booth, R. J. (2007). The development and psychometric properties of LIWC2007. *Austin, TX, LIWC. Net*.
- Pennebaker, J. W., & Francis, M. E. (1996). Cognitive, emotional, and language processes in disclosure. *Cognition & Emotion*, *10*(6), 601-626.
- Pennebaker, J. W., Francis, M. E., & Booth, R. J. (2001). Linguistic inquiry and word count: LIWC. *Computer software*. Mahwah, NJ: Erlbaum.
- Pennebaker, J. W., & King, L. A. (1999). Linguistic styles: Language use as an individual difference. *Journal of Personality and Social Psychology*, *77*(6), 1296-1312.
- Pennebaker, J. W., Mehl, M. R., & Niederhoffer, K. G. (2003). Psychological aspects of natural language use: Our words, our selves. *Annual Review of Psychology*, *54*, 547-577.
- Robinson, R. L., Navea, R., & Ickes, W. (2013). Predicting Final Course Performance From Students' Written Self-Introductions: A LIWC Analysis. *Journal of Language and Social Psychology*.
- Rude, S., Gortner, E. M., & Pennebaker, J. W. (2004). Language use of depressed and depression-vulnerable college students. *Cognition and Emotion*, *18*(8), 1121-1133.
- Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality & Social Psychology*, *39*(6), 1161-1178.
- Russell, J. A. (2003). Core Affect and the Psychological Construction of Emotion. *Psychological Review*, *110*(1), 145-172.
- Simmons, R. A., Gordon, P. C., & Chambless, D. L. (2005). Pronouns in marital interaction: What do "you" and "I" say about marital health? *Psychological Science*, *16*(12), 932-936.
- Slatcher, R. B., & Pennebaker, J. W. (2006). How do I love thee? Let me count the words: the social effects of expressive writing. *Psychological Science*, *17*(8), 660-664.
- Smyth, J. M. (1998). Written emotional expression: effect sizes, outcome types, and moderating variables. *Journal of consulting and clinical psychology*, *66*(1), 174.
- Stirman, S. W., & Pennebaker, J. W. (2001). Word use in the poetry of suicidal and nonsuicidal poets. *Psychosomatic Medicine*, *63*(4), 517-522.
- Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and

- computerized text analysis methods. *Journal of Language and Social Psychology*, 29(1), 24-54.
- Tsai, J. L., Knutson, B., & Fung, H. H. (2006). Cultural variation in affect valuation. *Journal of Personality and Social Psychology*, 90(2), 288-307.
- Twenge, J. M., Campbell, W. K., & Gentile, B. (2012). Changes in Pronoun Use in American Books and the Rise of Individualism, 1960-2008. *Journal of Cross-Cultural Psychology*, 44(3), 406-415.
- Van Boven, L. (2005). Experientialism, materialism, and the pursuit of happiness. *Review of general psychology*, 9(2), 132.
- Vanheule, S., Meganck, R., & Desmet, M. (2011). Alexithymia, social detachment and cognitive processing. *Psychiatry research*, 190(1), 49-51.
- Yarkoni, T. (2010). Personality in 100,000 words: A large-scale analysis of personality and word use among bloggers. *Journal of Research in Personality*, 44(3), 363-373.

收稿日期：2013 年 5 月 20 日

一稿修訂日期：2013 年 8 月 31 日

二稿修訂日期：2013 年 11 月 10 日

接受刊登日期：2013 年 11 月 10 日

From LIWC to C-LIWC: The Potentialities of Computerized Chinese Word Count Analysis

Wei-Fang Lin

Department of
Psychology
National Taiwan
University

Chin-Lan Huang

Department of
Humanities and Social
Sciences
National Taiwan
University of Science and
Technology

Yi-Cheng Lin

Department of
Psychology
National Taiwan University

Abstract

Previous research has demonstrated that expressive writing paradigm has positive effects on individual physical and psychological health. The cognitive and emotional processes related to this effect might be very well revealed in the language usage involved. However, although it seems intuitive that internal psychological processes are expressed in various forms of language, how to accurately translate natural language into operational indices remains to be a hard challenge to psychologists. To achieve this goal, Pennebaker and his colleagues psychometric properties of language use. LIWC2007 is the latest version which has been applied to various research topics and has demonstrated good reliability and validities. The current article reviewed the development and applications of both the original LIWC and also the Chinese version of LIWC (C-LIWC). Even though LIWC simply does word counting and reports the percentages of word categories from texts, the exploratory power of LIWC can be highly enriched in combination with theoretical and contextual sensitivities. The authors argue that C-LIWC builds a bridge between qualitative and quantitative research, which is especially useful in the field of counseling psychology.

Keywords: expressive writing, text-analysis, LIWC