

跨語句公式化語序的心理真實性： 以小學數學文字題為例之事件相關腦電位研究

姚甸京、呂菁菁

國立清華大學 臺灣語言研究與教學研究所

通訊作者：姚甸京

聯絡地址：330 新竹市南大路 521 號

電子郵件：itisno1@gapp.nthu.edu.tw

投稿日期：2020 年 3 月

接受日期：2020 年 5 月

摘 要

本研究運用事件相關腦電位之實驗研究法，以小學數學文字題作為實驗材料，驗證公式化語序在語言理解運作時所產生的影響，以檢視跨語句之公式化語序在語句理解運作時的心理真實性。本研究以不同問句類型在 500ms~900ms 的時間區段之間，所誘發之平均腦波波幅進行統計檢驗，結果發現相較於常見問句；罕見問句與不合理問句在平均振幅、潛伏期和峰值振幅上皆呈現出顯著的 P600 效應，因此我們推論語句理解運作時，跨語句之公式化語序的心理真實性是可能的，也希望藉由本研究結果，對於語句理解的認知運作相關理論，提供一些不同角度的思考。

關鍵詞：語言運作、語句理解、公式化語序、事件相關腦電位。

壹、緒論

「語言」這是人類演化史上重要的發明，更是推動人類文明進程的重要基石。對人類而言，「語言」不僅只是如同動物的溝通系統，也是人類認知思維運作的體現，是可以讓我們一窺大腦奧秘的窗口。除了一般所觀察、收集、分析出的形、音、義等語言規則之外，心理語言學所關心的議題是人類如何習得、學習語言，以及人類如何在不同的語境(context)下透過變動複雜的語言系統，來表達自己的想法或是理解他人的意圖，以進一步了解人們在進行語言相關活動時，大腦的思維認知功能的運作歷程又是如何？

現代的心理語言學領域對於人類的語言運作，在經過科學化的研究後發現，語言是經由習得(acquisition)、學習(learning)而來，「意義」(meaning)是語言使用者經由參與溝通

(communication)互動，以自己的理解推論所產生，因著日漸累積的語言活動經驗，進而習得所使用的語言系統，掌握語言系統中的的架構模式(Pattern)，如此人類語言的表達及理解才能更有效率及省力的進行，自然語言中的「公式化語序」(formulaic sequences)就是例證之一。

在日常語言使用中，不同類型的公式化語序現象出現的頻率相當高，一般人在使用語言及理解語言時，並非總是對語句中的各個詞彙逐一的進行解讀，同時也會運用公式化語序，進行語言理解的運作加工，以便更有效率的掌握多字串以提高語言使用的流暢性、更易於溝通以及更符合大腦的認知運作歷程(Erman & Warren, 2000; Schmitt & Carter, 2004)。小學數學文字題基本上是由幾個條件句組成題幹，再搭配上問句所組成，在基本的呈現形式與其所相對應的運算意義上，是跨越句子(條件句)與句子(問句)之間在相互的搭配上有

相關程度的關聯性，在文字敘述上較不涉及複雜的句法結構，因此在實驗設計上可以讓受試者著重於語句理解的任務，是適用以檢視跨語句之公式化語序，在語句理解運作時的心理真實性的實驗語料。

貳、文獻探討

一、「公式化語序」(formulaic sequence)

早在 19 世紀時失語症臨床研究的領域中，醫師從失語症病人的語言表現上發現到公式化語序(formulaic sequence)的現象，此後，研究者也漸漸察覺到在日常語言使用中，不同類型的公式化語序(formulaic sequence)現象出現的頻率也相當高，一般人在使用語言及理解語言時，並非總是對語句中的各個詞彙逐一的進行解讀，他們同時也會運用自然語言中的公式化語序(formulaic sequence)進行語言理解的運作處理，以便提高語言使用的流暢性、更有效率的掌握及更易於溝通，也符合大腦在認知心理運作上的省力原則。

在日常生活中，公式化語序(formulaic sequences)大量的使用在語言的溝通理解上，所涵蓋的範圍很廣，大多是固定搭配出現的語言表達方式，從片語詞組、慣用形式(conventionalised forms)、約定成俗的說話方式(ready-made utterances)、俗諺(proverbs)、俚語(idioms)、句型架構(sentence builders)或言談時成套搭配的表達形式(preassembled speech)...等。「公式化語序」(formulaic sequence)指的是那些在大腦中以整體形式儲存及提取的，不是通過既有的語法衍生或分析出來的。基本上就是以「整體存儲與提取」來界定所謂的「公式化語序」，也就是說，一組由多個詞彙所組成的語言單位、句式...等，只有被語言使用者熟練地使用並以整體的形式，存取在個體的心理詞彙庫中，才能稱為「公式化語序」(Schmitt & Carter, 2004; Wray, 2004, 2005)。

小學數學文字題基本上是由幾個條件句組成題幹，再搭配上問句所組成，條件句所提供的是相關已知數量的訊息，問句則是表示未知數量的訊息，題幹組搭配問句以產生運算相關概念的意義，讓解題者藉由閱讀語句，理解出題意後，推論已知數量和未知數量之間的關係，進而根據相關運算原則和解題策略解出答案。小學數學文字題在基本的呈現形式與其相對應的運算意義上，是跨越句子(條件句)與句子(問句)之間在相互的搭配上有相關程度的關聯性，用以作為本研究所討論的跨語句之公式化語言序列(the formulaic sequences of cross-sentence)。

二、事件相關腦電位與成分波 P600

事件相關腦電位(Event-Related Potentials, ERPs)是一項

運用腦電波圖(electroencephalogram, EEG)技術，在實驗進行中外加經過設計操控的特定刺激項，以觀察實驗參與者在接收到刺激或撤銷刺激時，在其腦部所誘發的大腦電位變化。心理語言學側重于研究個體在進行語言認知理解時大腦的運作歷程，因為大腦神經訊息傳遞的速度相當之快，而事件相關腦電位(ERPs)具有很高的時間解析度(ms)，因此事件相關腦電位的研究方法常為心理語言學領域所採用。因實驗設計所誘發的正負腦電波動可以被稱為「峰」、「波」或「成分」(component)，是事件相關電位研究探討實驗結果的主要參考指標，依照腦電波出現的時間以及極性，最常受到語言學領域討論的事件相關腦電位之腦電波成分有 N400 以及 P600(Kutas & Federmeier, 2011)。

以事件相關腦電位(ERPs)研究大腦進行語句理解運作時，P600 是常運用的觀察指標。P600 是腦電波之最大波峰(Peak)出現在約 600ms，其潛伏期(Latency)為 600 ms~900ms 的正向波，最大波幅差異通常會出現在頭皮頂部中後區的電極點區域。此 P600 效應主要出現在語句的單複數一致性、短語結構、動詞時態、格位一致性、句子成分移位等與句法相關的違反和花園徑句的語句，如果存在訊息整合上有困難，就會誘發 P600 效應，反映出大腦進行語言理解運作時，對於語句理解的整合和再分析的過程(Fisher, Bassok, & Osterhout, 2010; Friederici, 2002; Hagoort, 2003; Kim & Osterhout, 2005)。

過去的研究發現，N400 效應與語意上的處理有關，而在句法上的違反或整合有困難則會誘發 P600 效應，然而，這種單一線性式的句法加工階段和與 ERPs 成分之間一對一的對應關係，日漸受到討論。從近年來相關研究中發現的語意 P600 效應，對於語言運作的相關理論，提出了不同角度的觀點，所謂語意 P600 效應是指在語法上正確，但論旨角色出現搭配錯誤(通常是施事者和受事者的角色顛倒)的語句所誘發出的 P600 效應(Kim & Osterhout, 2005; Kuperberg et al., 2003)，而其他發現會誘發語義 P600 效應的因素，還有語義違反(semantic violation)、語義聯繫(semantic association)、有無生命性的違反(animacy semantic anomaly)、實驗任務要求，以及語境(context)的作用...等(Kuperberg, 2007; Nieuwland & Van Berkum, 2005)，此外，以認知監控理論的觀點認為，語意 P600 效應所反映的並不只是純粹的句法再分析的歷程，而是反映了在意義上更廣泛的語言理解過程中的認知監控，也就是大腦的認知監控系統，對可能出現的錯誤的檢測與衝突解決(Kuperberg, 2007; Shen, Fiori-Duharcourt, & Isel, 2016; Van Herten, Chwilla, & Kolk, 2006; Vissers, Chwilla, & Kolk, 2007)。

三、事件相關腦電位在數學文字題上的應用

Fisher et al. (2010)以事件相關腦電位，研究受試者在閱讀簡單的加法、除法和單詞範疇一致性等問題的句子，觀察他們協調語意和算術知識一致以及不一致的情況。(例如，一致：“十二朵玫瑰加三朵雛菊等於十五”；不一致：“十二個餅乾加三個罐子等於十五”)，此研究發現在進行加法的實驗中，受試者接收到錯誤的算術答案時，在電極點 Cz、Fz 及其周圍，腦波會誘發出顯著的 N400 效應，而當詞彙語意範疇(目標詞)不一致時，在電極點 Pz 會誘發出顯著的 P600 效應；在進行除法的實驗中，受試者接收到錯誤的算術答案時，在電極點 Cz 及 P4 周圍，腦波會誘發出顯著的 N400 效應，當詞彙語意範疇(目標詞)不一致時，在電極點 Cz、Pz、P3、P4 則誘發出顯著的 P600 效應，結果發現 P600 可以作為數學文字題中結構一致性的指標，也做為語句中語法一致性和其他有意義的序列的指標。

自 20 世紀末至今，隨著新的研究方法及技術的發展使用，例如 eye tracking、ERPs、fMRI...等，讓公式化語序的理論概念，進入到不同的語言學研究領域，如心理語言學、認知語言學、語言習得、第二語言學習...等，擴展出相關以大腦神經網絡連結為基礎的語言運作模型。本研究運用事件相關腦電位之實驗研究法，以小學數學文字題作為實驗材料，驗證公式化語序在語言理解運作時所產生的影響，以檢視跨語句之公式化語序在語句理解運作時的心理真實性。

參、研究方法

一、實驗參與者

實驗參與者共有 17 位成年人(女性: 13 人; 男性: 4 人)，年齡範圍為 28~52 歲 ($M = 40.71$, $SD = 5.32$)，所有實驗參與者皆為大專程度以上，經愛丁堡慣用手問卷(Edinburgh handedness inventory)¹評估皆為慣用右手者，無明顯的視力及聽力障礙，亦無語言障礙、腦部損傷、中風、失智症等病症。參與本實驗時聽力正常、視力正常或矯正正常，在了解實驗目的及流程後，均簽署實驗參與同意書後，才進行正式實驗。本研究經國立清華大學研究倫理審查委員會審查通過(REC 編號：10811HT105)。

二、實驗語料設計

本實驗以小學數學文字題(應用題)中添加型問題，做為設計實驗語料之參考依據，小學數學文字題的文字敘述主要

可分為條件句以及問句兩部分，本實驗所討論的添是由題目中添加型的條件句(起始量以及添加型的改變量)搭配詢問起始量改變後的結果量的問句而成，此題型的基本算式為 $a+b=c$ 。(a=起始量、b=改變量、c=結果量)。

我們將施測題目中的條件句，控制為添加型的條件句組合：起始量(條件句一)和添加型的改變量(條件句二)，並以題目的問句做為實驗的操控項。我們以添加型的條件句搭配三種不同的問句(詢問結果量、詢問改變量、詢問差異量)成為常見問句、罕見問句、不合理問句等三種不同問句類型，然後進行語料配置之編寫(參閱表一)。本實驗共設計 39 個不同情境的添加型條件句組合(起始量 + 添加型的改變量)，每個添加型條件句組合會搭配中三種問句類型的其中兩種問句類型，因此本實驗共有 78 個測驗項，並以隨機之方式呈現給實驗參與者。

三、實驗流程

為避免外在環境對研究參與者的干擾，本研究安排實驗參與者在舒適、光線適中且無聲音干擾的獨立實驗室空間，且在進行實驗的過程中，如果研究參與者有任何不適，可以隨時中斷實驗。在每一道測驗題開始前，螢幕上會先出現一個「+」的符號，表示測驗題目即將開始，提醒實驗參與者專注看著螢幕，接著螢幕上會依序出現一道數學文字問題，請實驗參與者在看完題目後，對於文字題句末所之搭配問句進行語句合理性判斷，並用滑鼠按鍵進行回答，如果呈現之題幹組與問句的搭配是合理的，則按滑鼠上之正確鍵(O)，表示符合；如果所搭配出現的問句是不合理的，則按滑鼠上之錯誤鍵(X)，表示不相符合。簡要流程如下圖一所示。

本研究的測驗作業是以 STIM 2 軟體編撰程式並進行隨機撥放，所有的刺激皆呈現在 22 吋電腦螢幕，其解析度為 1024×768 像素，實驗參與者以距離 75 cm 平視電腦螢幕進行測驗。本實驗使用 Neuroscan 公司所生產之腦波紀錄儀器，腦電位訊號的是透過實驗參與者佩戴之電極帽上的電極點進行收集記錄，電極帽上的電極點均為銀／氯化銀合金(Ag/AgCl-Sintered)材質，此 30 個電極點採用國際通用 10-20 腦波紀錄系統(the International 10-20 System)配置。電極帽以前額 GND 為接地電極，參照點 A1、A2 置於左右耳後之凸骨(mastoid)處，眼部移動時產生的肌電訊號(Electrooculogram, EOG)，由左右眼外側(canthal)的電極點 HEOL 和 HEOR 主要監控紀錄水平眼電(HEOG)，置於左眼上下的電極點 VEOU 和 VEOL 主要監控紀錄垂直眼電(VEOG)，用以記錄並排除會干擾 ERP 的眼動及眨眼行為。實驗過程中的腦電波訊號由 Neuroscan 公司所生產之腦波放大器 NuAmps 進行訊號放大，訊號的採樣率為 1000HZ，所

¹ Oldfield, R.C. (1971) The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*, 9, 97-113.

表一 實驗語料配置說明範例

條件句一 (起始量)	條件句二 (添加型改變量)	問句	備註
		阿明現在有幾顆彈珠？	常見問句 (詢問結果量)
阿明有 15 顆彈珠	小強再給他 7 顆	請問 小強少了幾顆彈珠？	罕見問句 (詢問改變量)
		阿明和小強相差幾顆彈珠？	不合理問句 (詢問差異量)



圖一 實驗試題呈現之簡要流程圖

有的電極皆與皮膚接觸良好，在電阻為 5kΩ 以下進行收錄，濾波設為 0.1-30 Hz band pass，並即時透過 SCAN 4.5 進行腦電波訊號的記錄。

四、資料處理與統計分析

(一) 行為反應數據篩選與分析

本研究收集之行為數據為答題選擇以及作答反應時間 (Reaction time, RT)，並將不具代表性的行為反應資料，例如：沒有對刺激項作答的按鍵反應、錯誤按鍵反應、超過待答時間(3000ms)的按鍵反應等三種情況，排除在統計分析之外。由於本研究的實驗作業設計特性，行為數據之資料不以正確與否來區分不同的操控項，主要紀錄實驗參與者的答題選擇傾向，因此答題選擇之依變項無統計上的檢驗意義，本研究僅能檢驗作答反應時間上的統計效果。本研究以統計軟體 SPSS 18，將三種不同問句類別(常見問句、罕見問句、不合理問句)之作答反應時間(ms)，進行單因子相依樣本變異數分析(Repeated measured ANOVA)。

(二) 事件相關腦電位之擷取與分析

完成收集腦波記錄之後，腦電波之事件相關電位使用 Neuroscan 4.5 分析軟體進行離線的數據分析，其分析的步驟如下：先進行眼部偽跡校正(Ocular artifact reduction)，接著進行腦波圖的分段擷取(Epoch)：本研究之目標訊號標記(trigger)為操控項(問句)出現時，我們擷取目標訊號標記出現前 100ms 至目標訊號出現後 1000 ms、濾波處理 (Filter)：頻寬設定於 0.1 Hz 至 30 Hz band pass、基線校正(Baseline correction)：目標訊號標記出現前 100ms，排除雜訊干擾(Artifact rejection)：排除振幅在- 50μv ~ 50μv 範圍之外者，疊加平均 (Average)：最後依各個操控項類別分別進行疊加平均。

肆、結果

一、行為反應數據之結果

由於本研究的實驗作業的設計特性，按鍵作答之行為數據資料，不以是否正確來區分不同的問句類型，主要是紀錄實驗參與者對於判斷不同問句類型之選擇傾向，以及作答反應時間，結果如表二所示。

表二 不同問句類型之答題選擇以及作答反應時間

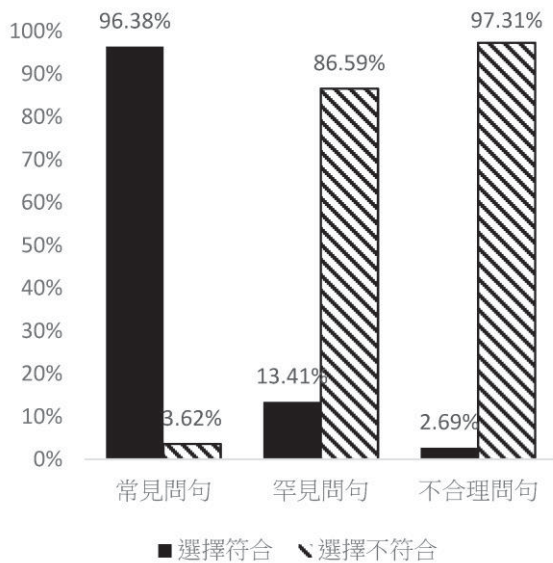
	選擇符合	選擇不符合	反應時間(ms)
常見問句	96.38%	3.62%	637±286
罕見問句	13.41%	86.59%	653±318
不合理問句	2.69%	97.31%	633±270

依本實驗參與者對於不同問句類型之答題選擇傾向(參閱圖二)，我們可以發現不同操控項是可以為實驗參與者所區辨的，傾向選擇常見問句與題幹組在語意搭配上相符合者最高(96.38%)，而傾向選擇在語意搭配上相符合者低的是罕見問句(13.41%)以及不合理問句(2.69%)。本研究以統計軟體 SPSS 18，將三種不同問句類別(常見問句、罕見問句、不合理問句)之作答反應時間(ms)，進行單因子相依樣本變異數分析(Repeated measured ANOVA)，結果發現反應時間的主要效果沒有達顯著的差異(P>.05)。

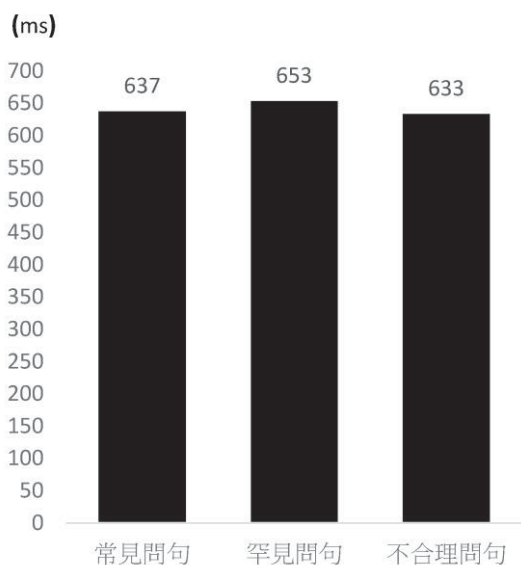
二、事件相關腦電位資料分析結果

(一) 平均腦波電位分析結果

我們將實驗中常見問句、罕見問句、不合理問句等三種不同問句之類別，所誘發的平均腦波電位的形態變化效應，經參考相關文獻的結果並透過觀察平均腦波電位資料之後，



圖二 不同問句類型之答題選擇傾向。



圖三 不同問句類型之答題反應時間。

表三 不同問句類型之平均振幅、潛伏時間與峰值振幅的平均數與標準差

	振幅 (Mean Amplitude)		潛伏時間 (Latency)		峰值振幅 (Peak Amplitude)	
	M	SD	M	SD	M	SD
常見問句	3.64	2.95	679.38	79.64	6.24	3.14
罕見問句	3.98	3.08	643.67	46.43	7.38	3.44
不合理問句	4.74	3.66	662.19	64.27	8.15	4.45

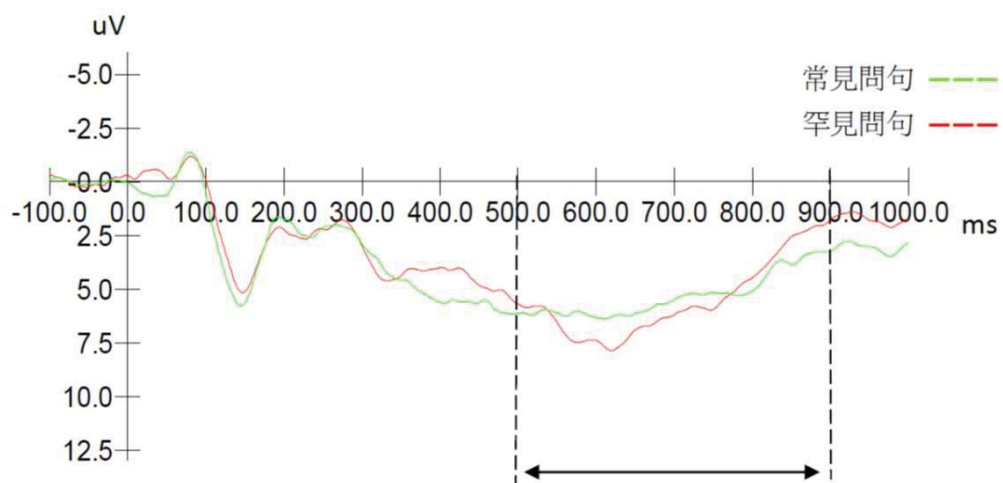
(N=153)

我們選取 F3、FZ、F4、C3、CZ、C4、P3、PZ、P4 等 9 個電極點位置的事件相關腦電位之平均腦波數值，進行統計分析。本研究之目標訊號標記(trigger)設在操控項(問句)出現時，從操控項呈現開始為目標訊號標記，以目標訊號標記之後的 500~900 毫秒之平均電位，做為觀察 P600 效應之時間區段(Time window)。

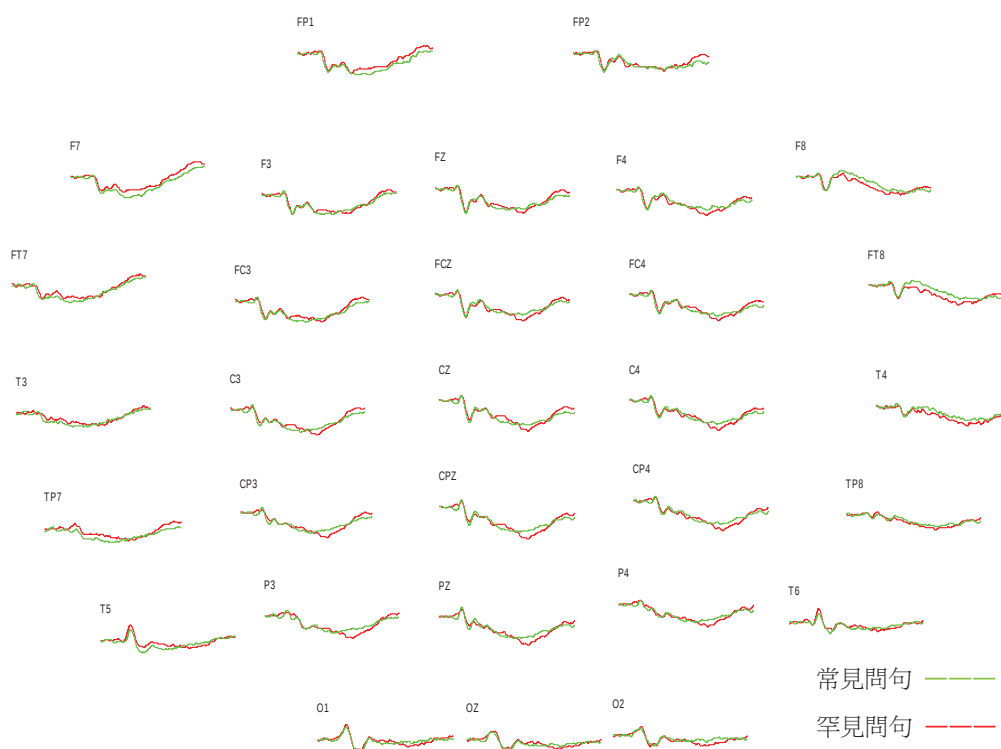
本研究以統計軟體 SPSS 18 進行統計分析，以三種不同問句類別(常見問句、罕見問句、不合理問句)，進行單因子相依樣本變異數分析(Repeated measured ANOVA)，在統計檢定上主要關注因為不同問句類別所誘發出的成分波(component)的振幅(Mean Amplitude)、潛伏時間 (Latency) 及峰值振幅 (Peak Amplitude)，以觀察不同問句類別在平均腦波電位是否達到顯著差異，實驗數據若違反 Mauchly 球形檢定假設，組內效應則以 Geisser-Greenehouse 作為校正結果，若因子之主要效果達到顯著水準，統計數值達顯著之變項則以 partial η^2 呈現其效果量，並以 Bonferroni 進行事後比較檢定分析，所有統計考驗顯著水準訂為 $\alpha=0.05$ 。

我們依不同問句類型所反映出的腦波波段(EPOCH)分別進行疊加平均 (Average)，並將三個實驗操控項進行比較，可以觀察到從目標訊號標記(trigger)起始後的 500~900ms 之間(圖上標示虛線之間的時間區段)，相較於常見問句的平均腦波電位，罕見問句及不合理問句的腦波皆呈現正走向之趨勢(參閱圖四之一 ~ 圖四之四)。

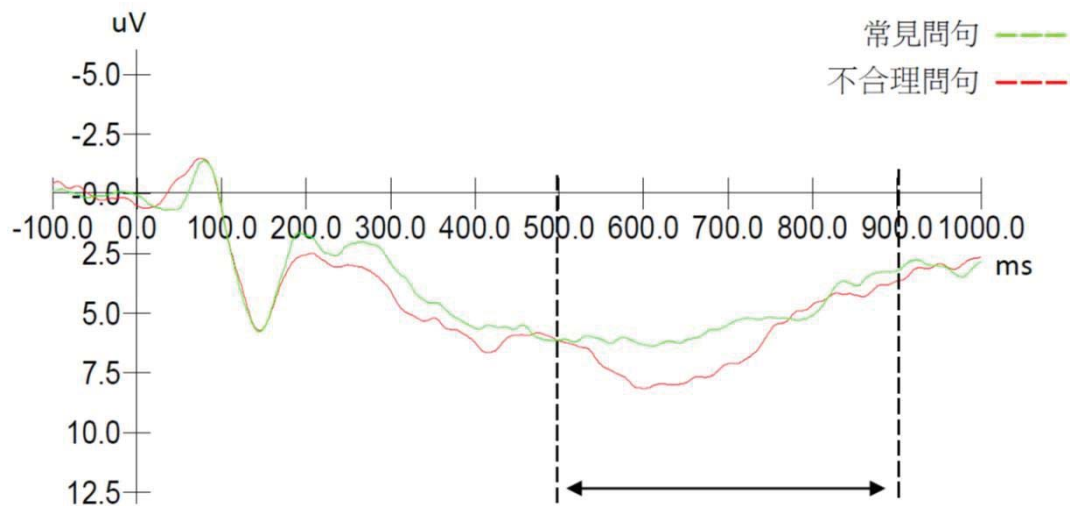
我們經參考文獻探討的結果並透過觀察平均腦波電位資料之後，選取目標訊號標記之後的 500~900 毫秒之平均電位作為觀察 P600 效應之時間區段(Time window)。此時間區段之平均振幅 (Mean Amplitude)、潛伏時間 (Latency) 與峰值振幅 (Peak Amplitude)平均數與標準差，參見表三。



圖四之一 常見問句與罕見問句在電極點:CZ 之平均腦波電位圖



圖四之二 常見問句與罕見問句之平均腦波電位圖



圖四之三 常見問句與不合理問句在電極點:CZ 之平均腦波電位圖



圖四之四 常見問句與不合理問句之平均腦波電位圖

我們以單因子相依樣本變異數分析(Repeated measured ANOVA)，分析不同問句類型在時間區段間的平均振幅、潛伏時間與峰值振幅之差異。結果發現，不同問句類型在平均振幅上之主要效果達顯著的差異($p < .001$)，效果量為低度關聯強度($F(1.708, 259.596) = 8.228, p < .001, \text{partial}\eta^2 = .051$)。經事後比較檢定顯示，常見問句($M=3.641, SD=2.946$)與罕見問句($M=3.980, SD=3.078$)之間沒有顯著差異($p > .05$)；不合理問句($M=4.738, SD=3.656$)顯著高於常見問句($p < .01$)；不合理問句顯著高於罕見問句($P < .05$)。

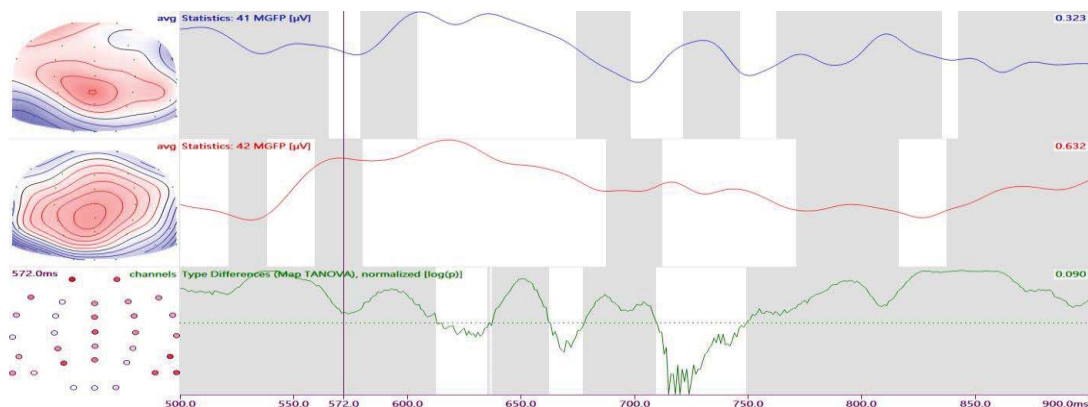
不同問句類型在潛伏時間之主要效果有達到顯著的差異($p < .001$)，效果量為中度關聯強度($F(1.821, 276.859) = 16.686, p < .001, \text{partial}\eta^2 = .099$)，經事後比較檢定顯示，常見問句($M=679.379, SD=79.642$)顯著高於罕見問句($M=643.667, SD=46.427$) ($p < .001$)；常見問句顯著高於不合理問句($M=662.190, SD=64.270$) ($p < .05$)；不合理問句顯著高於罕見問句 ($P < .001$)。

不同問句類型在峰值振幅上主要的效果達到顯著的差異($p < .001$)，效果量為中度關聯強度($F(1.709, 259.696) = 19.253, p < .001, \text{partial}\eta^2 = 0.112$)，經事後比較檢定顯示，罕見問句

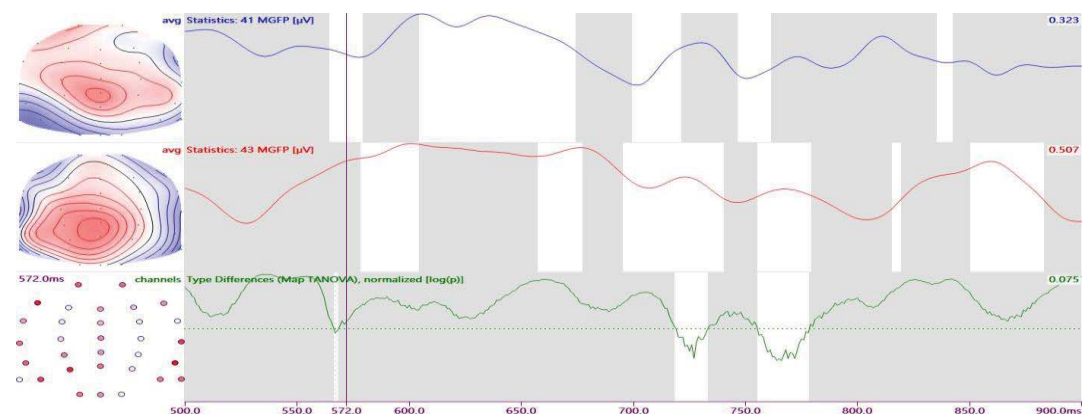
($M=7.384, SD=3.440$)顯著高於常見問句($M=6.244, SD=3.143$) ($p < .001$)；不合理問句($M=8.153, SD=4.450$) 顯著高於常見問句 ($p < .001$)；不合理問句顯著高於罕見問句($P < .05$)。

(二) 腦地形圖變異數分析(TANOVA)結果

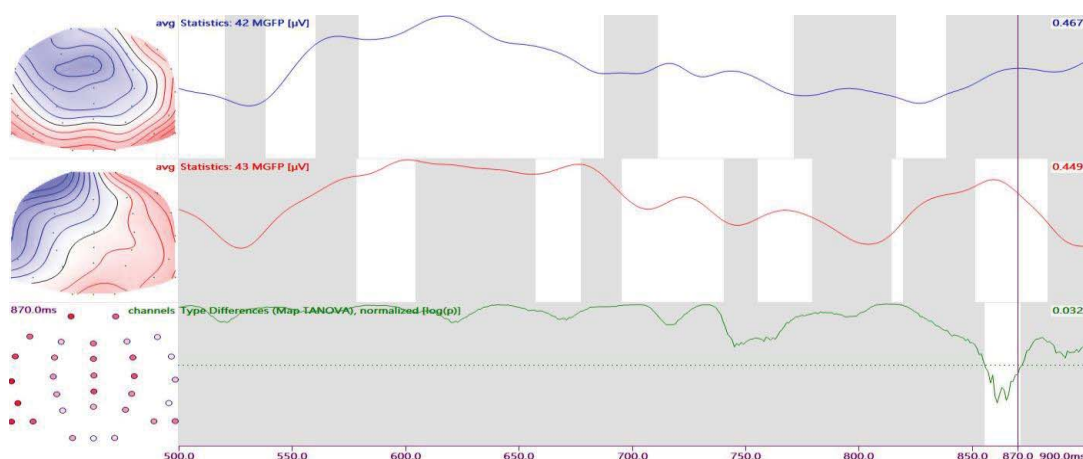
腦地形圖變異數分析(Topographic Analysis of Variance, TANOVA)，是以多重比較校正分析法，確定在所設定的時間區段中，各操控項之間平均潛伏時間和腦地形圖中主要成分波的潛伏時間，是否呈現顯著的差異。本研究運用 CURRY 8 中內建的 TANOVA 分析套件進行統計分析，相應的所需數字重複次數設定為 $R = 50 / p = 1000$ ，用於差異檢驗的腦地形圖先均一化，使每個取樣的腦地形圖 $MGFP=1$ ，統計分析之時間區段為 500ms 到 900ms，顯著差異值為 $p < 0.05$ 。由 TANOVA 分析結果，我們可以發現，常見問句與罕見問句之間有 613 ms ~ 635 ms、663 ms ~ 677 ms、710 ms ~ 749 ms 等三個時間區段，呈現顯著的差異(圖五)。常見問句與不合理問句之間有 719 ms ~ 733 ms、756 ms ~ 778 ms 等兩個時間區段，呈現顯著的差異(圖六)。罕見問句與不合理問句之間有 856 ms ~ 871.0 ms 時間區段，呈現顯著的差異(圖七)。



圖五 常見問句與罕見問句之 TANOVA 分析之差異結果分布圖。



圖六 常見問句與不合理問句之 TANOVA 分析之差異結果分布圖。



圖七 罕見問句與不合理問句之 TANOVA 分析之差異結果分布圖。

伍、討論

我們針對數學文字題題幹組所搭配的問題句進行操控，使用三種不同的問題句類型進行搭配，包括常見問題句、罕見問題句、不合理問題句，觀察實驗參與者對於搭配不同問題句類型所誘發的事件相關腦電位，在腦波波型上是否有顯著的差異。本研究以統計軟體 SPSS 18，將三種不同問題句類別(常見問題句、罕見問題句、不合理問題句)之作答反應時間(ms)，進行單因子相依樣本變異數分析(Repeated measured ANOVA)，結果發現反應時間的主要效果沒有達顯著的差異($P>.05$)。相較於過去相關文獻實驗的結果，本實驗的作答反應時間並沒有出現實驗的效應，我們的討論可能因為本實驗設計是以延遲反應的按鍵方式，使得實驗參與者對於各實驗操控項的反應時間上產生天花板效應的限制，因此，本實驗結果在反應時間的主要效果上沒有達到顯著的差異。

小學數學文字題中的添加型問題，是由已知的起始量，搭配使起始量增加的添加型改變量，以便求出增加後的結果量的答案，此基本算式為 $a+b=c$ 。(a=起始量、b=添加型改變量、c=結果量)。以本實驗所設計的施測語料之語句範例及呈現語序如下：

合理問題句：「阿明有 15 顆彈珠，小強再給他 7 顆，請問阿明現在有幾顆彈珠？」，其語序為：「起始量/改變量/詢問結果量」；罕見問題句：「阿明有 15 顆彈珠，小強再給他 7 顆，請問小強少了幾顆彈珠？」，其語序為：「起始量/改變量/詢問改變量」；不合理問題句：「阿明有 15 顆彈珠，小強再給他 7 顆，請問阿明和小強相差幾顆彈珠？」，其語序為：「起始量/改變量/詢問差異量」。

從本研究的結果發現(參閱表七)，罕見問題句之波段腦波的峰值呈現大於常見問題句的波段腦波的正走向，但是二者在平均振幅上沒有顯著的差異，而在潛伏時間及峰值振幅上呈

現極為顯著的差異。以 TANOVA 分析之結果，有 613 ms ~ 635 ms、663 ms ~ 677 ms、710 ms ~ 749 ms 等三個時間區段呈現顯著的差異。不合理問題句的波段腦波呈現大於常見問題句的波段腦波的正走向，二者在平均振幅、潛伏時間及峰值振幅上皆呈現顯著的差異。以 TANOVA 分析之結果，有 719 ms ~ 733 ms、756 ms ~ 778 ms 等兩個時間區段呈現顯著的差異。因此我們發現，當實驗參與者接收到的問題句訊息，不是問結果量未知的提問搭配形式(pattern)，例如「請問阿明現在有幾顆彈珠？」，而是詢問改變量的問題類型，例如「請問小強少了幾顆彈珠？」，或是屬於詢問差異量的問題類型，例如「請問阿明和小強相差幾顆彈珠？」便會誘發出 P600 效應。

此外，罕見問題句與不合理問題句的波段腦波，在平均振幅、潛伏時間及峰值振幅上皆呈現顯著的差異，以 TANOVA 分析之結果，有 856 ms ~ 871.0 ms 時間區段呈現顯著的差異，表示罕見問題句與不合理問題句的波段腦波皆誘發了 P600 效應，都是不符合跨語句的公式化語序的操控項，但兩者在 P600 效應的幅度上，還是有層次上的不同。

從實驗中觀察到的 P600 效應，我們推論當實驗參與者對所呈現的條件句例如「阿明有 15 顆彈珠，小強再給他 7 顆」進行語句理解後，解讀出條件句所提供的條件訊息有一個起始量「阿明有 15 顆彈珠」以及一個改變量「小強再給他 7 顆」，在進行語句理解的認知運作時，會啟動跨語句公式化語序的相關自動化歷程，同步協助閱讀理解的進行。在數學文字題中的添加型問題中，語序成分的搭配依序為「起始量、改變量、結果量」，我們推論實驗參與者，以此跨語句公式化語序的整體形式，做為理解題目的預設值(default)，因此預設之問句為詢問結果量的「阿明現在有幾顆彈珠？」，但是當實驗參與者所接受到的刺激訊息，是詢問改變量的「小強少了幾顆彈珠？」或是屬於詢問差異量的「阿明和小強相差幾顆彈珠？」，與其預設之公式化語序無法相容(match)，與題目

的情境(context)語意的理解，產生整合上的衝突而需要進行再確認，因此誘發出顯著的 P600 效應。

表七 不同問句類型之間在平均振幅、潛伏時間與峰值振幅的差異顯著性

	平均振幅	潛伏時間	峰值振幅
常見問句 VS 罕見問句		***	***
常見問句 VS 不合理問句	**	*	***
罕見問句 VS 不合理問句	*	***	*

*: P < 0.05 (significant), **: P < 0.01 (highly significant); ***: P < 0.001 (extremely significant)

我們所觀察到的 P600 效應與 Shen et al. (2016)所提出的觀點相近，認為語意 P600 反映的是一種更為廣泛的衝突監控(conflict monitoring)機制，Shen et al. (2016)發現，當大腦的前扣帶皮層(anterior cingulate cortex, ACC)偵測到的語言輸入和人們所預期的不一致，產生認知監控上的衝突時，便會誘發出顯著的 P600 效應，導致為了檢索長期記憶中的世界知識(world knowledge)，以便重新解讀相關語句所產生的語意，因而大腦的右前額葉(RaPFC)等區域也會被激活。進而我們推論，跨語句之公式化語序的心理真實性是存在的，且當人們在語句理解的認知運作歷程時，應是多工、多層次平行處理且相互影響，是相當複雜的運作歷程。

陸、結語

本研究運用高時間解析度的事件相關腦電位(ERPs)研究法，以數學文字題中的添加型問題為實驗材料，對於大腦在語言理解之運作歷程進行觀察。小學數學文字題在基本的呈現形式與其所相對應的運算意義上，是跨越句子(條件句)與句子(問句)之間在相互的搭配上有相關程度的關聯性，本研究之結果發現當問句的訊息與條件句所提供的訊息意義在整合上產生困難，或是與解題者從長期記憶所提取出來的訊息不一致(違反)，與其預設之提問搭配形式(pattern)無法相容(match)時，或是因為實驗操控項是一般很少見乃至於不曾出現的樣態，便會產生 P600 的效應，此結果與 Shen et al. (2016)提出的衝突監控(conflict monitoring)機制的觀點以及 Fisher et al. (2010)發現當詞彙語意範疇(目標詞)不一致時，會誘發顯著的 P600 效應，做為數學文字題中的結構一致性的指標，可以互為參照。

我們在相關文獻的基礎上，進一步對於「跨語句之公式化語序」以整體形式儲存在腦中的心理真實性進行探討，運用事件相關腦電位之腦電波成分 P600 為參照指標，從實驗結果觀察到語句理解運作時，各操控項與常見問句相比較，罕見問句與不合理問句在 500ms~900ms 的時間區段之間的平均振幅、潛伏期和峰值振幅上皆呈現出顯著的 P600 效應，因

此我們推論語句理解運作時，跨語句之公式化語序的心理真實性是可能的。我們建議未來的相關研究，可以擴展至不同的數學文字題類型或是討論不同年齡層的實驗參與者例如學齡兒童...等，或是將大腦的語句理解運作歷程進行溯源分析(source reconstruction)，以將研究的基礎加深加廣，也希望藉由本研究結果，對於語句理解的認知運作相關理論，提供一些不同角度的思考。

參考文獻

Erman, B., & Warren, B. (2000). The idiom principle and the open choice principle. *Text*, 20(1), 29-62. doi:10.1515/text.1.2000.20.1.29

Fisher, K., Bassok, M., & Osterhout, L. (2010). *When two plus two does not equal four: event-related potential responses to semantically incongruous arithmetic word problems*. Paper presented at the Proceedings of the Annual Meeting of the Cognitive Science Society.

Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in cognitive sciences*, 6(2), 78-84.

Hagoort, P. (2003). How the brain solves the binding problem for language: a neurocomputational model of syntactic processing. *Neuroimage*, 20, S18-S29.

Kim, A., & Osterhout, L. (2005). The independence of combinatory semantic processing: Evidence from event-related potentials. *Journal of memory and language*, 52(2), 205-225.

Kuperberg, G. R. (2007). Neural mechanisms of language comprehension: Challenges to syntax. *Brain research*, 1146, 23-49.

Kuperberg, G. R., Holcomb, P. J., Sitnikova, T., Greve, D., Dale, A. M., & Caplan, D. (2003). Distinct patterns of neural modulation during the processing of conceptual and

- syntactic anomalies. *Journal of Cognitive Neuroscience*, 15(2), 272-293.
- Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Annual review of psychology*, 62, 621-647.
- Nieuwland, M. S., & Van Berkum, J. J. (2005). Testing the limits of the semantic illusion phenomenon: ERPs reveal temporary semantic change deafness in discourse comprehension. *Cognitive Brain Research*, 24(3), 691-701.
- Schmitt, N., & Carter, R. (2004). Formulaic sequences in action. *Formulaic sequences: Acquisition, processing and use*, 1-22.
- Shen, W., Fiori-Duharcourt, N., & Isel, F. (2016). Functional significance of the semantic P600: evidence from the event-related brain potential source localization. *NeuroReport*, 27(7), 548-558.
- Van Herten, M., Chwilla, D. J., & Kolk, H. H. (2006). When heuristics clash with parsing routines: ERP evidence for conflict monitoring in sentence perception. *Journal of Cognitive Neuroscience*, 18(7), 1181-1197.
- Vissers, C. T. W., Chwilla, D. J., & Kolk, H. H. (2007). The interplay of heuristics and parsing routines in sentence comprehension: Evidence from ERPs and reaction times. *Biological Psychology*, 75(1), 8-18.
- Wray, A. (2004). Here's one I prepared earlier. *Formulaic sequences: Acquisition, processing, and use*, 9, 249.
- Wray, A. (2005). *Formulaic language and the lexicon*: Cambridge University Press.

The Psychological Reality of The Formulaic Sequences of Cross-Sentence : An Event-Related Potential Study of Math Word Problem

Yao, Tien-Ching 、 Lu, Ching- Ching

National Tsing Hua University Institute of Taiwan Languages and Language Teaching.

Abstract

The aims of the present study were to use the experimental research method of event-related brain potentials to investigate the psychological reality of the “Formulaic Sequences of Cross-Sentence”. Therefore, we take the math word problems as the experimental materials and use the ERPs' P600 component to verify. In this study, the statistical analysis shows a significant P600 effect of the average brainwave waveform induced by different question types between 500ms to 900ms. Therefore, the psychological reality of the formulaic sequences of cross-sentence is possible. We can provide some different perspectives on the theory of cognitive operation of sentence comprehension by the results of this study.

Keywords: language processing, sentence comprehension, Formulaic sequences, event-related potentials