

「2017 AI の語言世界」工作坊

陳純音*

一、緣起

人工智慧 (Artificial Intelligence, AI) 近年來一直受到重視，在 Google AlphaGo 戰勝棋王那一刻後，AI 一詞迅間更成為時下流行語。前人文及社會科學研究發展司 (簡稱人文司) 洪世章司長於 2017 年 9 月 5 日召開科技部人文司科技發展會議並指示，當工程技術研究發展司 (簡稱工程司)、前瞻及應用科技司 (簡稱前瞻司) 與產學與園區業務司 (簡稱產學司) 等司對 AI 投入心力探索的同時，人文司更應該踴躍響應並奉獻所學，進行跨領域對話，因此人文司於去年下半年舉辦了一系列的 AI 工作坊。法律學門工作坊訂在 9 月，語言學門安排在 10 月，11 月分有心理學門，還有 12 月的哲學學門壓軸，可說是人文司總動員的活動。

有鑑於 AI 的應用對人類生活與社會可能帶來的便利，以及未來人類使用語音控制 (voice control) 機器機會的增高，更加堅定我們對 AI 的認知，如：機器學習、深度學習、自然語言理解、環境意識和數據隱私。其中，自然語言理解跟語言學門息息相關。人類語言是一種自然語言 (natural language)，是一個溝通工具，也是一套系統，包含語音／音韻、構詞、語意、語法的組合律、言談篇章規則與語用行為條件等。人與人的互動與溝通是相當複雜的行為，語境中的性別、輩分、社會地位、親疏關係與文化背景等皆可影響說話者的選詞與用字，決定聽話者對隱喻或幽默等策略之理解，這些也是 AI 目前面臨的最大的挑戰之一。本工作坊以 AI 與語言學次領域的相關性作為主題，針對語言學對 AI 發展趨勢可提供的助益，與 AI 在語言學相關議題的應用研究，進行了一天的討論與交流，期能提供研究 AI 領域學者專家與業界一些建議，並作為政府相關部會決策之參考。

* 國立臺灣師範大學英語學系教授、科技部人文及社會科學研究發展司語言學門召集人

二、工作坊規劃與舉辦

這次的工作坊規劃時間短，但相當感恩下面這幾位好友情義相挺。一天的工作坊我規劃了 6 個場次的主題，由語音、語意、句法、手勢、文本分析、科技應用到綜合座談，盡量涵蓋到語言世界的點線面，與與會人士交流。工作坊由科技部人文司語言學門與國立臺灣師範大學英語學系共同主辦，於 2017 年 10 月 21 日（星期六）在國立臺灣師範大學綜合大樓國際會議廳 509 隆重登場：

- | | |
|--------------------------------|-----|
| 場次 1 語音技術在 AI 上的應用 | |
| 主持人：中央研究院語言學研究所 | 鄭秋豫 |
| 主講人：國立臺灣師範大學英語學系 | 甯俐馨 |
| 國立臺灣大學語言學研究所 | 邱振豪 |
| 場次 2 AI 的語意網絡探討 | |
| 主持人：國立中正大學語言學研究所 | 吳俊雄 |
| 主講人：國立臺灣大學語言學研究所 | 謝舒凱 |
| 國立政治大學英國語文學系 | 鍾曉芳 |
| 場次 3 句法規則與手勢研究對 AI 的價值 | |
| 主持人：國立政治大學資訊科學系 | 劉昭麟 |
| 主講人：臺北醫學大學通識教育中心 | 周昭廷 |
| 國立臺北科技大學應用英文系 | 林彥良 |
| 場次 4 文本分析、機器翻譯與 AI 的未來 | |
| 主持人：慈濟大學英美語文學系 | 周中天 |
| 主講人：國立臺灣科技大學應用外語學系 | 王世平 |
| 國立臺灣大學外國語文學系 | 高照明 |
| 場次 5 AI 科技在語言學相關議題的應用研究 | |
| 主持人：國立臺灣師範大學英語學系 | 陳純音 |
| 主講人：國立臺灣師範大學華語文教學系 | 籃玉如 |
| 國立成功大學外國語文學系 | 劉繼仁 |
| 場次 6 綜合座談：與 Siri 對話 嗨一下 | |
| 主持人：實踐大學應用外語學系 | 黃美金 |
| 主講人：國立臺灣大學語言學研究所 | 馮怡蓁 |
| 國立清華大學語言學研究所 | 蘇宜青 |
| 國立臺灣師範大學英語學系 | 蘇席瑤 |

這次工作坊線上報名 188 名，參加人數有 171 人，算是相當踴躍。與會人士有近 4 成是學者專家，5 成多是大學生和研究生，有近 1 成屬社會人士（含業界）。學者專家中除了語言學門師長們，還有其他學門（如：文學、資科、醫學與哲學等領域）的學者蒞臨指導，由此可見，各領域對 AI 議題相當感興趣，與會學子人數過半，更可看出 AI 時代是與年輕人息息相關的時代，他們的好奇心各個寫在臉上，各場主題皆帶來新知，著實感動。其中值得一提的是，哲學學門召集人謝世民教授與相關師長也出席本工作坊，了解語言學門與 AI 之關聯性，相當感謝。我於閉幕時，也幫忙播放了一段動畫，希望能行銷 AI 時代所帶來的發人深省的哲學議題，接軌 12 月分哲學學門的 AI 工作坊。

三、工作坊各場次紀要

這次工作坊能順利圓滿舉辦，除要謝謝助理夥伴們同心協力完成任務，讓與會人士有賓至如歸的感覺外，更要感謝講者們的大力幫忙，各場次的主題內容豐富且多元，知識性與趣味性兼具，他們真是工作坊的亮點。以下簡單摘要各場次之重點，與大家分享：



圖一：「2017 AI の語言世界」工作坊

場次 1 語音技術在 AI 上的應用

場次 1 主持人鄭秋豫老師一開場就直接切入主題，點出「跨學科」的重要，鼓勵研究結合跨領域學科，踏出腳步才能看到世界，並提醒大家研究必須面向

寬廣，且有 focus，才不會閉門造車。接著，甯俐馨老師以語音結合 AI 出發，用貼近生活的例子帶入主題，點出臺灣目前 AI 語音發展的最大問題：資料庫不足，並指出兩點語言學家能做的，也就是建立資料庫及進行錯誤分析。最後她以 “Linguists play an important role in cleaning data.” 說明機器可以做到 90%，剩下的 10% 就是靠語言學家，勉勵在場語言學者能做的事是把 AI 推向更完美完善的境界。

最後，邱振豪老師進一步用三個部分來說明語音與 AI：機器學習 (Machine learning)、人類機器 (Human-machine) 和時間 (Time)。關於機器學習，邱老師介紹了 Deep neural network (DNN) 模擬人腦，處理多個輸入 (input) 經過多道隱藏的處理過程，變成簡化的輸出 (output)，並舉例將 DNN 用於標音與舌位高低的結果：DNN 能告訴我們滿好的結果，雖然有些無法辨別出來，但誤差極小，且辨別不出來的部分跟人辨別不出來的部分一樣，因此還是可以在多方面運用機器 DNN。接著，邱老師提到了氣流對皮膚的感知，並舉例說明此研究可用於手機感知收音，就算有很多雜音還是能辨識說話發音，清楚傳達通話。而發音的肌肉形狀也會因世界語言而有所不同，也是值得研究發掘的主題，並總結我們需要更多的資料來建構資料庫，而且因為模型不會思考，我們更需擔任思考的角色，勉勵大家要更了解自己的語言。

場次 2 AI 的語意網絡探討

第二個場次由吳俊雄老師主持，邀請謝舒凱老師和鍾曉芳老師就 AI 的語意網絡這主題進行分享。謝舒凱老師以最常被誤解的觀念帶入主題：(1)「語言辨識」不等於「語言理解」。(2) 語音辨識 + 字串理解 = AI。先介紹了目前 AI 最常做的運用包含知識搜尋處理、感知 (人臉辨識)、生成 (圖文詩詞等自動生成)。接著介紹深度學習 (Deep learning) 和語言科技的應用例子，如：手語翻譯 (手勢辨識)、及作者自己正在建構的網路機器人——LopeChat。令人印象深刻的是，作者建構的網路機器人準確度驚人，上傳自拍照片即能判斷出作者的基本背景 (心情、年齡……等)。最後，作者提出需要再討論的相關議題，包含先前提到的數據取得、「演算法正義」相關議題 (對人種的判斷彈性)、資料庫再大也難以解決的問題等。最後，謝老師提出想介入 AI，須使用向量語意學，將不同方面不同領域詞語所表達的意思建構完全，而語言學所做的研究都將可能成為 AI 的系統資料之一。

鍾曉芳老師則進一步提及 AI 要完善，必須結合語用、語意、文化，她介紹了數個語意相關的有趣例子，如：隱藏意義 (hidden meanings)。在語料庫搜尋

關鍵字「連…都」與「連…都…知道」找到的句子，雖都使用「連…都…（知道）」，但形成的句子有各種意思，有些意思是語句中看不出來的。另外，說話者有時說的句子聽者必須在場，才能理解說話者的情緒，以及語義歧異可能造成的混淆等現象。這些在 AI 研究中，目前應該都還有發展空間，需要足夠的專業知識來建構語料庫，並提升 AI 技術。

場次 3 句法規則與手勢研究對 AI 的價值

場次 3 由有資料背景的劉昭麟老師主持，邀請兩位年輕學者周昭廷老師和林彥良老師談談句法規則與手勢研究對 AI 的價值。周昭廷老師承接前場次，先以常識（common sense）對電腦來說很難起頭，談到人認為理所當然的事，因電腦不了解句法結構與隱藏意義不易克服。接著，以 Google 翻譯操作為例，說明即使句子簡單仍可能有很多暗示（hints），Google 依舊無法翻譯成自然句子來討論人類語言的屬性：結構依賴性（structure dependence）和極小性（minimality），並介紹聊天機器人（chatbot），說明人類語言非句子表面、不是每個字單獨存在，而是有層次的組成。

林彥良老師則介紹了手勢與 AI 的結合。一開始先提醒大家，手勢分析的語料可從很多地方取得，例如：電影、影片、演講紀錄等，與我們的生活息息相關。手勢在 AI 中可運用在教室內學習建構、手機懸浮控制等地方。而手勢也可以與應用語言學與溝通結合，例如不同的主題對話、場合、性別等皆會影響手勢。林老師也分享了自己目前做的與手勢相關的研究所使用的工具（Anvil, Elan）等，並強調手勢在語言溝通的重要性，例如有時候一句話「他往那兒走了！」會搭配手勢指某個方向（speech coding）。最後，他概述了目前研究努力方向與手勢結合 AI 的未來展望：從手動工具得到輸出，再將輸出給 AI 學習，進一步變成控制的手勢。研發 App 與聾啞人士的互動、利用手勢的大數據診斷人健康不健康、有無疾病以及智慧家電等操作。

場次 4 文本分析、機器翻譯與 AI 的未來

由慈濟大學周中天老師主持，高照明老師和王世平老師談論文本分析、機器翻譯與 AI 的未來。高照明老師先概述了電腦演算法的演變，接著舉了幾種翻譯工具作為例子，包含 Google 翻譯，因有足夠對應的平行語料，強項是詞組翻譯。但對應網路上的平行語料有時也會有問題，例如：對岸與臺灣的用語不同，但翻譯出來以對岸用語為主，原因出在翻譯系統以網路有資料或資料多者取勝。RNN 系統演算法可解決多種問題，除了機器翻譯外，還有聊天機器人，白



圖二：會場一隅

話文文言文，簡繁轉換等。最後，他總結雖然系統有複雜的數學公式，我們可以做的是將句對應、詞對應、音節對應用於語言教學中。

接著，王世平老師以貼近生活的實例：“Linsanity”為例，說明生活四周充滿了語言學，包含了 Lexical chain、Move、Text analysis 等，並提供了網路統計 27 種與文本分析 (text analysis) 有關的分析軟體，建議大家可以多加使用於學術論文寫作中。

場次 5 AI 科技在語言學相關議題的應用研究

場次 5 由我主持，邀請身為研究資訊科技相關到語言教學的跨領域學者籃玉如老師，和資訊與英語教學背景的劉繼仁老師談 AI 科技在語言教學上的應用。籃玉如老師以幾部電影的片段為例，說明人機對話、環境轉換、模擬情境角色扮演等，並說明這些例子能如何用於語言教學。例如揣摩母語人士與學習者做對話練習；自然語言處理 -data 與 VR 結合，模擬學習者目標語言的環境；後設認知能力的部分：偵測學習者的學習狀況，給予鼓勵與情感支持等等；精準測驗出學習者的能力程度，給予學習者客製化的學習。

劉繼仁老師強調語言是用於溝通、語言十分重要之外，承續籃老師說明語言學習結合 AI 有三個參考指標：有效性、效率、吸引力。接著，談到 E-learning 的思維轉換三階段：E-learning (包含很早就出現的 BBS)、Mobile learning (手機 App 語言學習)、Context-aware U-learning (環境感知、互動學習)。科技結合語言學習受歡迎是因為有許多優點：客製化、能讓人無限反覆重複練習。劉老師

也比喻過去生活三要素：陽光、空氣、水，現代社會更多了一元素：電。科技可做的事情，包含最常見的搜尋、修改、聯絡等，目前世界上有些國家都已大力推廣支持資訊學習資源，並建議結合 AI 做兩大方向的研究：Assessment 和 Evaluation。

場次 6 綜合座談：與 Siri 對話 嗨一下

場次 6 是綜合座談，由黃美金老師主持，讓三位不同語言學次領域的老師們與 Siri 現場對話，並從語言習得、社會語言學及語音學等方面，探討人與 AI 的不同以及 AI 未來能做的應用與發展。

蘇宜青老師從母語習得的角度講述人工智慧與人腦的差異：是否具有心智？具有心智的人，即使是學齡前的小孩子仍能在語料輸入 (input) 後產生無限多種、無限多長的語句回應，且能知道語法限制。而沒有心智的 AI (如 Siri) 卻不會知道語法限制，因此常答非所問。蘇席瑤老師從社會語言角度來探索 AI，她認為 Siri 不能給予期待的答案與語言的社會性的關係，並舉了她朋友的生活例子——「從教小孩不吃手論訓練 AI 的難度」為例，說明文化、社會的互動固然複雜，但仍有其規則存在。不同人群間的語言表達與個體語言風格皆有其差異，蘇老師也建議可將社會語言學結合 AI 在生活上應用，例如服務業 (機器售票) 改變語音更貼近消費者。馮怡蓁老師則提供了 AI 發展之後可能產生的現象與反思。

在座談會尾聲，三位老師再次深入說明了 AI 與人的不同之處，這些地方也是 AI 未來可發展更完善的方向，並以 AI 發展可能產生的利與弊做總結。



圖三：與會工作人員合影

四、展望

這次的工作坊讓與會人士在歡笑中開啟了一扇知識之門，走進未來世界，讓大夥學到一些 AI 相關的術語，例如：AI 不等於機器學習 (Machine Learning)，後者可以說是前者的一部分，透過算法給機器新數據，讓機器做出決定。此外，各場次學者的分享，也可看出自然語言處理 (Natural Language Processing) 是 AI 研究仍需努力的地方，也是語言學家最能貢獻一己之力的地方，畢竟人與人的溝通不像跟機械下指令，有其複雜的網絡。享受知識饗宴的同時，我們也感受到隨著新世紀的來臨，我們也要有新思維。只有 AI 技術沒有語言學學理，無法將 AI 優化升級，而只有語言學學理無 AI 技術，也僅是紙上談兵。未來的時代顯然是跨領域合作的時代，讓我們彼此多一點對話，多一些交流，我們將可以更互補地勾繪更幸福的未來。

參考資料

A 6 minute Intro to AI <https://snips.ai/content/intro-to-ai/>