

數位語藝批評與文本探勘工具—以反核臉書粉絲團形塑幻想主題為例

曹開明、黃鈴媚、劉大華

摘要

本文數位語藝具有「互動性」與「互文性」特質，形成龐大的網路文本量，導致數位語藝批評研究的重大挑戰。本研究主張進行數位語藝研究時，可採取「混合取徑」原則：提供結合文本探勘工具與語藝研究者評斷分析的步驟，以擴大語藝批評在數位環境下的實用性。本研究並以反核臉書粉絲團進行案例分析，將資料科學發展出之文本探勘工具結合人力瀏覽、評判方式，進行幻想主題分析。研究結果指出，採取文本探勘工具，除有利於確認被「覆誦」的關鍵字，更有助於語藝研究者論證社群成員共享與共創的「幻想主題」。

- ◎ 關鍵字：數位語藝、幻想主題分析、覆誦、文本探勘、混合取徑
- ◎ 本文第一作者曹開明為國防大學新聞學系助理教授；第二作者黃鈴媚為世新大學口傳系教授；第三作者劉大華為國防大學新聞學系講師。
- ◎ 通訊作者為黃鈴媚，聯絡方式：E-mail：lmhuang@mail.shu.edu.tw；電話：02-22368225 轉3165；通訊處：世新大學116台北市木柵路一段17巷1號。
- ◎ 收稿日期：2016/06/30 接受日期：2017/01/24
- ◎ 本研究獲得科技部「台灣能源政策溝通與建構核能共識：氣候變遷因應與核能安全之論述策略分析」(MOST103-2511-S-128-001-MY3) 專題研究計畫之經費補助，研究成果初稿〈數位語藝批評方法結合電腦文本探勘工具初探—以反核臉書粉絲團形塑幻想主題為例〉曾發表於中華傳播學會2016年年會。

An Exploration of Text Mining Tools on Digital Rhetoric Criticism – How Anti-nuclear Facebook Fanpages Shaped Fantasy Themes

Kai-Ming Tsao, Lin-Mei Huang, Ta-Hua Liu

Abstract

The rhetorics in the digital world demonstrate interactivity and intertextuality, generating the massive volume of internet texts. This study suggest that the frame of “hybrid approach” combine text mining tools with manual browsing and coding, may be better utilized in the digital world to confirm the analytical results and to rectify any weaknesses,thereby expanding the usability of rhetorical criticism. Moreover, This study takes on the subject of an anti-nuclear Facebook channel. The computer text mining tool is used to determine and analyze the keywords that are chained out in the text. The rhetorical collocation of the main entries and responses is found having extension of meanings; through them, the fantasy themes and rhetorical vision are further strengthened.

⊙ Keywords: digital rhetoric, fantasy theme analysis, chain out, text mining, hybrid approach

⊙ Kai-Ming Tsao is Assistant Professor in the Department of Journalism, National Defense University, Lin-Mei Huang is Professor in the Department of Speech Communication, Shih Hsin University, and Ta-Hua Liu is Lecturer in the Department of Journalism, National Defense University.

⊙ E-mail: lmhuang@mail.shu.edu.tw

⊙ Received: 2016/06/30 Accepted: 2017/01/24

壹、前言

「數位語藝」(digital rhetoric)一詞,早於1989年由美國學者Richard Lanham提出(Eyman, 2015:24),強調針對電腦中介傳播情境產製的數位文本進行語藝探究時,必須留意其「超文本」(hypertext)的特質;因而在研究方法上,也必須與時俱進。正因為線上文本具有「互文性」(intertextuality)及「互動性」(interactivity)的特性(Warnick, 2007);「數位語藝」的意涵,即在於探究、分析其文本的語藝策略,以及形塑認同與建構網路社群的特徵。譬如,沈錦惠(2009)提出「電子符號行動者」(electronic symbolic actors)概念,強調網路上的活躍言者,有能力運用網路科技傳布訊息互動、避免傳統大眾傳播單向式溝通與擺脫媒體接近性的侷限;尤其在面對實際「語藝情境」(rhetorical situation)的「急迫性」(exigence)時,具備提出符號行動的「語藝能動者」(rhetorical agent),得以召喚認同與凝聚共識,彰顯傳播行動主體的具體實踐。

然而,網路社群交織討論議題時,文本數量龐大加上不斷更新累積;以傳統人力方式直接瀏覽搜尋,容易疏漏或忽略,恐導致文本分析時缺乏足夠論證依據,成為當前網路時代進行數位語藝批評的主要困境。為了解決語藝研究面臨網路龐大資料的問題,除了限縮文本蒐集範圍、研究小樣本文本外,大數據研究方法或可提供解決方法。大數據(big data)又稱巨量資料,一般以5個V來標誌大數據的特性,分別是Volume(容量)、Velocity(速度)和Variety(多樣性)、Veracity(真實性)和Value(價值)。研究者可以透「資料探勘」(data mining)、「文字探勘」(text mining)、社群網絡分析、資訊視覺化及統計學等工具技術,從資訊大海中挖掘有價值的資料與知識。然而,假如研究主題研究涉及大量文字資料的分析判讀與應用,「文字探勘」就會扮演比較重要的角色(李欣宜,2015;劉吉軒,2016)。

換言之,為了有效掌握網路社群的意見領袖,並確知其具備「電子符號行動者」的語藝特徵,若能善用現有的大數據研究方法中的「文字探勘」,選擇適當的分析工具,將有益於解決上述數位語藝批評需要面對的難題。「文字探勘」係指從非結構化的文本資料中挖掘與建立資訊的過程,其技術結合資訊檢索、自然語言處理、統計分析和機器學習等多個學科領域。一般認為「文字探勘」是「資料探勘」技術的延伸,

兩者的差異在於，「資料探勘」是針對資料庫為基礎的結構化資料進行分析，探求其趨勢、類型、相關性和預測模型；相對地，「文字探勘」處理的資料大多屬於無固定格式與架構之文字資料檔案。因此，「文字探勘」需要針對文字內容進行處理，透過斷詞處理、關鍵字篩選等技術，轉換成適合後續處理與分析的結構式資料（Liu & Lu, 2011；陳良駒、傅振華、楊誌瑋，2010）。

值得注意的是，語藝研究取徑十分重視內、外在脈絡分析，如何結合大數據研究之分析工具，俾利提供詮釋時之論證與舉證效力，成為數位語藝研究者在研究方法上的重大挑戰。例如，Ridolfo（2006）提出C.O.D.E.—Comprehensive Online Document Evaluation的理論概念，強調數位語藝研究，不應僅引述網路文本，而應該善用網路分析工具，以進行批判性網路文本研究¹。然而，Ridolfo所建議的網路分析，主要是在確認網路流量以及取得網路言者身份的資訊，在進行網路文本探究上，仍顯不足；也凸顯目前數位語藝研究如何結合大數據分析，在研究方法上尚仍處於初探階段。

因此，參酌Lewis、Zamith和Hermida（2013）所提出的「混合取徑」（hybrid approach），本研究意欲探討語藝批評方法結合大數據研究方法中的「文字探勘」時，如何在不同研究階段的網路文本選擇、分析以及詮釋上，透過研究者的適時與適度介入，逐步重構文本脈絡，以提供研究者有效的論證或詮釋所需要的依據資料。即如陳佳靖（2004：119-120）指出，進行網路文化語藝分析時，內、外在脈絡分析容易淪為過度簡化或標籤化的界定；除非研究者能夠清楚界定兩者之間的範疇（亦即明確的概念化定義），並提供一套探討外在脈絡的系統性分析方法（亦即明確的操作化定義），否則對數位語藝批評學者而言，內在分析還是首要重點（林靜伶，2000）。因此，針對現階段的數位語藝批評研究，本研究的「混合取徑」主要還是聚焦於協助研究者掌握內在脈絡的相關特徵，以利於文本的詮釋與論證。

1 Ridolfo（2006）曾討論到可用於數位語藝研究的網路分析工具，包括全球WHOIS查詢系統。其中，WHOIS是一個用於網頁IP查詢的網路平台；ARIN（American Registry of Internet Numbers）則是北美地區的網路IP註冊機構，研究者可透過此一機制，查詢網址註冊者的資訊，以及分析網路流量。

貳、數位語藝特徵與研究方法

Zappen (2005:323-324) 指出：「數位語藝理論是將著重探討說服的傳統語藝研究，轉移到數位空間，藉以擴展語藝理論與批評方法的研究領域。」易言之，數位語藝並非單一的理論概念，而是摻合傳統語藝理論概念與其他社會學與文化研究等理論思維，進一步提升語藝科學與批評方法的效用，進而成為整合型的理論概念。Zappen 並指出，數位語藝研究主要聚焦在下列四個主要範疇：

- 一、探究與分析數位文本中所使用的語藝策略
- 二、確認新媒體的語藝特徵、作用與限制
- 三、探究數位認同的形塑
- 四、探究建構語藝社群的效用

由此可知，進行數位語藝研究時，必須留意數位文本的特性，進而探究網路言者的語藝策略，以及形塑網路社群認同的效用。相對地，數位語藝研究必須在方法取徑上有所因應，俾能有效論證與詮釋。Eyman (2015:65) 即指出，傳統語藝五要素—創作、組織、風格、傳遞與記憶，應用在數位實踐中，其意義即有所變化（參見表一）；亦即在數位空間中進行語藝研究與批評，數位科技發展下形成的新媒體溝通平台以及可資運用傳播工具，形成了數位實踐的語藝五要素。因此，在探究數位空間中，網路活躍言者如何達到說服、獲致認同，必需重視研究取徑與數位語藝特徵相互扣連。

以 Warnick (2007) 出版之《線上語藝：全球資訊網的說服與政治》專書為例，為探究政黨、抗議團體如何在網路上引發公眾注意，並形成主流的社會或政治議題，Warnick 採取一系列針對網路政策討論的個案研究，依此取徑加以標示出數位語藝研究應重視數位文本所具有的「互文性」(intertextuality) 及「互動性」(interactivity) 的特性。Warnick 強調，網路上的言者涉及運用數位文本「互為言者」特性，使得網路言者與閱聽眾間交互詮釋議題的文本，成為語藝策略的資源；同時，Warnick 也援引語藝學者 Kenneth Burke 所提出語藝的主要作用在於形塑「認同化」(identification)²的

概念，強調網路論述在於網路言者提出議題觀點後，引發其他網路言者參與互動與關切，交織形成「符號行動」（symbolic action）進而形塑社群認同。

此外，數位語藝的「互動性」與「互文性」特質，也促成對既有語藝理論的反思。例如Pashaei（2010）以美國「911事件」後網路部落格對於穆斯林的認知情形為例，指出「語藝情境」（rhetorical situation）理論所強調特定的「緊急狀態」才引發語藝論述的論點，在面對網路情境時應加以修改；並提出相關部落格的討論為例證，認為是因為閱聽眾持續瀏覽相關網路文本並加回應意見後，才引發對穆斯林的關切。因此，Pashaei（2010）認為，應該以「語藝生態學」（rhetorical ecology）取代「語藝情境」概念，更能詮釋言者與閱聽眾互為文本的特性。然而，「語藝生態學」概念，是否真能符合數位空間的人際互動，並作為數位語藝的理論架構，因為缺乏更多研究文獻佐證，目前尚未有定論。然而，至少可確知的是，上述數位語藝的特質對語藝研究學者而言，無論在理論上或方法上，都帶來亟須突破的難題。

依據Eyman（2015）對數位語藝研究方法的綜總結果，除了前述指出的

表一：語藝要素的數位實踐

語藝要素	傳統定義與使用概念	數位實踐
創作	尋求可資說服的方法	網絡資訊彙整與協商；結合新媒體展現
組織	形成論說次序	選擇與重組既有的網路資訊，形成數位媒體控制
風格	言說方法	善用數位文本與影像（色彩、字型、使用適合的多媒體呈現）
傳遞	言說展演	善用與連結各種網路平台傳遞訊息
記憶	受眾記憶	具備數位資訊儲存、移除與控制的知能

2 Burke（1969）指出，語藝的符號行動功能就是在於產出「認同化」，也就是形塑出「我群感」的想像概念。易言之，語言即是一種行動，因為我們總是在言說中展現出特定動機籲求，藉以激勵他人與我們共享，這正是「動機的話藝」（the motives of rhetoric）的作用。所以，當我們討論語言即是行動時，重點已非聚焦於如何說服他人的表象作法，而是在如何召喚人們「認同化」（identification）的架構下，用以促成合作。

C.O.D.E.方法，藉以確認特定網站的流量，以及運用ARIN或是全球WHOIS查詢與統計網路使用者的身份與資訊外；也可以使用「社會網絡分析」（Social Network Analysis, SNA），亦即質化資料蒐集（例如訪談法及觀察法）結合電腦的量化數據分析工具，以探討網路人際互動情形；或是使用量化內容分析（Content Analysis）方式，進行文本特定概念的頻次分析以及文本與使用者間的相互關聯性。由上述方法之簡述可知，數位語藝研究因為面臨龐大繁雜的資料數據，恐非人力所非解決；運用電腦運算工具，有助於解決此一難題。誠如盧安邦（2014）指出，質化研究可與大數據分析相互為用，並「使用資訊科學分析工具，以『人工智慧』取代『工人智慧』，幫助處理大量資訊。」

然而，儘管蔡鴻濱（2004）曾指出，常用的十一種語藝批評方法³，都有應用在網路文化研究的合理性；但是，國內既有語藝批評研究在蒐集、編碼網路文本時，仍傾向於依賴人力（少數例外請參見黃鈴媚、曹開明、沈錦惠、劉大華，2015）。語藝批評大致區分為內在文本與外在脈絡分析，目前國內針對網路文本進行內在分析之研究，主要由研究者直接進入研究場域，也就是直接在網路平台中審視文本，加以歸納、舉證外在脈絡特徵加以論證後，進而綜整出語藝策略；最後再加以詮釋與批評，或是分析其言說動機。但是，若僅靠人工瀏覽、編碼的方式，使得數位語藝批評的範疇只能侷限於處理小樣本資料；相對地，假如能夠運用文本探勘工具，便能擴大文本處理數量。目前結合文本探勘工具進行數位語藝研究，尚在起步階段，是一個值得探究的範疇。比較困難地是，研究者如何善用文本探勘工具，以便能夠同時擴大文本處理數量並兼顧語藝批評的詮釋與批評精神？這正是本研究希望透過分析反核臉書粉絲團來回答的問題。

3 包括新亞里斯多德批評、類型批評、隱喻批評、戲劇五因批評、幻想主題分析、女性主義批評、敘事批評、論辯批評、意念分析、群集分析與批判語藝。

參、數位語藝批評揉合文本探勘工具之「混合取徑」

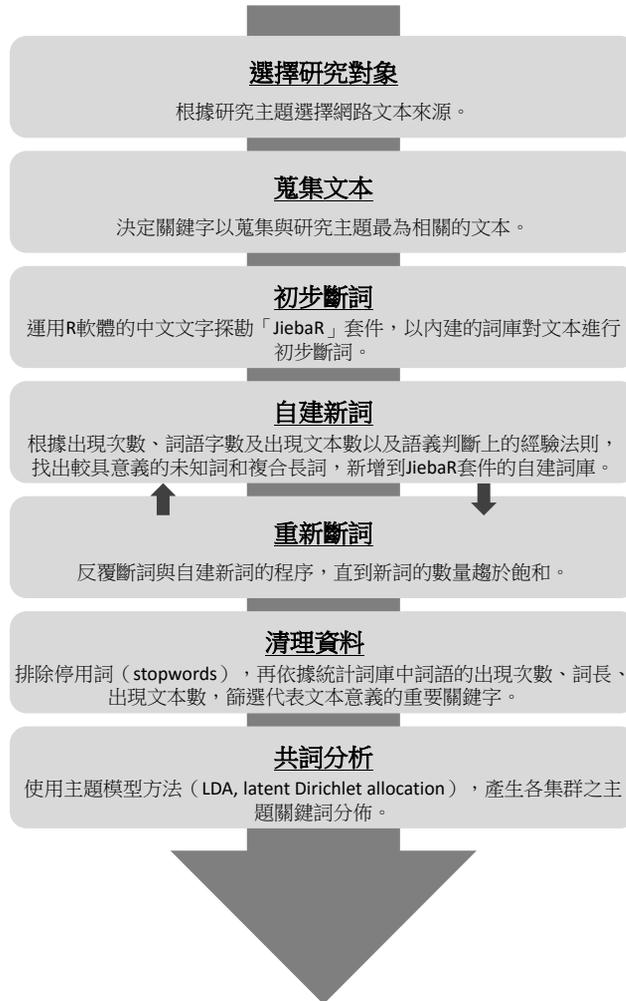
一、數位語藝批評揉合文本探勘工具之關鍵問題

首先，以文本探勘工具的效用而言，主要是減少人力耗費；研究者透過電腦演算法以挖掘文本意義，將文本內容數據化為詞彙出現頻率（陳世榮，2015：688）。然而，陳世榮（同上引）也指出，文本探勘工具的運用，仍需在自身學術訓練的前提下，慎選與研究主題有關的文本，並明白揭露推論設計與判讀程序，作為合理論證的依據。換言之，研究者在運用文本探勘工具時，研究者在各個研究階段的角色，或許很難如同盧安邦（2014）文中所謂的「以『人工智慧』取代『工人智慧』」般截然劃分。也就是說，當研究者運用『人工智慧』（文本探勘工具）時，『工人智慧』（研究者）決定『人工智慧』是否能夠得出對研究主題最相關的研究資料與結果。因此，本研究認為，即使借重資訊科學帶來的便利，研究者在各個研究階段都必須能夠保持適時、適度的參與。即以本研究為例，在圖一所建議的『人工智慧』研究階段，都有研究者必須扮演的角色。

其次，李蔡彥、鄭宇君（2011）討論資訊科技與新媒體研究發展時，指出一個有待進一步探索的問題：「在新媒體網路化、行動化、個人化的發展趨勢下，如何選擇一個合適的分析單位，進而整合質化研究與使用者操作記錄兩種不同研究架構所收集的分析資料？」因此，文本探勘工具能否與語藝批評取徑相互整合，以獲致合適的分析單位與資料，是數位語藝研究應加探究的面向。

也就是說，將文本探勘工具運用到數位語藝研究，需要思考文本探勘對特定語藝批評方法的協助程度。因此，本研究認為，當語藝學者思考如何運用文本探勘工具時，語藝學者必須特別關注，哪些語藝批評方法比較適合透過文本探勘的協助來進行後續分析？以本研究為例，本研究選擇「幻想主題分析」（FTA），就是因為考量透過文本探勘分析，可以確認在反核臉書主文與回文中共同出現的重要關鍵字，便於驗證「幻想主題」的「象徵性線索（symbolic cues）之「覆誦」。

圖一：反核臉書「文字探勘」分析步驟



此外，Boyd和Crawford（2012: 666-670）對於大數據分析曾提出批判與警告，認為過度標榜大數據分析，恐怕會被客觀性與精確性所誤導；尤其是質性研究著重在故事詮釋，涉及文本脈絡分析，大數據研究恐怕不如縮小數據範圍進行分析，以免獨尊量化統計，反而喪失研究意義。再者，大數據分析所蒐整的資料，是否真能回答研究問題（諸如資料是否具有信、效度等問題），均是研究者在運用大數據分析應加留意的面向。對於特別重視內、外脈絡的「厚描」（thick description）的語藝研究，更應留意上述大數據分析的侷限或量化偏見。因此，本研究認為，當語藝學者思考如何運

用文本探勘工具時，語藝學者必須特別留意，哪些文本探勘工具比較有助於文本內、外脈絡之重構？以本研究為例，本研究選擇「主題模型分析」（LDA）而非「社會網絡分析」（SNA），就是因為前者能夠更明確呈現臉書主文與回文之間詞語共現情形，並根據具有「覆誦」特徵的重要關鍵字來審視主、回文詞群間的異同，有利重構其內在脈絡。

簡言之，將文本探勘工具運用到數位語藝批評研究上，除了借重其有效處理大量網路文本的長處外，也必須了解其限制，同時思考在不同的研究階段，研究者如何將其優點發揮到最大而又如何補強其不足之處？本研究認為，除了研究者在借重文本探勘工具的『人工智慧』研究階段中，必須能夠保持適時、適度的參與；本研究也認同Lewis、Zamith和Hermida（2013）所建議的「混合取徑」（hybrid approach），針對『人工智慧』研究階段成果，假如數位語藝研究者可以佐以適量的『工人智慧』分析，研究者能夠更有信心地提出研究結果。然而，相對於Lewis、Zamith和Hermida（2013）的「混合取徑」，本研究更加強調無論在『人工智慧』研究階段或『工人智慧』研究階段（盧安邦，2014），研究者都必須保持適時、適度的參與。具體做法簡述如下，實際操作步驟則將在反核臉書粉絲專頁分析中呈現。

二、文本探勘工具揉合語藝學者論證之「混合取徑」

大數據研究必須直接面對資料蒐整、歸納過程中可能產生的刪減與遺漏，誠如Sweden等人（2015:852-853）指出的「大數據總謬誤」（Big Data Total Error, BDTE），諸如資料蒐整時不當省略、刪除、不當命名與描述或是命名正確但資料錯置等加總後，均會影響到大數據分析的信效度。因此，在使用文本探勘工具進行分析的歷程中，研究者必須謹慎思考自己究竟應該扮演何種角色，方能確保文本分析的完整性與避免論據不足；尤其是語藝批評方法重視探知文本與脈絡特徵，以歸納出相關的語藝策略及其背後的言說動機。假如只是運用文本探勘工具得出關鍵字、頻次以及語意共現情形，便據此推論研究結果，難免受到質疑；因為關鍵字詞頻次與共現情形，仍不足以論證語藝理論與推知語藝策略背後的脈絡問題。

誠如Lewis、Zamith和Hermida（2013）所建議，進入Web 2.0 傳播時代後，既有

的傳播內容分析（content analysis）方法的可行性也面臨挑戰；因而三位作者建議研究者採用「混合取徑」（hybrid approach）因應上述挑戰，亦即將大數據的「文本分析」方式結合內容分析方法，如此便能做到「保存傳統內容分析方法的系統性嚴謹與脈絡敏感度，同時也極大化資料數量與計算方法的演算精確度」（p. 34）。本研究認為，數位語藝批評也應參酌Lewis、Zamith和Hermida（2013）之建議，採取「混合取徑」；但是Lewis等人（同上）也強調「謹慎的混合」（careful combinations）之重要性。因此，針對進入Web 2.0 傳播時代的數位語藝批評研究，本研究嘗試提出適用的「混合取徑」；具體研究步驟與流程如圖一、圖二所示，文字說明如下。此外，在本文下一節，將以語藝批評方法中之「幻想主題分析」（fantasy theme analysis）為例，實際操作本文所建議的「混合取徑」，亦即如何結合文本探勘工具、人力瀏覽、人力評判的數位語藝研究方法。

（一）「文本探勘」工具分析為主階段

根據圖一所示，在主要運用「文本探勘」工具進行分析的階段，可以簡單分成七個步驟：選擇研究對象、蒐集文本、初步斷詞、自建新詞、重新斷詞、清理資料以及共詞分析。在上述七個步驟中，語藝學者必須保持適時、適度的介入文本探勘過程；尤其是當語藝學者無法自己操作「文本探勘」工具進行分析，因而必須委託資料分析公司或資料科學專家代為處理時，語藝學者如何有效地與其合作對象溝通其研究需求，往往會影響整個研究資料的效度（譬如「主題模型分析」比資料分析公司常用之「社會網絡分析」更有助於本研究幻想主題之確認）與信度。因此，本研究建議，即使不是親自操作「文本探勘」工具，在運用「文本探勘」工具進行分析的階段，語藝學者也必須了解每個分析步驟中，語藝學者需要介入判斷或決策的時間點或項目。

根據本研究作者之前的研究經驗，我們發現在運用「文本探勘」工具進行分析的階段，下列項目是語藝學者需要介入判斷或決策的時間點。首先，在選擇研究對象與蒐集文本階段時，並非所有的網路本文皆納入分析，而是仰賴研究者綜合考量下列因素：（1）電腦搜尋引擎的運算機制與限制、（2）網站貼文與回文的互動情況、（3）文本內容與研究主題的契合度，才能下達正確的指令與搜尋條件，以獲得具有

代表性的樣本。換句話說，必須注意所謂的Garbage in Garbage Out (GIGO) 的可能性；研究者除了知道自己要的是什麼（譬如如何設定關鍵字以蒐集最契合研究主題的文本），在借助資料科學技術輔助的當下，更重要的是清楚該排除些什麼。

其次，在進行斷詞與清理資料步驟時，語藝學者必須讓演算過程的黑盒子透明化，以分析工具既有的內建辭典搭配語藝學者經驗法則的檢視，自建符合研究需要的新建詞庫。也就是說，在初步斷詞時，研究者使用資料科學技術有效率地完成大量地、單一性、通則性的工作；但是在重新斷詞時，研究者需要反覆檢視斷詞的結果，介入對特殊性個案（譬如對未知詞與停用詞的判斷、複合詞的接合）的判斷，以達到最佳的斷詞結果。再者，在使用LDA（主題模型方法）的共詞分析過程中，需要研究者依據文本特性（譬如文本數量的大小，以及詞彙的多樣性）來設定TF（term frequency）和DF（document frequency）的門檻值，以選擇投入LDA分析的詞彙數量。此外，在完成主題（詞群）分類後，則需要研究者的詮釋，以完成主題（詞群）的辨識（譬如各詞群所代表的意義為何與研究主題的關聯性）、命名與編碼。

值得注意的是，爲了提升資料品質與後續語藝批評的品質，在時間以及研究資源許可情況下，研究者應該參考Denzin（1978）提出的三角檢驗方法（triangulate methods），根據研究屬性，進行不同的理論、資料、方法或研究者的三角檢驗（轉引自Johnson, Onwuegbuzie, & Turner, 2007:114；以本研究爲例，本研究主要是透過「研究者」之間的三角檢驗，亦即透過三位作者之間的相互討論，以確認相關分析結果。具體而言，除了本研究的上述七個分析步驟是三位作者共同討論的成果外，在進行個案分析時，儘管「文本探勘」工具的主要執行工作者是第三作者，但是三位作者還是共同討論與決定未知詞、停用詞與自建新詞，並來回反復檢視LDA分析結果，以避免單一語藝研究者的疏漏，影響最終研究成果。

（二）語藝學者論證、評斷爲主階段

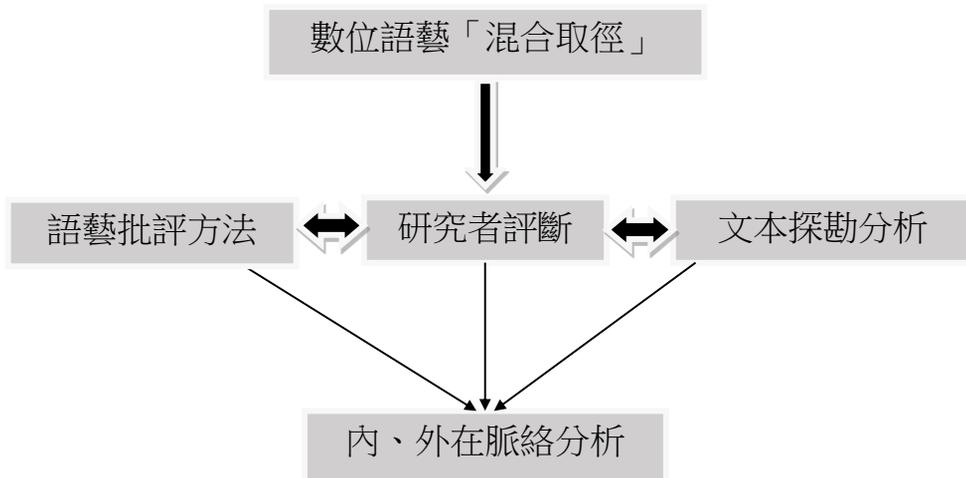
儘管資料科學技術可以協助語藝學者因應巨量網路文本帶來的研究挑戰，資料科學技術也並非萬靈丹；尤其是運用到特別重視內、外脈絡的「厚描」（thick description）的語藝批評研究，更應留意資料科學技術的侷限或量化偏見。陳譽晏

(2015)指出，想從大數據中非結構化資料中挖掘出有助於理解研究文本的特質，方法上確實有其侷限；但是他建議採用「文本探勘」工具，以利挖掘新聞或網路文本中的文本涵義。以「文本探勘」而言，上述七個分析步驟幾乎都需要語藝學者持續保持適時、適度的參與。換言之，以數位語藝批評研究而言，比較難截然劃分哪些研究階段是『人工智慧』研究階段而哪些是『工人智慧』研究階段（盧安邦，2014）。也就是說，區分『人工智慧』研究階段以及『工人智慧』研究階段似乎並沒有其必要性。

然而，根據本研究作者之前的研究經驗發現，只是靠「文本探勘」工具來進行數位語藝批評研究，確實有其限制；尤其在處理文本的內、外脈絡上。譬如在黃鈴媚、曹開明、沈錦惠、劉大華（2015.12）研究中，藉由資訊科學分析工具的協助，探討從網路少數的「活躍言者」到「反核社群」的認同擴散歷程，在分析網路聲量、社會網絡分析、詞頻分析與語意共現分析後，也只能做到分析少數「活躍言者」之論述如何被其他網民「覆誦」的過程；但無法從分析結果中進一步推論當中存在哪些「幻想主題」或「幻想類型」。儘管本研究嘗試一種新的「文本探勘」工具（LDA），確實可以協助數位語藝批評學者重構部分文本脈絡；但是最後還是必須仰賴研究者回到完整的文本中，方能論證「文本探勘」工具分析階段之研究成果。

因此，本研究認為，一個完整的數位語藝批評研究流程，必須同時包括「文本探勘」工具分析階段以及語藝學者論證階段（如圖二所示）；亦即是一個語藝批評方法結合「文本探勘」工具之「混合取徑」。換言之，數位語藝研究必然要結合文本探勘工具，以解決巨量資料蒐整的難題；但是，語藝批評本質上具有人文科學的特性，強調符號真實而非客觀真實，語藝研究者必需充分熟悉語藝情境，才能進行脈絡分析。然而，語藝批評方法也會「隨著分析文本的特性有相當大的彈性與應用空間」（林靜伶，2000），以及Foss（2004:21-22）強調評判語藝批評良窳的標準，在於是否能具備足夠的舉證（justification）、合理（reasonable inference）以及評判邏輯的一致性（coherence）。因此，在下一單元，本研究根據圖一、圖二所建議之研究流程，對11個反核臉書粉絲專頁的主文與回文進行分析，以探討本研究建議之「混合取徑」對數位語藝批評研究的可能貢獻。

圖二：數位語藝「混合取徑」操作示意圖



肆、反核臉書粉絲專頁「幻想主題分析」

一、幻想主題分析

誠如蔡鴻濱（2004）曾指出，常用的十一種語藝批評方法（包括新亞里斯多德批評、類型批評、隱喻批評、戲劇五因批評、幻想主題分析、女性主義批評、敘事批評、論辯批評、意念分析、群集分析與批判語藝），都有應用在網路文化研究的合理性。本研究選擇當中的「幻想主題分析」做為示範，主要是提出「幻想主題分析」之Bormann（2003:367）曾指出，採取「幻想主題分析」進行相關領域研究的期刊論文及書籍文章已近500篇；後續也有不少學者持續運用「幻想主題分析」進行研究（Arsenault, 2005; Bishop, 2003; Sovacool & Brossmann, 2010）。近年來陸續有學者將「幻想主題分析」用來分析網路文本，文本來源包括仇恨團體網站（Duffy, 2003）、新聞媒體網站（Greer, 2008; Perry & Roesch, 2004; Vultee, 2012）、社會支持團體網站

(Aleman, 2005; McCabe, 2009)；但是上述研究出現幾個值得探討的問題。

首先，針對數位語藝的「互動性」與「互文性」特質 (Warnick, 2007)，其中只有少數研究 (e.g., Greer, 2008; Vultee, 2012) 的研究文本反應數位語藝的「互動性」特質。其次，即使只是處理貼文 (或主文) 而不包括回文，也只挑選少數貼文進行分析；譬如Duffy (2003) 從四個仇恨團體網站各自選擇三篇文本進行分析。此外，當研究者 (Greer, 2008) 同時分析主文與回文時，研究者只能選擇極為少數的主文 (譬如Greer選擇10%的主文) 以利於後續的回文分析。換言之，Web 2.0時代的數位語藝產製大量的網路文本，對於既有的「幻想主題分析」方法形成很大的挑戰，誠如Duffy (2003: 295) 指出以上述方式，即使只是研究單一網站，分析整個網站的文本對幻想主題分析方法而言仍是不可行的。

面對Web 2.0時代對「幻想主題分析」方法帶來的挑戰，Vultee (2012) 採用「扎根理論 (Glaser, 1978) 回應。具體而言，研究者從2009年七月開始分析福斯新聞網上關於美國歐巴馬總統新聞報導以及讀者回應，透過不斷閱讀與比較，研究者登錄「幻想主題」，接著將類似「幻想主題」整併成「幻想類型」；上述工作一直持續到2009年結束，因為已經出現「扎根理論」所謂的「理論飽和度 (theoretical saturation) (Glaser & Strauss, 1967)，亦即不再出現新的「幻想主題」。本研究認為，當網路文本來自於單一網站時，Vultee (2012) 的做法具備參考價值；但是假如研究主題無法只是分析來自於單一網站的文本 (譬如本研究分析反核臉書粉絲專頁) 時，Vultee (2012) 的做法就會發生Duffy (2003) 所指出的「可行性」問題。因此，結合資料科學運算技術成為克服上述「可行性」問題的一種可能做法。

相對地，國內學者 (趙雅麗, 2001; 陳佳靖, 2002; 蔡鴻濱, 2004; 張裕亮、陳奕任, 2008; 秦琍琍, 2010) 也持續採用「幻想主題分析」作為語藝批評方法 (Arsenault, 2005; Bishop, 2003; Chen, 2007; Sovacool & Brossmann, 2010)。國內研究中，有兩篇針對網路社群進行研究 (陳佳靖, 2002; 張裕亮、陳奕任, 2008)。其中，陳佳靖 (2002) 選擇特定BBS站 (「花魁藝色館」) 情色討論區，作為研究網路情色場域；張裕亮、陳奕任 (2008) 則是針對中國大陸BBS站 (「天涯社區」的「關天茶舍」版) 作為研究場域，探究大陸網民在此論壇中共同形塑對日認知。這兩篇論文均未明確說明所取樣的文本量有多少，但可確知的是主要是透過研究者人力瀏覽、

歸納出的幻想主題，藉以掌握文本內在脈絡，藉以推論出其「覆誦」情形。

簡言之，1972年Ernest Bormann提出「符號融合理論」，預設溝通創造真實與成員共享此一真實；人們使用符號互動與詮釋的過程，猶如是種戲劇形式，藉以分享彼此的情感、動機與意義，進而產生共享的真實。用以檢視這個溝通創造真實的戲劇形式與內涵的語藝研究方法，則稱為「幻想主題分析」（Fantasy theme analysis）。此一理論與研究方法，已在人文科學與口語傳播學門中，獲得確認並持續援引作為研究工具（Bormann,1972, 1980a,1982b, 1982c, 1983,1985a, 1985b, 1990; Shields & Preston,1985; Cragon & Shield, 1992）。由於「符號融合理論」重點在於強調所謂的「真實」（reality）是經由符號所創造，因此人們行為的意義、情感及動機，會在言說脈絡中明顯表露。

Bormann（1983:73）也曾指出，若是愈多社群成員參與這場說故事的劇情中，結果是產製出「具爆炸性的溝通情節」（explosive communication episode），並透過符號串聯整合社群，依循共同經驗產製出符號融合的作用。

因此，Foss（2004:109）歸納出「符號融合理論」兩項研究前提，一是強調人們藉由溝通創造真實；二是個人不僅採取符號創造真實，個體藉由符號融合、創造出社群共享的意義。至於所謂的「幻想主題」則是人們在詮釋事件的過程中，所使用的一個字、一句話或是一種說明方式，重點在於結合群體經驗建構出一套故事，用以創造真實。「幻想主題」主要透過描述「人物」、「行動」及「場景」的方式來形成，並透過「覆誦」作用形塑出「語藝視野」（rhetoric vision）。

Shields（1981:6-7）則指出「幻想主題」具有四個戲劇要素，即「角色」、「情節」（plotline，意指「行動」）、「場景」以及「合法化機制」（sanctioning agent）。藉由對照Foss及Shields所提的「幻想主題」所具備的戲劇要素，其實兩者大致相符；但Shields特別標示出「合法化機制」為幻想主題的要素，在於強調在語藝行動的劇碼中，為達到說服目的，溝通傳播內容涉及到須具有正當性或合法化的來源。由於本研究探討的是核能議題的公共溝通情形，相關文本如何形塑「合法化機制」，也是值得關切的問題，故應加以探討。基此，參照Foss及Shields所提「幻想主題」的要素，可分別再加說明如下：

（一）「角色」主題：意指在語藝劇碼中，出現的人物被賦予的人格特質，例如

在文本中什麼人被描述為英雄、惡徒或支持者，藉以指稱其言行動機。

(二)「行動」主題：涉及在語藝劇碼中所發生的行動，也就是何者對誰做了什麼事以及如何達成的過程，進而也形成整個「語藝視野」的戲劇樣態（主動的或是被動負向的、歡喜收場還是一場悲劇）。

(三)「場景」主題：主要是在文本中所標示語藝事件發生的所在地。而且重點不僅在於指出地點而已，而是在於彰顯發生地點的特徵以及意涵。

(四)「合法化機制」主題：則是指在語藝劇碼中，論點可被接受的合理化理由。例如有時文本強調具有較高位階的權力依據（像是指稱上帝、正義、民主等）；另外，則是當下所面臨的特殊情境（如冷戰時的核子攻擊威脅、戰爭衝突、宗教的最後審判等）。也因此，「合法化機制」是語藝劇碼「幻想主題」中，最為顯著明確的行動動機，藉以強調為何行動的原由。

二、「文本探勘」工具分析為主階段

(一) 選擇研究對象、蒐集文本

由於本研究選擇語藝批評方法中的「幻想主題分析」，而此一方法特別適合用來探討從網路少數的「活躍言者」到群眾的認同擴散歷程；2014年4月28日「政府宣佈核四封存記者會」則是台灣反核運動的關鍵事件⁴，因而讓本研究作者好奇：我們是否有可能從分析反核臉書粉絲專頁在上述關鍵事件發生前後的論述，看到網路少數的反核「活躍言者」的論述如何被其他網民「覆誦」，最終形成一個共享「反核真實」的「象徵社群」的認同擴散歷程？決定好研究對象後，本研究碰到第一個難題：因為研

4 2014年4月28日，時任行政院長江宜樺召開記者會表宣佈重大決議，第一，「核四一號機不施工、只安檢，安檢後封存；核四二號機全部停工」，第二，「儘速召開全國能源會議，以確保未來供電無虞」。行政院（2014年4月28日）。「江揆：核四停工並非停建，為下一代保留一個選擇權，並儘速召開全國能源會議」。檢索時間：2015.7.20。網址：http://www.ey.gov.tw/News_Content.aspx?n=F8BAEBE9491FC830&s=34040F31D4C717E6

究並非在上述事件發生當下進行，無法即時儲存資料，如何蒐集網路文本成爲一個挑戰。

因此，本研究在比較坊間可以合作的對象（譬如IBM、網路溫度計及意藍科技）後，決定委託意藍科技公司以其網路爬文技術（web crawler），協助我們取得網路文本的原始資料（raw data）；因爲意藍科技採取的文本探勘工具及雲端儲存功能，已逾兩年以上期間全面蒐整及儲存網路使用者的各類文本記錄，諸如公開的FB貼文及回文、PTT及網路新聞等，足供歷時性的研究所需。簡言之，我們根據網路聲量數據決定資料蒐集時間爲2014年3月28日至4月28日；接著以「核能發電的立場」、「成立的時間」和「按讚人數」爲篩選條件，選擇具有代表性的社群媒體樣本，共計選取11個標榜反核的臉書粉絲專頁⁵作爲分析對象（如表二）。針對文本蒐集，也是委由「意藍科技」公司以「核能」、「核電」、「核四」、「反核」、「廢核」等關鍵字搜尋，在選定的11個標榜反核的臉書群組中共擷取了3424則貼文，6443則回文。

（二）關鍵詞抽取：初步斷詞、自建新詞、重新斷詞、清理資料

在意藍科技公司完成蒐集網路文本的原始資料步驟後，根據本研究作者之前的研究經驗，我們決定由本文第三作者接手完成後續「文本探勘」步驟；因爲唯有如此，研究者才能在後續的幾個關鍵步驟中，保持適時、適度的介入。具體而言，本文第三作者在熟悉資料科學技術學者協助下，根據本研究研究主題的需要，採取R軟體的Rwordseg分析套件（package）⁶，進行關鍵詞與分類。換言之，詞語是語言意義的最小O單位，因而中文文本探勘首先須對蒐集的文本進行斷詞（word segmentation），也就是分析文本上每個出現的詞語；但是由於蒐集的文本中包含許多未被收錄到詞典中的未知詞（unknown words）與複合詞（compound words），因而必須進行以下程序，以便順利斷出重要的詞語。

1. 初步斷詞：利用R軟體3.3.2版，安裝中文斷詞套件JiebaR，以其內建的詞庫進

5 分析資料來自本文第二作者執行的科技部研究案：「台灣能源政策溝通與建構核能共識：氣候變遷因應與核能安全之論述策略分析」中，第一年計畫的部分研究成果。

表二：反核臉書研究樣本

反核臉書粉絲專頁					
項次	臉書群組	成立時間	按讚人數	貼文數	回文數
1	反核部隊	2012.03.12	21,538	870	724
2	全國核廢行動平台	2013.01.20	100,166	340	3336
3	我是人，我反核！	<u>2012.05.27</u>	35,299	58	332
4	媽媽監督核電廠聯盟	2013.03.08	59,431	102	1437
5	我愛台灣反核行動聯盟	<u>2011.06.09</u>	2,586	248	43
6	爸爸非核陣線	2013.08.07	3,790	25	36
7	終結核四催票大聯盟	<u>2013.07.26</u>	7,170	75	35
8	地球公民基金會	<u>2011.03.16</u>	83,687	15	380
9	綠色公民行動聯盟	<u>2010.11.04</u>	113,693	102	973
10	綠色和平Greenpeace	<u>2010.09.07</u>	425,135	9	337
11	台灣環境保護聯盟	<u>2012.07.17</u>	17,765	63	31

註：資料收集時間為2015年7月28日

行初步的斷詞，貼文計得出23766詞彙，回文得出15764。

2. 自建新詞：將初步斷詞的結果，統計任意兩個前後緊連詞語（bigrams）的出現次數，根據出現次數、詞語字數及出現文本數以及語義判斷上的經驗法則等訊息，找出較具語意的未知詞和複合長詞，將其新增到JiebaR套件的自建詞庫。以本研究為

6 R軟體是1970年代AT&T貝爾實驗室建構而成的免費科學與統計軟體，包括資料處理、統計分析、模擬、科學運算與圖形功能。R軟體目前為開放原始碼的自由軟體（open-source software），並已逐漸普及應用在學術研究（陳景祥，2010）。為了掌握關鍵字的分類與計算詞頻，本研究使用R軟體中的Rwordseg分析套件，Rwordseg是R環境下專門用來進行中文斷詞的工具（Jian Li，2013），可在R網站免費下載https://r-forge.r-project.org/R/?group_id=1054。

例，透過上述過程，得出許多與研究主題相關的新詞彙，譬如「林義」+「義雄」。

3. 重新斷詞：反覆上述斷詞與自建新詞的程序，直到新增的新詞數量趨於飽合。以本研究為例，前後來回斷詞4次後，最終共計得出貼文23896個詞，回文則有15894個詞。

4. 清理資料：經過上述反覆斷詞程序後，為取得較佳的分類結果，避免過多詞彙造成分類的複雜，以及兼顧電腦運算效益考量，必須進一步過濾出重要詞彙。本研究結合詞頻分析與「詞頻-逆向文件頻率」（TF-IDF, Term Frequency-Inverse Document Frequency）⁸，以避免高估或是低估關鍵字與語意共現情形，協助研究者聚焦在重要的關鍵詞上（林頌堅，2010；陳世榮，2015）。具體而言，本研究依據新建詞庫中每個出現詞語的出現次數、詞長、出現文本數，篩選其中出現次數大於20次，詞長為2個字以上，以及出現文本數大於10篇的詞語，移除英文字、數字與停止詞（stopword）後，最終共得到貼文計2893個詞，回文1702個詞語，列為具備文本意義的重要關鍵字。

（三）共詞分析 (Co-words Analysis)：「主題模型方法」之應用

儘管Lewis、Zamith與Hermida（2013:38）曾指出，運用資料科學工具進行文本探勘以及分析，最主要的做法是找出文本內高頻次出現的關鍵字，同時探究上述高頻次關鍵字之「語意共現」（co-occurrence）特徵，以便歸結出網路文本的論述本質。然

7 依據劉大華等（2016）的研究指出，以bigrams方法加上人工檢視的相互參照，有助於發現較具語意的未知詞和複合長詞，例如「林義」+「義雄」、「非核」、「家園」等。另一項優點是區分不同複合詞在語義上的歧異，例如「輻射汙染」與「輻射劑量」。

8 TF-IDF ($TF-IDF = tf \times idf$) 是一種篩選重要關鍵詞的指標，用以評估一個字詞對於一篇文件或在一個語料庫中的其中一份文件的重要程度。字詞的重要性隨著它在文件中出現的次數 (TF, Term Frequency) 成正比增加，但同時隨著它在語料庫中出現頻率 (DF, Document Frequency) 成反比下降。可修正單靠詞頻分析可能誤判「高頻次但不重要的詞」（如但是、一起、因為…等），或是「低頻次但深具意義的詞」（如重要的地名、人名）。

而，本研究認為，頻次固然有其意義；但是假如研究者想進一步辨識大規模文件群集（document collection）或語料庫（corpus）中潛藏的主題訊息，那就需要採用不同的文本探勘工具。因此，本研究在完成清理資料步驟後，經由長期協助本研究的資料科學技術學者建議，選擇「主題模型方法」（Topic Modeling）進行接下來的文本分析；因為「主題模型方法」是協助研究者從文件群集中抽取隱藏「主題結構」（thematic structures）的技術方法，一般被應用來抽取一篇文章或一段話所表達的「中心思想」，而這也正是「幻想主題分析」研究之目的。

簡言之，「主題模型方法」主要使用LDA（latent Dirichlet allocation）統計模型描述文件（document）產生的過程，揭露文件集中可能出現的各個主題（topic），以及這些主題出現在每一筆文件上的機率。這個模型假設1. 每一筆文件都包含多個主題，可以利用機率混合（probabilistic mixture）來表示文件內的主題組成比例。2. 每一個主題都是由一組詞彙（vocabulary）上的詞語依據不同的機率組成（Blei, Ng, & Jordan, 2003；林頌堅，2014）。依據上述的定義，本研究認為應用LDA協助本研究進行文本探勘有兩個優點：（1）以主題呈現潛藏在文本的共詞（co-words）特徵（依各個詞彙彼此間共同出現的機率，加以類聚與分類成為各個詞群／主題），所歸類出的主題間並不互斥，因而更可以看出頻次以外的共詞特性；並依據主題彼此間詞語重疊程度，以呈現出主題的強度特性。（2）由於LDA同時計算每一筆文件中各個主題的分佈機率，研究者可針對特定主題，回到各則文件中查找代表性文本（各文件出現機率最高的主題），有助於質性的文本檢視。

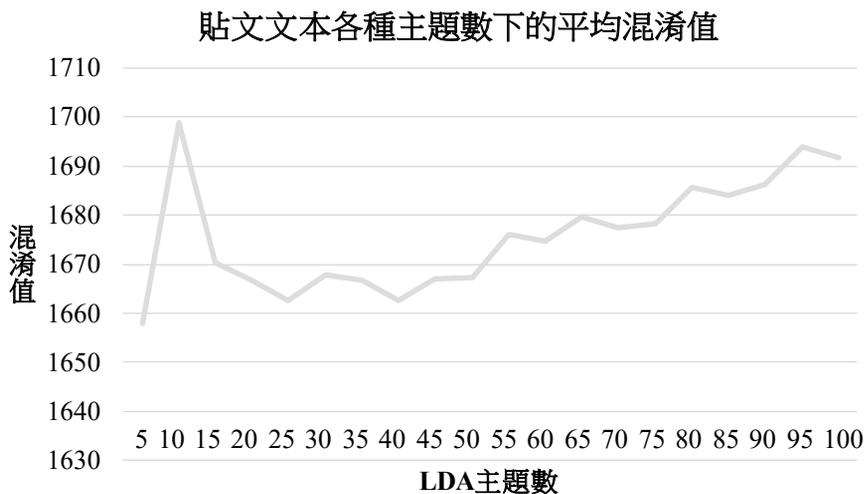
在實際操作上，本研究先將11個臉書的文本分成貼文及回文兩個部份，然後各別投入代表其文本的關鍵字（貼文2893個詞，回文1702個詞）進行分析。所投入的關鍵詞語，以文件—詞語矩陣（document-term-matrix）輸入Gibbs 演算法推導文件特徵向量與主題特徵向量等模型參數。另外採用機器學習（machine learning）與資料探勘（data mining）常用的k次交叉驗證（k-fold cross validation）方法，以10次交叉驗證估計混淆度（perplexity），決定主題模型方法最佳的主題的數目。經過演算法的估計，貼文的部份以5個主題數的混淆度最低，表示該主題數的模型具較佳的分類結果（混淆度平均值如圖三所示）。經過相同的驗證，回文的部份以15個主題數的混淆度最低（混淆度平均值如圖四所示）。換言之，經由上述之LDA共詞分析步驟，本研究

得出貼文15個主題及回文5個主題的共詞特徵（如表三及表四），作為本研究作者進行後續歸類、命名其中人物、場景、行動及合法化機制等「幻想主題」的依據。

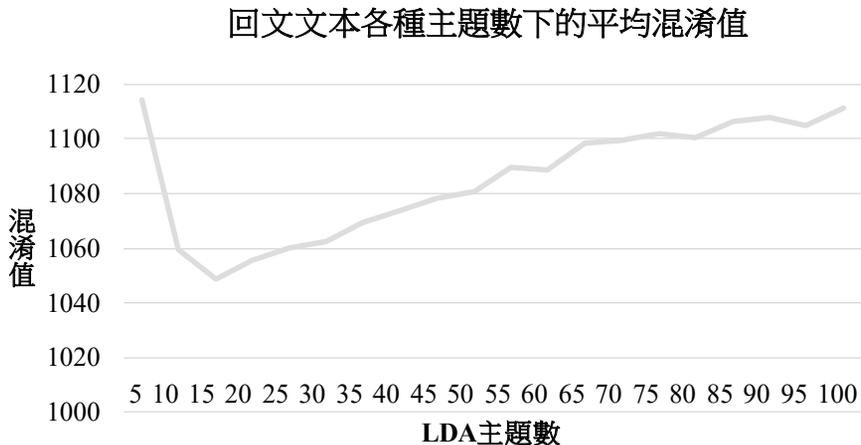
透過LDA獲得共詞分析結果後，本研究認為，此時就需要研究者對於所研究的語藝事件有足夠的掌握，方能評判重要關鍵字；同時研究者也要用人力檢視、對照反核臉書貼文與回文使用詞語，是否具有共通情形，以確認存在「覆誦」主題。簡言之，本研究在走完以「文本探勘」工具分析為主階段後，進入語藝學者論證、評斷？為主階段，在這個階段中，本研究分成兩個步驟進行「幻想主題分析」：研究者確認「覆誦」關鍵字及進行文本舉證、研究者歸類幻想主題，分別說明如下。

誦」關鍵字及進行文本舉證、研究者歸類幻想主題，分別說明如下。

圖三：反核臉書貼文共詞分析主題數混淆度



圖四：反核臉書回文共詞分析主題數混淆值



三、語藝學者論證、評斷為主階段

(一) 研究者確認「覆誦」關鍵字及進行文本舉證

根據LDA的共詞分析結果，本研究發現，貼文Topic4（表三）所呈現的共詞特性，可推知其中隱藏號召公民參與反核遊行的「行動主題」，因為「反核」、「遊行」、「廢核」與「行動」出現高頻次的語意共現情形。更值得注意的是，在此次反核運動中，林義雄先生絕食抗議的苦行，這種不惜犧牲生命健康來反對核四廠的精神，在當時引起新聞媒體大幅報導與大眾的注意，顯見「林義雄先生」或「林義雄」一詞更具有行動召喚作用。因此，「林義雄先生」或「林義雄」一詞雖非此一主題內最高頻次的詞語；但卻是研究者必須關注的關鍵字，此種分析敏感度必須來自研究者對此一語藝事件的掌握，實屬文本探勘工具無法代勞之處。事實上，「林義雄先生」或「林義雄」一詞在回文Topic3、Topic7、Topic14、Topic15中也被「覆誦」。凸顯反核臉書中的活躍言者為促使非核家園，因此要進行反核遊行，強調以停建核四的重要性，經由重要關鍵字—「林義雄先生」或「林義雄」的「覆誦」後，更具有合理性與

必要性，以召喚更多的人們參與這項行動。

表三：11個反核臉書貼文主題

貼文共詞分析主題	
Topic1	<u>台灣</u> ，核四， <u>核電</u> ， <u>政府</u> ，能源， <u>反核</u> ，核能，人民，台電， <u>廢核</u> ，國家，公投，民眾，發電，電力， <u>服貿</u> ， <u>再生能源</u> ，現在， <u>發展</u> ， <u>成本</u>
Topic2	<u>台灣</u> ，森林，道路，空氣污染，手術，室外，研究，環境，增加，空氣，記者，分享，暴露，新聞，大夫，有關，監控，中國，廢棄物，告白
Topic3	輻射， <u>日本</u> ， <u>福島</u> ，核電廠，核電，發生，台灣，地震， <u>核災</u> ，美國，造成，輻射污染，影響，核反應爐，劑量，研究， <u>福島核災</u> ，核廢料，核能，安全
Topic4	<u>反核</u> ， <u>遊行</u> ，台灣， <u>廢核</u> ， <u>行動</u> ，活動，一起，停建核四，現場，廣場， <u>林義雄先生</u> ，凱道，大家，下午，核四，非核家園，立法院，公民，地點，參與
Topic5	省電， <u>台灣</u> ，台電，美國，隨手省電，核電廠，技術，實現非核家園，節能的， <u>福島</u> ， <u>核災</u> ，媽媽監督核電廠聯盟，發電裝置容量，節能，提升， <u>日本</u> ，廢除核四，謝謝，核四廠

註：主題內詞語是由左上第一個詞語為出現頻次最高者依序由左至右向下排列

表四：11個反核臉書回文主題

回文共詞分析主題	
Topic1	核四,總統,億餘元,興建,行政院,陳水扁,預算,台電,億元,指示,復工,立法院,國民黨,行政院長,核能,計畫,核四工程,損失,李登輝,時任
Topic2	拒絕,核四,叫做,民進黨, <u>台灣</u> ,英文,現在,反對,億元,經濟, <u>服貿</u> ,人民,團體,人家,馬英九,政府,投資,警察,人才,執政
Topic3	核四,公投,國民黨,停工,停建核四, <u>台灣</u> ,安檢,運轉,停建,反核,政府,馬英九, <u>林義雄</u> ,通過,人民,支持,民主,燃料棒,全民,簡單
Topic4	反核,警察,政府,我反核,大家,支持,人民,抗議,行動,民眾,出來, <u>台灣</u> ,看到,我是人,影響,擁核,別人,抗爭,訴求,需要

Topic5	反核,遊行,廢核,晚會,台灣,現場,台北,連署,加油,萬人,網路,大家,核四,高雄,凱道,公民,今天,直播,看到,台北場
Topic6	台電,台灣,核電,發電,核四,電廠,核能,美國,核安,機組,系統,燃料,容量,安全,核電廠,接受,資訊,電價,人民,便宜
Topic7	反核,民進黨,核四,執政,國民黨,現在,林義雄,廢核,大家,選舉,政治,立委,停建核四,人民,支持,政黨,立法院,市長,總統,一起
Topic8	台灣,日本,大陸,大陸人,優勢,貿易,免稅,商品,民進黨,經濟,成本,利用,大家,中國,美元,各國,自己的,台灣人,學生,多少個
Topic9	台灣,核電,反核,廢核,發電,政府,核能,核四,核廢料,核電廠,用電,太陽能,大家,安全,能源,德國,現在,電費,火力發電,電力
Topic10	核能,核四運轉,不必,出事,日本,早就,不該,國家,正當,說不定,無處,事實,俄羅斯,懲罰,理念,支持核四,支持,萬一,全台,意志
Topic11	反核四,關閉,政客,建立,政治,核四廠,司法,黃牛,先生,分享,核一廠,全面,核二廠,爺爺,無核,國度,家園,核能發電廠,操作,無核家園
Topic12	台灣,人民,加油,臺灣,反核,國家,大家,政府,台灣人,核災,發生,核四,台灣人民,中共,國民黨,現在,一起,土地,希望,站出來
Topic13	近親繁殖,馬英九,違法,違憲,入學,關係,媒體,建中,對號入座,秦始皇,犬馬,黑道,太監,總管,國民黨,無恥,亂象,國父,果然,可恥
Topic14	台灣,林義雄先生,核四,停建核四,聯盟,革命,總統,馬英九,社長,林義雄,台北,生命,台灣人民,民主,立法院,拿錢,出來,希望,禁食,抗爭
Topic15	反核,竹林,台電,回應,危害,民意,生命財產,鄉親,電纜,高壓,林義雄,核四,抗議,擁核,綠色,政客,聯署,探視,縣市長,反對

註：主題內詞語是由左上第一個詞語為出現頻次最高者左至右向下依序排列

研究者確認在貼文及回文的特定詞群中出現的重要關鍵字「林義雄先生」或「林義雄」後，本研究認為，上一階段的LDA共詞分析結果，也提供研究者能夠找到各個詞群主題曾經出現的全文文本之路徑，讓研究者可以透過人力瀏覽方式舉證出含有「林義雄先生」或「林義雄」此一關鍵字的貼文或回文之論述特徵，以協助研究者得以重構比較完整的文本內在脈絡。例如在回文Topic 15中編號3490及編號3491的網友在回應貼文標題為「面對多數的廢核民意，遍地烽火的行動，執政黨依然故我、冥頑不靈」的文章時，分別回文：

我們所敬愛的林義雄先生，為了解決核四紛爭、落實民主體制，禁食已經進入第四天，身體虛弱，新陳代謝開始進入衰竭，危及生命。但馬英九仍然一意孤行，不願面對停建核四的廣大民意。今天台灣人民站出來，

我們要守護台灣、守護林義雄。林義雄先生以禁食、禁語的方式，呼籲政府停建核四，我們也將採取非暴力抗爭的方式，要求政府停建核四，我們要求政府要立即對於林義雄先生及多數台灣人民的訴求積極處理。我們在此宣告，如果因政府的麻木不仁，致使林義雄先生發生不測，我們將放棄非暴力抗爭的方式，化悲憤為革命的行動，採取各種積極有效的抗爭手段，讓馬英九為此付出代價，直到馬政府傾聽民意、停建核四。（編號3490）

國民黨殺了民主先驅林義雄的母親13刀斃命，兩個讀幼兒園的雙胞胎小女兒各一刀斃命，台灣沒有林義雄等民進黨民主先驅犧牲性命台灣人哪來的民主自由??恐怕到現在還是蔣家王朝，台灣欠林義雄一家太多了，是時候發動流血革命將中國國民黨和這些吸台灣血的高級外省人趕回中國了!!!!!!!林義雄倒下革命的號角就響起。（編號3491）

Topic 15中編號4364的網友在回應主文標題為「林義雄先生發出聲明，宣布停止禁食行動」的文章時，則是回文表示：

終於可以不用再擔心著了!謝謝林義雄先生。（編號4364）

簡言之，在LDA詞群分析結果的基礎上，還是需要仰賴研究者對所研究的語藝事件的了解，才能在決定形塑幻想主題的關鍵字時，不受限於詞頻；然而，LDA的特性，也協助研究者可以根據所選定的關鍵字，回溯、審視完整文本，讓研究者能夠從實際由文本的內在脈絡中，確認經由覆誦「林義雄先生」或「林義雄」此一關鍵字所形塑的幻想主題，主要是出於敬重當時進行絕食抗議的林義雄，強調追求正義，透過「動之以情」的方式形塑相互認同。

其次，從貼文Topic1（表三）呈現的共詞特性，在「核電」、「政府」、「反核」、「廢核」高頻次出現情形下，「服貿」、「發展」、「再生能源」以及「成本」在此一主題詞群中則是別具意義。事實上，在2014年4月一連串的反核遊行前，台灣社會才因為「太陽花反服貿運動」（2014年3月18日至4月10日），出現「反傾中」與「反經濟掛帥」的公民籲求。故在反核臉書貼文中，則是強調反對發展核能作

為發展兩岸經貿合作的基礎電力來源，並要求政府擴大使用再生能源，以效法先進國家乾淨能源的環保作為，能源政策不應只為了服膺兩岸統合的經濟需求！對照回文Topic9（表四）出現「太陽能」（再生能源的一種），以及Topic2（表四）出現「服貿」關鍵字以及Topic8（表四）出現「成本」、「大陸」、「貿易」等關鍵字覆誦下，可以推知是建基在太陽花運動有效阻擋服貿協議在立法院闖關的成功效用下，強調反核即是「反傾中」經濟，與其經濟掛帥推展核能，僅是成為傾中助力，更願意選擇推展再生能源，避免損害台灣自主與普羅大眾的權益。

為論證LDA分析結果所呈現之主題詞群特徵的意義，本研究依據貼文Topic1（表三），回溯言者貼文，例如編號34號及編號151號的臉書貼文分別表示：

服貿如嫁女？不嫁活不下去！他還說：核四不運轉會缺電。有這種經濟部長，真的覺得台灣人很倒霉！（編號34）

三月十八日晚上，上百反對兩岸服務業貿易協議的青年與社會公民攻破了立法院議場的大門，開始台灣史無前例的「佔領國會」行動，從這個晚上起，議場外陸續聚集了上萬的聲援民眾，一同守護這場行動，持續訴求「退服貿，守護民主」。綠色公民行動聯盟，長期推動核電與能源政策改革，以及其他環境、資源治理問題。看似和這次兩岸服貿協議爭議無直接關聯，但如果我們細看當前反核運動的處境，就會發現其中共同面對的政治與經濟結構壓迫是多麼的相似。...（編號151）

也正由於反核與反服貿已形成政治意義上共通性，在經由反核社群間相互覆誦歷程下，形成對執政階級對抗的宣示及共同立場。在覆誦「反核」及「反服貿」的歷程中，相對也就形成透過「反服貿」的成功社會運動經驗，將「反核」與「反傾中」與反統治階層壓迫相互扣連，形成一種「合法化機制」的幻想主題，形塑認同。例如本研究依據回文Topic2（表四）回溯相關文本，編號2852網友在回應貼文標題為「江揆：停工不代表停建 不會廢棄核四！」時則說：

您是有思想的台灣人民!~您了解核四拼裝一號機二號機的內幕嗎?您放心

您的家人一直受到核災的威脅嗎?您信任核廢料的處理能力嗎?您認為抗爭是為選票嗎?您理解現在民主時代誰逼人民站出來?您知道反服貿是爭民主自由尊嚴敲破賣台的黑箱嗎?您相信停建核四是為了家人的生命孩子的未來嗎?您可以去感受一下那種風吹雨淋的辛酸嗎?您想想今日可以過這樣的日子是多少搗蛋人民笨出來的?是人終究自私的~真的建議您多了解台電到底在騙人民甚麼?都是鄉愿害了台灣幾十年了!(編號2852)

上述發現(亦即在反核事件中出現「服貿」及「台灣人民」等關鍵字),某種程度說明,網路社群中使用的特定重要關鍵字在經由「覆誦」歷程中,確實會出現「意義展延」(Burke, 1937: 230)情形。因此,透過文本探勘工具的語意共現分析,確能有效地釐清具有「意義展延」的關鍵字,並可舉證相關文本,以利論證。也因此,面對大量文本時,若是只採行人力瀏覽、編碼方式,恐無法有效論證社群成員共享與共創幻想主題的特徵。

其三,從貼文Topic3(表三)及Topic5(表三)所呈現的共詞特性,「日本」、「福島」、「核災」以及「福島核災」成爲重要關鍵字,無疑是在召喚社群成員記取日本核災殷鑑,凸顯台灣無法承受如此災變;相對地,「日本」一詞在回文詞群Topic8(表四)及Topic10(表四)被覆誦日本福島核災後,強調的是「場景主題」,凸顯日本核災殷鑑不遠,台灣地小人稠更是無法承擔核災風險,要求政治人物不要再欺瞞百姓。依據Topic8(表四)回溯相關網友文本,例如編號928網友表示:

星雲支持建核四,吳敦義:確保核安才運轉」星雲25日在佛光山會見吳敦義和楊秋興,表示支持興建核四的理由竟是因為天災難免!這邏輯很奇?天災就是因為躲不掉,核災又具毀滅性才要廢核,希望您早日得"道"?核安是騙人的,沒絕對的,日本核災是故意蓋不良核電廠嗎?支持「非核家園」吧?
(編號928)

第四,從表三貼文Topic1、Topic2及Topic5所呈現的共詞特性,「台灣」一詞是重要關鍵字,「台灣」一詞在表4回文Topic2、Topic4、Topic6、Topic8、Topic9、Topic12、Topic14中被覆誦,尤其在Topic12及Topic14中,出現「台灣」與「台灣人

民」的共詞情形，相關反核、省電及環保節能等概念，均是在召喚台灣人主體意識，要求一起站出來抗爭。例如依據回文Topic14（表四）回溯相關網友文本，其中編號6239網友表示：

是受傷的腳讓步調緩慢才剛走到凱道就開始下起大雨現場大家有秩序的就地而坐肅靜的氣氛是尊敬林義雄先生！他願意用自己的生命來保護台灣母親他願意用自己的生命來捍衛台灣兒女當他的生命隨著時間一點一滴流逝看著滂沱大雨內心無助吶喊著！身為台灣人民的我們能夠為台灣做什麼！？我們又為台灣做了什麼！？明天就讓我們再次上凱道！讓92%民意下令8%馬英九！立即給我停建核四！（編號6239）

（二）研究者歸類幻想主題

因此，本研究認為經由LDA的共詞分析，可以看出在反核臉書社群間的貼文與回文中互動及互文歷程，確認已被覆誦的相關重要關鍵詞語；並透過文本探勘工具回溯詞群文本的作用，提供研究者以人工瀏覽、評斷以及舉證文本，加以類歸提出11個反核臉書社群間的幻想主題如下：

人物主題：反核是展現台灣人主體意識

場景主題：日本福島核災殷鑑不遠

行動主題：支持林義雄先生的正義遊行

合法化機制：反核即是反傾中、反壓迫

由上列幻想主題可知，反核臉書社群成員已將反核議題提升至政治、兩岸經貿合作以及台灣主體意識的層次，並非僅是對於核安問題的討論。經由本研究透過「混合取徑」方式歸類出上列的幻想主題，可發現反核臉書社群共享與共創出諸如反核即是反傾中、反壓迫及展現台灣人主體意識的論述特徵，凸顯反核議題已然政治化，並與國族認同與大陸威脅等認知相互扣連；用以抵制核能發電取得價廉電力以促進經濟的

論述策略。此外，誠如林靜伶（2000：81）強調，Bormann的「幻想類型」「是由一連串共享的、相關的幻想主題所構成」；但是研究者需要結合外在脈絡分析，進行後續詮釋與分析，才能進一步推論出這11個反核臉書頻道中出現的共同「幻想類型」，並且進而提出反核臉書社群成員「語藝視野」的樣態。

伍、討論與建議

針對進入Web 2.0 傳播時代的數位語藝批評研究，本研究提問：研究者如何善用「文本探勘」工具，以便能夠同時擴大文本處理數量並兼顧語藝批評的詮釋與批評精神？也就是說，將「文本探勘」工具運用到數位語藝批評研究上，除了借重其有效處理大量網路文本的長處外，也必須了解其限制，同時思考在不同的研究階段，研究者如何將其優點發揮到最大而又如何補強其不足之處？回應上述問題，本研究探討語藝批評方法結合「文字探勘」工具時，如何在不同研究階段的網路文本選擇、分析以及詮釋上，透過研究者的適時與適度介入，逐步重構文本脈絡，以提供研究者有效的論證或詮釋所需要的依據資料。

簡言之，本研究建議語藝學者採取語藝批評方法揉合「文本探勘」工具的「混合取徑」。之所以選擇「揉合」而非「結合」，主要是本研究強調，即使是在「文本探勘」工具分析為主的七個步驟中，各個步驟都需要語藝學者持續保持適時、適度的參與；相對地，在進入語藝學者論證、評判為主階段後，研究者也需要在上一階段LDA研究成果的基礎上，進行主題(詞群)辨識、命名與編碼工作。也就是說，以數位語藝批評研究而言，比較難截然劃分哪些研究階段是盧安邦（2014）的『人工智慧』研究階段而哪些是『工人智慧』研究階段。因此，本研究認為，對數位語藝批評研究而言，宜避免以「非黑即白」的態度來區分『人工智慧』研究階段以及『工人智慧』研究階段。

為了論證上述建議之可行性，本研究選擇常用的十一種語藝批評方法中之「幻想主題分析」，分析11個反核臉書粉絲專頁的3424則貼文與6443則回文。在「揉合」「主題模型方法」之LDA分析後，本研究發現，「主題模型方法」確實有助於研究者從文件群集中抽取隱藏的「主題」。簡言之，即使在LDA主題（詞群）分析結果的基

礎上，還是需要仰賴研究者對所研究的語藝事件的了解，才能在決定形塑幻想主題的關鍵字時，不受限於詞頻；然而，LDA的特性，也協助研究者可以根據所選定的關鍵字，回溯、審視完整文本，讓研究者得以重構比較完整的文本內在脈絡。因此，對於面對Web 2.0傳播時代挑戰的數位語藝批評研究，本研究所建議的「混合取徑」提供一個可行的研究取徑。當然，本研究只是一個新的嚐試，後續研究還有很多可以開拓與精進的空間。

首先，爲了進一步理解數位語藝的「互文性」特徵，可嘗試其他不同的詞語共現分析方法，例如運用文本探勘的「點對點交互資訊」（PMI, Pointwise Mutual Information）和「詞彙共現網路」（co-occurrence network）；所謂「詞彙共現網路」是以各個詞彙爲網絡中的節點（node），而共同出現在同一篇文章的詞彙以「邊」（edge）來聯結，繪製詞彙共現網絡圖，再依據網絡中心性（centrality）的指標解各節點（詞彙）在整個網絡中的重要程度（施旭峰等，2014）。上述分析工具可以確認詞組社會網絡，再加進行詞群分析以及代表性樣本的人工瀏覽，或也有助於論證特定的重要關鍵字，是否具有貫穿文本核心的意義。

其次，Van Dijk（2003: 349）曾強調「脈絡」（context）無疑是人們的「心靈模式」（mental models），是人們理解事物的社會認知準據，推估脈絡的方式，主要是由文本中言者論述的關鍵話語來加以確認；也就是說，言者強調、重視的脈絡要項，必然會在其所述文本中出現。基此，Van Dijk（2003:368）更指出，確認脈絡關係的方式可由文本中歸類出溝通的「範疇」（domain，意指面臨的溝通情境）、「背景」（setting，意指實際所處場域）、「擔任的「角色」（role，意指言者支持或反對的立場，諸如如何表述正向自我呈現與負向他者）、「訴諸「行動」（action，其中又包括全球行動、在地行動及當前行動等）及強調「目標」與「知識」（goal & knowledge）等類目，俾利印證內、外在脈絡間的互動關係。本研究建議後續研究，可以針對Van Dijk的「社會認知脈絡理論」所提的脈絡主題類目，也可透過文本探勘工具（如LDA分析）獲致相關主題類目的關鍵字串，進而釐清內在文本與外在脈絡間的特徵。

此外，對於兼具圖像、影音及語言文字之文本時，如何運用「文本探勘」進行分析，本研究並未處理，也是本研究的限制。然而，本研究仍建議後續研究採取「混合取徑」，針對圖像或是影音內容可先進行人工瀏覽與語藝分析，並以語藝批評理論爲

依據，運用文本探勘工具針對網路回文，進行爬文、斷詞與詞語共現等步驟；俾利論證圖象、影音內容，為何會引發認同的原由。總之，本研究強調，若能採取「混合取徑」原則，善用電腦文本探勘工具，將有益於解決網路巨量文本的問題。

參考資料

- 牛隆光（1998）。《新聞論述再現的哲學觀與價值觀—以海峽兩岸政權透過中央日報及人民日報進行「江八點」及「李六條」的對話行為為例》。台北：輔仁大學大眾傳播研究所碩士論文。
- 吳佩謙（2013）。〈以行為展演與敘事經驗建構身分認同：馬來西亞檳榔嶼華人穆斯林之個案研究〉，《新世紀宗教研究》，12（1）：89-118。
- 李欣宜(2015.04.01)。一次搞懂大數據（上）。數位時代。2016年12月28日，取自 <https://www.bnext.com.tw/article/35807/BN-2015-03-31-151014-36>
- 李蔡彥、鄭宇君（2011）。〈資訊科技與新媒體研究之發展〉，《傳播研究與實踐》。1（1）：75-80。
- 林頌堅（2014）。〈資訊科學期刊的主題分布與多樣性研究〉，《圖書資訊學研究》。9（1）：171-200。
- 周緝妮（2008）。《從食品謠言解讀當代飲食觀點》。台北：淡江大學大眾傳播學系碩士論文。
- 林頌堅（2010）。以詞語共現網絡分析探勘資訊傳播學領域的研究主題與關係。圖書資訊學研究，4（2），123-148。
- 林靜伶（2000）。《語藝批評理論與實踐》。台北：五南。
- 施旭峰、李蔡彥、鄭宇君、陳百齡（2014.05）。〈災難事件中社群媒體訊息之自動分類設計〉。ICIM2014第25屆國際資訊管理學術研討會，台中：國立中興大學，2014年5月24日。
- 施懿珊（2001）。《李登輝總統言論之戲劇觀—從柏克戲劇五因語藝觀點分析》，台北：世新大學傳播研究碩士論文。
- 秦琍琍、黃瓊儀、陳彥龍、張嘉予（2010）。〈組織認定、企業論述與組織文化的變革：從語藝觀點檢視公廣集團的整併過程〉，《新聞學研究》，104：67-111。
- 張秀蓉（1998）。《口語傳播概論》台北：正中。
- 張裕亮、陳奕任（2008）。〈爭辯中的中國網路民族主義論述〉，《香港社會科學

- 學》，35：103-142。
- 陳中淋（2013）。《紀錄原鄉：台灣運動紀錄片的再現模式分析》，台北：國立台北教育大學體育學系碩士論文。
- 陳世榮〈社會科學研究中的文字探勘應用：以文意為基礎的文件分類及其問題〉，《人文及社會科集刊》，27（4）：683-718。
- 陳佳靖（2002）。〈網路情色的符號地景：一個從語藝批評出發的觀點〉，《資訊社會研究》，3：25-62。
- 陳佳靖（2004）。〈喃喃嚙語話語藝：管它新瓶裝舊酒，定生出新味道〉，《資訊社會研究》，6：114-126。
- 陳譽晏（2015）。〈運用R Shiny建立文字探勘平台之語意分析及輿情分析〉，《數據分析》，10（6）：51-78。
- 陳良駒、傅振華、楊誌璋（2010）。〈詞彙共現分析在中國大陸信息作戰領域發展之實證研究〉，《中國大陸研究》，53（2）：115-116。
- 黃新生（1992）。《媒介分析方法》台北：遠流。
- 黃鈴媚、曹開明、沈錦惠、劉大華（2015.12）。〈社群媒體的「微」傳播影響力：從符號融合理論看反核臉書中核電真實之共構〉，「第二屆兩峽兩岸口語傳播學術研討會」論文。中國，廣西。
- 趙雅麗（2001）。〈民進黨社會運動的「語藝」批評〉，《新聞學研究》，68：151-192。
- 齊偉先（2013）。〈媒體災難敘事的社會意義建構：日本福島核災的戲劇分析〉，《思與言》，51（1）：103-134。
- 劉大華、黃鈴媚、林頌堅與曹開明（2016.12）。〈從數位人文觀點看語藝在運算轉向的機會與挑戰〉，「2016第七屆數位典藏與數位人文國際研討會」論文。台北：台灣大學。
- 劉吉軒(2016)。大數據分析與人文社會科學跨領域研究應用。傳播文化。2016年12月28日，取自<https://nccur.lib.nccu.edu.tw/bitstream/140.119/104032/1/412225.pdf>
- 劉名峰（2010）。〈現代性與國族認同的建構：從日治時期到民主鞏固期間之想像中國時的道德視域〉，《台灣社會研究季刊》，79：159-202。

- 劉德昌 (2004)。《美國總統布希在2003年美伊戰爭前演說之宣傳策略與戲劇五因分析》，台北：政治作戰學校新聞研究所碩士論文。
- 蔡哲明 (2011)。《臺灣選舉詆毀性訊息的語藝呈現(1998-2009)-柏克「戲劇五因」之觀點》，台北：中國文化大學新聞系碩士論文。
- 蔡鴻濱 (2004)。〈語藝批評方法在網路文化研究之應用與探討〉，《資訊社會研究》，6：91-126。
- 蔡鴻濱 (2011)。〈從語藝觀點探討荒誕宗教廣播節目之認同策略〉，《傳播研究與實踐》，1 (2)：211-243。
- Aleman, M. W. (2005). Embracing and resisting romantic fantasies as the rhetorical vision on a SeniorNet discussion board. *Journal of Communication*, 55, 5-21.
- Anderson, F., & Prelli, L. (2001). Pentadic cartography: Mapping the universe of discourse. *Quarterly Journal of Speech*, 87(1), 73-95.
- Arsenault, D. (2005). Rhetorical vision of the independent and sovereign nation of Hawai'i: A fantasy theme analysis. *Journal of Critical Postmodern Organization Science*, 3(2), 57-73.
- Blei, D. M., Ng, A. Y. & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, 3, 993-1022.
- Birdsell, D. (1987). Ronald Reagan on Lebanon and renada: Flexibility and Interpretation in the Application of Kenneth Burke's Pentad. *The Quarterly Journal of Speech*, 73, 267-279.
- Bishop, R. (2003). The world's nicest grown-up: A fantasy theme analysis of news media coverage of Fred Rogers. *Journal of Communication*, 53(1), 16-31.
- Bormann, E. G. (1972). Fantasy and rhetorical vision: The rhetorical criticism of social reality. *Quarterly Journal of Speech*, 58, 396-407.
- Bormann, E. G. (1973). The Eagleton affair: A fantasy theme analysis. *Quarterly Journal of Speech*, 59, 143-159.
- Bormann, E. G. (1982a). A fantasy theme analysis of the television coverage of the hostage release and the Reagan inaugural. *Quarterly Journal of Speech*, 68, 135-145.

- Bormann, E. G. (1982b). Fantasy and rhetorical vision: Ten years later. *Quarterly Journal of Speech*, 68, 288-305.
- Bormann, E. G. (1982c). The symbolic convergence theory of communication: Applications and implications for teachers and consultants. *Journal of Applied Communication Research*, 10, 50-61.
- Bormann, E. G. (1983). Symbolic convergence theory: Organizational communication and culture. In L. Putnam & M. E. Paconowsky (Eds.), *Communication and organizations: An interpretive approach* (pp. 99-122). Beverly Hill, CA: Sage.
- Bormann, E. G. (1985a). Symbolic convergence theory: A communication formulation. *The Journal of Communication*, 35, 128-138.
- Bormann, E. G. (1985b). *The force of fantasy: Restoring the America dream*. Carbondale, IL: Southern Illinois University Press.
- Bormann, E. G. (1990). *Small group communication: Theory and practice*, 3rd ed. New York: Harper & Row.
- Bormann, E. G., Cragan, J. F., & Shields, D. C. (1994). In defense of symbolic convergence theory: A look at the theory and its criticisms after two decades. *Communication Theory*, 44, 259-294.
- Bormann, E. G. (2003). Defending Symbolic Convergence Theory From an Imaginary Gunn. *Quarterly Journal of Speech*, 89(4), 366-372.
- Boyd, D., & Crawford, K. (2012). Critical questions for big data. *Information, Communication & Society*, 15(5), 662-679.
- Burke, K. (1937). *Attitudes Towards History* (3rd ed.). Berkeley, California: University of California Press.
- Burke, K. (1968). *Counter-Statement*. Berkeley, Ca: California Press.
- Burke, K. (1969). *A Rhetoric of Motives*. Berkeley, Ca: California Press.
- Burke, K. (1989). *The Legacy of Kenneth Burke*. Wisconsin: University of Wisconsin.
- Burke, K. (1996). *Language as Symbolic Action*. Berkeley, Ca: California Press.
- Church, K. W., & Hanks, P. (1990). Word association norms, mutual information, and

- lexicography. *Computational Linguistics*, 16(1), 22-29.
- Crable, B. (2000a). Burke's perspective on perspectives: Grounding dramatism in the representative anecdote. *Quarterly Journal of Speech*, 86, 318-333.
- Crable, B. (2000b). Defending dramatism as ontological and literal. *Communication Quarterly*, 48, 323-342.
- Cragan, J., & Shields, D. (1992). The use of symbolic convergence theory in corporate strategic planning: a case study. *Journal of applied communication research*, 20(2), 199-218.
- Darr, C. R., & Strine, H. C. (2008). *A Pentadic Analysis of Celebrity Testimony in Congressional Hearings: Giving Voice to the Voiceless*. Paper presented at the annual meeting of the National Communication Association.
- Denzin, N. K. (1978). *The research act: A theoretical introduction to sociological methods*. New York: Praeger.
- Duffy, M. E. (2003). Web of hate: A fantasy theme analysis of the rhetorical vision of hate groups online. *Journal of Communication Inquiry*, 27, 291-312.
- Edbauer, J. (2009). Unframing Models of Public Distribution: From Rhetorical Situation to Rhetorical Ecologies. *Rhetorical Society Quarterly*, 35(4), 155-71.
- Elwood, W., & Greene, K. (2005). "Risks both known and unknown": A qualitative method to assess the role of situation in HIV/STD risk and prevention. *Journal of Homosexuality*, 50(1), 135-154.
- Eaves, H., & Savoie, M. (2005). Big Brother: Merging reality and fiction: An application of the narrative paradigm. *Texas Speech Communication Journal*, 29(2), 91-97.
- Eyman, G. (2015). *Digital Rhetoric: Theory, Method, Practice*. Ann Arbor: University of Michigan Press.
- Fisher, W. R. (1984). Narration as a human communication paradigm: The case of public moral argument. *Communication Monographs*, 51, 1-22.
- Fisher, W. R. (1985a). The narrative paradigm: In the beginning. *Journal of Communication*, 35, 74-89.

- Fisher, W. R. (1985b). The narrative paradigm: An elaboration. *Communication Monographs*, 52, 347-367.
- Fisher, W. R. (1987). *Human communication as narration: Toward a philosophy of reason, value, and action*. Columbia: University of South Carolina.
- Fisher, W. R. (1988). The narrative paradigm and the assessment of historical texts. *Argumentation & Advocacy*, 25, 49-53.
- Fisher, W. R. (1989). Clarifying the narrative paradigm. *Communication Monographs*, 56, 55-58.
- Foss, K., & Griffin, L. (1995). Beyond Persuasion: A Proposal for an Invitational Rhetoric. *Communication Monographs*, 62, 2-18.
- Foss, S. (1996). *Rhetorical criticism : Exploration & Practice*(2nd ed). Prospect Heights, ILL : Waveland Press.
- Foss, S. (2004). *Rhetorical criticism: Exploration & practice*. Prospect Heights, IL: Waveland Press.
- Foss, S., Foss, K., & Trapp, R. (2002). *Contemporary perspectives on rhetoric*. Prospect Heights: Waveland.
- Glaser, B. G. (1978). *Theoretical Sensitivity: Advances in the Methodology of Grounded Theory*. Mill Valley, CA: The Sociology Press.
- Glaser, B. G. & Strauss, A. L. (1967). *The Discovery of Grounded Theory: Strategies for Qualitative Research*. Chicago: Aldine.
- Golden, J., Berquist, G., & Coleman, W. (1993). *The Rhetoric of Western Thought*. Iowa: Kendall/Hunt Pub .Co.
- Greer, C. (2008). Media literacy and internet discussion lists: A fantasy theme analysis of online postings following the death of Peter Jennings. *Studies in Media and Information Literacy Education*, 8, 1-11.
- Hart, R. P. (1990). *Modern Rhetorical Criticism*. Austin: Harper Collins.
- Heath, R. L. (1984). Kenneth Burke's break with formalism. *Quarterly Journal of Speech*, 70, 132-143.

- Hochmuth, M. (1952). Kenneth Burke and the "new rhetoric". *Quarterly Journal of Speech*, 38, 133-144.
- Johnson, R. B., Onwuegbuzie, A. L., & Turner, L. A. (2007). Toward a definition of mixed methods research. *Journal of Mixed Methods Research*, 1(2), 112-133.
- Lewis, S. C., Zamith, R., & Hermida, A. (2013). Content analysis in an Era of big data: A hybrid approach to computational and manual methods. *Journal of Broadcasting & Electronic Media*, 57(1), 34-52.
- Liu, F. & Lu, X. (2011). Survey on text clustering algorithm. In *Proceedings of 2nd International IEEE Conference on Software Engineering and Services Science (ICSESS)*, 901-904.
- Lucas, S. E. (1986). Book review on the Force of Fantasy: Restoring the American Dream. *Rhetoric Society Quarterly*, 16, 199-205.
- McCabe, J. (2009). Resisting alienation: The social construction of internet communities supporting eating disorders. *Communication Studies*, 60(1), 1-16.
- Meisenbach, R., Remake, R., Buzzanell, P., & Liu, M. (2008). They Allowed": Pentadic mapping of women's maternity leave discourse as organizational rhetoric. *Communication Monograph*, 75(1), 1-24.
- Mohrmann, G. P. (1982). An essay on fantasy of theme criticism. *Quarterly Journal of Speech*, 4(1), 1-6.
- Pashaei, F. H. (2010). *Unstable Situations: A Rhetorical Approach to Studying Blogs about Muslims*. Unpublished master's thesis, George Mason University, Virginia.
- Perry, S. D., & Roesch, A. L. (2004). He's in a new neighborhood now: Religious fantasy themes about Mister Rogers' Neighborhood. *Journal of Media and Religion*, 3, 199-218.
- Ridolfo, J. (2006). (C)omprehensive (O)nline (D)ocumen (E)valuation. *Kairos: A Journal of rhetoric, technology, and pedagogy*, 10(2). Retrieved January 8, 2016, from <http://kairos.technorhetoric.net/10.2/binder.html?praxis/ridolfo/index.html>.
- Shields, D. (1981). A dramatistic approach to applied communication research: Theory,

- methods, and applications. In Cragan, J.(Eds.), *Applied communication research- A dramatic approach* (pp5-13). Prospect Heights: Waveland Press.
- Shields, D., & Preston, Jr., C. (1985). Fantasy theme analysis in competitive rhetorical criticism. *National Forensic Journal*, 3, 102-115.
- Sovacool, B., & Brossmann, B. (2010). Symbolic convergence and hydrogen economy. *Energy Policy*, 38, 1999-2012.
- Smudde, P. (2004). Implications on the practice and study of Kenneth Burke's idea of a "Public relations counsel with a heart". *Communication Quarterly*, 52(4), 420-432.
- Steward, J. C., Smith, A. C., & Denton, E. R. (2001). *Persuasion and Social Movement*(4th ed.). Prospect Heights, Ill: Waveland Press.
- Sweden, L., Kreuter, F., Berg, M., Biemer, P., Decker, P., Lampe, C., Lane, J., O., Neol, C., & Usher, A. (2015). Bid data in survey research. *Public Opinion Quarterly*, 79(4), 839-880.
- Tell, D. (2007). Kenneth Burke and the politics of "Pure Persuasion". Paper presented at the annual meeting of the National Communication Association, Washington.
- Temple, A. N. (2011). *Reconceptualizing Fantasy Theme Analysis in a Digital Online Context*. Unpublished doctoral dissertation, State University of North Carolina, North Carolina.
- Van Dijk, T. A. (2003). Text and context of parliamentary debates. In: P. Bayley (ed.), *Crosscultural perspectives on parliamentary discourse* (pp. 339-372). Amsterdam: Benjamins.
- Vultee, F. (2012). Man-child in the White House: The discursive construction of Barack Obama in reader comments at foxnews.com. *Journalism Studies*, 13, 54-70.
- Warnick, B. (1987). The narrative paradigm: Another story. *Quarterly Journal of Speech*, 73, 172-182.
- Warnick, B. (2007). *Rhetoric online: Persuasion and Politics on the World Wide Web*. NY: Peter Lang.
- Williams, C. E. (1987). Fantasy theme analysis: Theory vs. Practice. *Rhetoric Society*

Quarterly, 17, 11-20.

Zappen, J. P. (2005). Digital rhetoric: Toward an integrated theory. *Technical Communication Quarterly*, 14(3), 319-325.

