

南 華 大 學

資訊管理學系

碩士論文

語音使用者介面設計之互動式多媒體資訊站

-以博物館導覽為例

Voice User Interface at Kiosk in Museum Navigation



研 究 生：蔡明勳

指導教授：楊聰仁

中華民國 97 年 6 月 30 日

南 華 大 學

資訊管理研究所

碩 士 學 位 論 文

語音使用者介面設計之互動式多媒體資訊站

-以博物館導覽為例

研究生：蔡明勳

經考試合格特此證明

口試委員：

簡文良
邱宏村
楊聰仁

指導教授：張聰仁

系主任(所長)：鍾國貴

口試日期：中華民國 97 年 6 月 13 日

誌 謝

本篇論文能夠順利完成，主要仰賴指導教授-楊聰仁老師的細心指導，在每一次的討論過程中，老師皆提供我論文方向與其他的不同想法，讓我可以從另一個方向去詳加思考，使的本篇論文的方向更加確定與縝密，在南華大學經過兩年的碩士班生活，特別感謝成均館 C208 的各位夥伴，宗亨、怡君、劉董，宗亨那搞笑功力真的是無人可比，也帶給我們實驗室無限的歡樂，也感謝學弟躍進、士峰的幫忙，其它實驗室 C222、C204 的光南、宣均、淳淳，育弘學長，宵夜團的我們就會集合一起出門去減壓，華山、半天岩都有我們的蹤跡，也感謝我的女朋友瑞琪，在這兩年來辛苦的陪伴著我，在我最需要幫忙時候伸出援手，最後，也是最感謝的人，就是我的父母，他們提供我在嘉義的所需，回家後也提供我一個可以休息的避風港，也感謝在本篇論文進行中有所貢獻的每一位朋友，謝謝你們。

語音使用者介面設計之互動式多媒體資訊站

-以博物館導覽為例

學生：蔡明勳

指導教授：楊聰仁

南 華 大 學 資 訊 管 理 學 系 碩 士 班

摘 要

本篇論文主要在探討使用者對於現有觸控式多媒體互動系統與加入語音辨識後的多媒體互動系統的差異性，並將系統運用在博物館導覽上面，系統透過 Microsoft SAPI 語法的撰寫與 XML 文法的設計，本研究所設計的語音介面改善了觸控式多媒體互動系統上所沒有的服務，電腦可以利用語音合成技術發出聲音(TTS)、改善使用者短暫記憶的缺失(STM)、使用者可以隨時說出求救(Help)的指令來幫助瀏覽…等相關功能，並在系統完成之後做實驗以分析使用者對於系統的使用滿意度與效能，透過實際的系統測試與事後問卷調查，藉由分析結果來討論是否加入語音辨識後效能會更佳，與使用者對於新的語音使用者介面的使用滿意度高與低。因此，本研究提出了三個研究假

設，透過系統的實測與問卷的調查，分別就各個假設去驗證，以證明我們所設計的語音使用者介面系統可以讓使用者快速且容易的取得博物館的相關資訊。

關鍵字：語音辨識、語音使用者介面、博物館導覽、XML 文法設計

Voice User Interface at Kiosk in Museum Navigation

Student : Tsai Ming-Shun

Advisors : Dr. Yang Tsung-Jen

Department of Information
Management
Nan-Hua University

Department of Accounting &
Information
Nan-Hua University

ABSTRACT

Speech Recognition has been developed for many years. It becomes popular in the recent years to apply speech technology in information systems. In this paper, we design a new user interface within voice recognition to navigate museum. We use Microsoft Speech SDK to be engine and XML Grammar designed to have higher recognition rate. We also improve touch to voice control to make navigation smoothly and design XML Grammar to make user be able to speak Natural Language. It shows that the voice user interface navigation system provides a more user friendly interface and enhances the effectiveness and efficiency for museum navigation. We test linear and non-linear tasks of wasting time after finishing tasks. By the test, we wish that voice user interface can reduce wasting time and navigate easily.

We design the system and evaluate the account to people. We wish that people be able to navigate and learn on Kiosk within voice

recognition. Otherwise, focus on getting information and making navigation easily and friendly are the important things.

Keywords : Speech Recognition, User Interface, Museum Navigation,
XML Grammar

目 錄

書名頁	i
博碩士論文授權書.....	ii
論文著作財產權同意書.....	ii
論文指導教授推薦書.....	iii
論文口試合格證明.....	iv
誌謝.....	v
中文摘要.....	vi
英文摘要.....	viii
目錄.....	x
表目錄.....	xii
圖目錄.....	xiii
第一章、緒論	1
1.1 研究背景：	1
1.2 研究動機：	2
1.3 研究範圍與限制：	4
1.4 研究步驟：	4
1.5 論文架構：	7
1.6 研究假設：	9
第二章、文獻探討	10
2.1 語音辨識：	10
2.2 語音合成發音.....	14
2.3 XML 文法設計：	19
2.4 語音使用者介面：	23
2.5 互動式多媒體資訊站：	28

2.5.1 Kiosk 的 SWOT 分析.....	30
2.5.2 Kiosk 介面需要的功能.....	32
2.6 線性與非線性任務：.....	33
2.7 使用者滿意度：.....	36
第三章、研究方法	39
3.1 系統設計與組成架構：.....	39
3.2 可延伸標記語言設計：.....	44
3.3 開發工具：.....	46
3.4 研究設計：.....	46
3.4.1 第一階段研究：.....	47
3.4.2 第二階段研究：.....	56
第四章、系統實作與結果分析	57
4.1 系統實作：.....	57
4.2 時間效能檢定：.....	60
4.3 問卷分析：.....	72
第五章、結論	84
第六章、問題與討論	87
參考文獻	91
一、中文部份.....	91
二、西文部份.....	91
附錄	93
附錄一：語音(觸控式)使用者滿意度調查之問卷題目	93

表 目 錄

表 2.1：語音合成與預錄聲音之比較.....	15
表 2.2：Pearson 與 Bailey 提出的 39 個問卷問項	37
表 3.1：使用者放聲思考法紀錄表.....	52
表 3.2：使用者意見整理表.....	55
表 4.1：任務規劃表.....	62
表 4.2：時間整理表(秒).....	65
表 4.3：非線性任務介面比較表.....	68
表 4.4：兩種介面所需平均時間(秒).....	69
表 4.5：觸控式輸入與語音輸入的 Z 檢定分析	71
表 4.6：問卷各個問項的平均值.....	74
表 4.7：系統特性檢定表.....	78
表 4.8：使用者介面檢定表.....	80
表 4.9：使用者介面的操控性檢定表.....	81
表 4.10：使用者介面的使用感覺檢定表.....	82
表 6.1：建置 Kiosk 需要的成本表	88

圖 目 錄

圖 1.1：研究流程圖.....	5
圖 1.2：使用者任務分組圖.....	6
圖 2.1：類比訊號轉換成數位訊號圖.....	11
圖 2.2：SAPI 架構圖[1].....	14
圖 2.3：XML 與 SAPI 之間的關係[21].....	19
圖 2.4：Grammar 設計圖[17].....	20
圖 2.5：XML 文法設計圖.....	22
圖 2.6：系統介面導覽圖.....	24
圖 2.7：使用者使用 Help 指令圖[16].....	25
圖 2.8：使用者使用線性選項.....	25
圖 2.9：線性任務下的選擇模式.....	33
圖 2.10：非線性任務下的選擇模式.....	34
圖 3.1：使用者資料流向圖.....	40
圖 3.2：網頁使用者提示圖[21].....	41
圖 3.3：系統等待時間圖.....	42
圖 3.4：博物館資訊簡圖.....	43
圖 3.5：系統發聲控制圖.....	44
圖 3.6：語音路徑選擇圖一.....	44
圖 3.7：語音路徑選擇圖二[21].....	45
圖 3.8：第一階段實驗流程圖.....	49
圖 3.9：使用者語音訓練圖.....	50
圖 4.1：博物館導覽資訊圖.....	58
圖 4.2：任務選擇圖.....	58
圖 4.3：XML 文法選擇圖.....	60

圖 4.4：使用者分組與執行任務圖.....	61
圖 4.5：非線性任務所需時間圖.....	66
圖 4.6：觸控式輸入與語音輸入的平均花費時間圖	70

第一章、緒論

本章節主要分為研究背景、研究動機、研究範圍與限制、研究步驟、論文架構、研究假設等六個小節，茲分別說明如各節。

1.1 研究背景

語音辨識的目的在於讓電腦可以聽懂人們的語句輸入，當電腦可以聽懂人們說的話之後就可以命令電腦執行相對應的工作。當聲音藉由類比訊號轉變到數位訊號裝置輸入電腦內部，並以數值方式儲存後，語音辨識程式便開始已事先儲存好的聲音樣本與輸入的測試聲音樣本進行比對工作。比對完成後電腦收集並比對一個它認為最“像”的聲音樣本序號，我們就可以知道使用者剛剛唸進去的聲音代表什麼意思，進而透過聲音命令電腦做事。

而博物館導覽系統的開發已經行之有年了，從早期的紙本 DM 發放、志工協助導覽、耳機配合錄音帶導覽，轉變成以電腦或是互動式多媒體資訊站(Kiosk)為主的數位導覽方式，提供使用者資訊的查詢與服務、館區的介紹與導覽服務，作品的說明簡介等。根據黃光男先生在「博物館行銷策略」[8]一文中提出，博物館是一個獨立完整的運作空間，博物館的參觀族群，涵蓋了各種年齡、性別、族群與職業，

因此在此環境下提供的一種自由性的、自願性的學習活動。博物館不像是學校機關有課程規劃學習，在沒有課程制度下，任何一個活動都需要符合三個原則：

- 1、自由性：兼顧到大眾各個階層的需求。
- 2、開放性：讓民眾在進入參觀時候，感受到自我學習的目的。
- 3、經驗性：環境的佈置，活化參觀者的經驗累積。

另外，博物館也是一個具有溝通、資訊與休閒目的的場所，是民眾終生學習的地方。所以我們在設計博物館導覽系統的時候，也要考慮到民眾在使用導覽系統時候能不能確實得到所要的資訊服務，在瀏覽時候流程控制順不順暢，利用語音辨識 Kiosk 當系統平台，是否可以將所要傳達的訊息傳達給使用者知道...等問題，所以這之間包含的範圍相當廣泛，我們分別在以下作深入討論。

1.2 研究動機

在人工智慧的運用上，有很多東西原本都只是一個理念，由於電腦硬體的功能越趨強大，原本只是理念的想法也逐漸被慢慢實現了，語音辨識技術的應用也是其中一例，試想，如果可以對著電腦直接利用語音下達命令，不需透過鍵盤滑鼠等輸入裝置，將可大大的提升友善度，減少一般使用者對於對腦使用的障礙。對於聽力障礙的人來說，直接透過語音辨識技術的應用，將所需資訊等服務用電腦合成語

音(TTS , Text To Speech)說出，也可以讓使用者在訊息接收上有另一種新的選擇。

近年來隨著資訊化程度越來越高，很多政府機構都導入資訊化的服務，如建置服務網頁、資料處理電腦化、建置 Kiosk 導覽系統(互動式多媒體資訊站)等，讓一般使用者對於欲知道的資訊取得更容易，也可以節省機構對於人事方面的開銷，另外，Kiosk 系統可應用在許多的場合，例如：博物館資訊導覽、旅遊資訊的導覽、機場劃位、訂票服務等，本篇研究主要是探討將語音辨識技術加入博物館資訊的導覽之後，產生出一個新的使用者介面(UI ,User Interface)，並與現有的觸控式 Kiosk 系統作效能與使用者滿意度之比較，看加入語音控制後的介面是否有比原有觸控式使用者介面更易於使用。

使用者在參觀博物館時候，往往需要參考資料來搭配參觀的景點，通常博物館會在旁邊配置一台觸控式 Kiosk 系統來導覽使用者，使得使用者可以透過此設備來瀏覽整個博物館地理位置，並取得使用者所需要的博物館的資訊。但是 Kiosk 一直以來使用率都不高，相關研究顯示出，大多是宣導不足與使用習慣的問題，因此，我們在原有系統上加入了語音辨識的功能，藉此增加並吸引使用者對於系統的好奇，另外，此研究之目的在於在探討現有 Kiosk 系統加入語音辨識之後，使用者對於系統的使用滿意度與效率是否有比原有系統更佳，所以系統建置完成後，會讓使用者測試其功能與方便性，透過實驗後的問卷調查使用者對於語音導覽系統的滿意度與其優缺點，進而改善系統，達到系統之易用性。

此研究以現有博物館 Kiosk 系統-科技台灣驚嘆號[7] 為平台，加

入語音辨識之技術，來探討使用者對於現有 Kiosk 系統與加入語音辨識之 Kiosk 系統的滿意度調查，與加入語音辨識後，觸控式系統與語音辨識系統完成一個任務所使用的時間比較，並透過問卷方式調查使用者對於語音使用者介面的接受度如何，與觸控式的介面作相互比較，並利用統計方式取得使用者對於語音使用者介面的滿意度，在本研究我們分別針對效度與滿意度提出三個假設，並在之後的章節會加以驗證。

1.3 研究範圍與限制

在著手進行使用者介面與設計時候，我們為了避免背景噪音影響語音辨識的結果，所以我們在進行相關實驗的時候，我們選擇噪音量較少的南華大學 C208 實驗室來進行實驗，且因為 Kiosk 系統的建置需花費許多資源，所以我們以華碩 A3H 筆記型電腦來設計使用者介面並以 HP 觸控面板電腦來代替 Kiosk 系統來模擬測試時的情況，以解決建置 Kiosk 所帶來的花費問題。另外，在本研究當中，我們以限定範圍的方式來提升語音辨識的辨識率，利用 XML 文法設計並限定在博物館導覽的資訊上，並非完全利用自然語言的處理方式來完成本研究，所以當使用者說出預期之外的語句，系統將會無法做相關處理。

1.4 研究步驟

如下圖 1.1 所表示，一開始先確認欲將語音辨識技術運用在哪

邊，經過研讀相關文獻後發覺，現有觸控式 Kiosk 導覽系統使用率不高之問題，所以欲將語音辨識技術導入，希冀可以讓使用者對於新的語音 Kiosk 導覽系統能有更多的興趣。

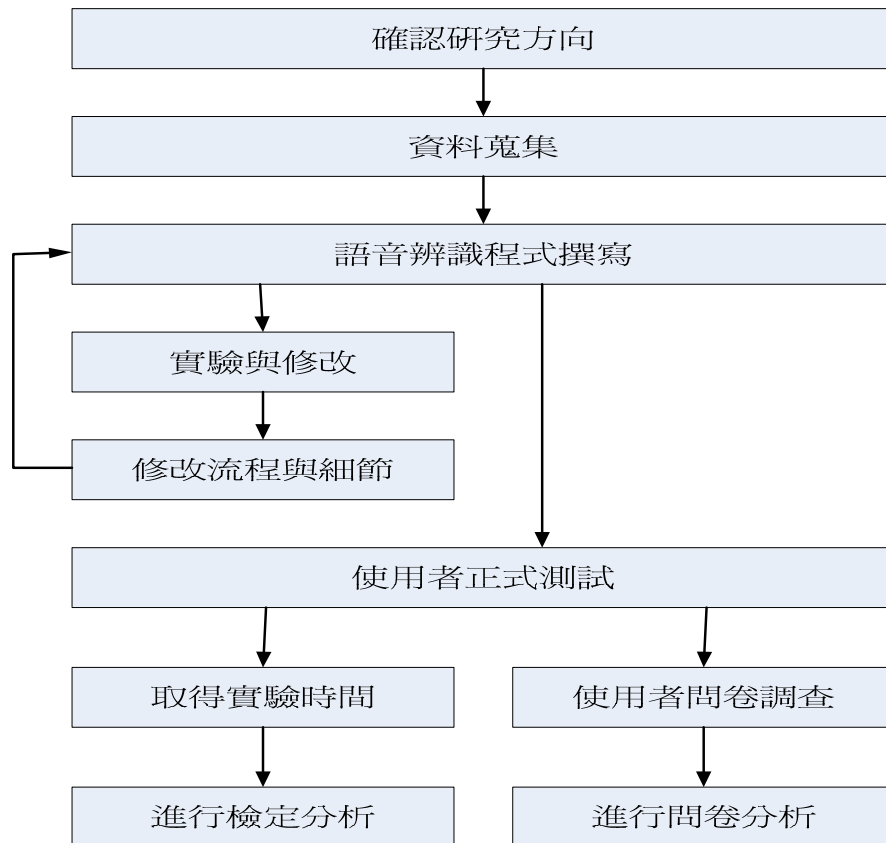


圖 1.1：研究流程圖

確認方向後，並開始收集相關文獻，許多學者對於使用者介面有提出許多看法，這些在文獻探討內的語音使用者介面內會有所說明，另外也有收集一些國外 Kiosk 系統設計的觀念文獻，將以上資料整理後便著手開始設計語音使用者介面之程式，經過不斷的嘗試錯誤與更正，語音導覽介面的雛型已經完成，接下來透過放聲思考法(Thinking aloud)來檢驗此系統，先將系統雛型給予受測者作第一次的測試，程

式設計者在身旁提問，讓受測者根據測試果並結合自身的經驗與想法，把系統之優缺點告知程式設計者，根據受測者的意見進而更改系統本身的流程或增加細節，把系統完整度更加提升以便達到容易使用的結果。

經過不斷修改後，將系統給予受測者測試，取得系統時間值並作 Z 檢定分析，並在使用者測試完系統之後給予問卷以方便調查使用者對於系統的滿意度高低，取得實驗時間與問卷後，會將數據資料等作相關分析以證明使用語音使用者介面比傳統觸控式使用者介面易於使用且時間效率較高。

透過統計分析的方式來驗證效率好壞，如下圖 1.2，透過使用者分組分別測試語音與觸控式在線性任務與非線性任務之差別，目的在於了解加入語音使用者介面之後，語音使用者介面在線性任務或是非線性任務可以讓語音辨識技術發揮最大效用。

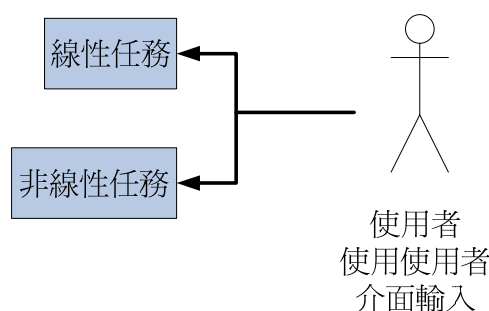


圖 1.2：使用者任務分組圖

1.5 論文架構

本篇論文架構如下：

第一章：序論，介紹為什麼想從事語音辨識導覽系統之研究，語音辨識系統的大致開發過程與博物館導覽所需要的背景知識，並探討為什麼要實作一個博物館導覽語音使用者介面。

第二章：文獻探討，介紹語音技術的發展與其應用，如語音辨識的歷史與其發展，語音合成技術的應用、XML 語法設計理念、語音使用者介面設計需要注意的事項、多媒體互動站的應用、線性與非線性任務的定義與介紹與使用者滿意度的定義與規劃。

第三章：研究方法，介紹系統的組成架構與詳細的系統功能，並利用放聲思考法來收集使用者對於系統的使用觀感，透過使用者的親自使用並提出系統的優缺點改進方法，進而改善系統的缺失。

第四章：研究結果，根據使用者測試系統所得的結果來做比較，當收集完使用者使用系統後的時間資料，利用統計方法

Z 檢定來檢測語音使用者介面與觸控式使用者介面的優劣，並實驗語音使用者介面在線性或是非線性任務底下的效率，之後透過問卷設計取得使用者對於系統的滿意度調查。

第五章：問題與討論，討論本篇研究當中，所可能遭遇到的困境與真正建置後所可能遇到的困難點，並提出可能解決問題的方法。

第六章：結論，本篇論文的結語，並統整此論文的優點與語音使用者介面的好處，並希冀可以對未來研究的人們有一些貢獻。

1.6 研究假設

我們在這篇研究裡面提出三個假設，在後面的章節會針對以下各個假設提出證明：

假設 1：使用者使用語音導覽介面可以有更佳的時間效率，透過語句來控制瀏覽流程，語音的控制介面可以比觸控式使用者介面更節省取得資訊的時間。

假設 2：在非線性任務的安排底下，利用語音使用者介面瀏覽系統會比利用觸控式使用者介面瀏覽系統更加有效率。

假設 3：使用者使用完系統之後，對於語音使用者介面會比觸控式使用者介面有較佳的滿意度。

在使用者完成系統測試之後，我們讓受測者填寫問卷，透過問卷調查方式讓受測者針對語音導覽、觸控式使用者介面、線性任務與非線性任務等作滿意度的調查，並分析其原因。

第二章、文獻探討

本章將針對語音辨識(ASR , Automatic Speech Recognition)、XML 文法設計(XML Grammar)、語音使用者介面(VUI , Voice User Interface)、語音合成發音(TTS , Text To Speech)、互動式多媒體資訊站(Kiosk)、線性與非線性任務與使用者滿意度等七個相關議題進行介紹與討論。

2.1 語音辨識

利用電腦技術來處理語言問題，是現代科技中一項重要的發展，其最終目的就是希望人類可以與機器利用自然語言(Natural language)來做溝通，在許多科幻電影中常常可以看到機器人可以知道人們所表達的語句，進而去完成人們所要機器人去完成的目標，這其中牽扯到許多科技技術的應用，語音辨識技術也是其中一個重要的指標。利用裝置將類比的聲音轉變成數位訊號讓電腦可以讀取並作出指令，如下圖 2.1，每隔一段距離取得一組樣本，並將此類比樣本資料經過二進制轉換，變成數位訊號以便電腦讀取，經過轉換之後在將訊號送到辨識引擎裡面作機率分析，引擎會根據訊號的傳遞選出機率最高的辨識字元。

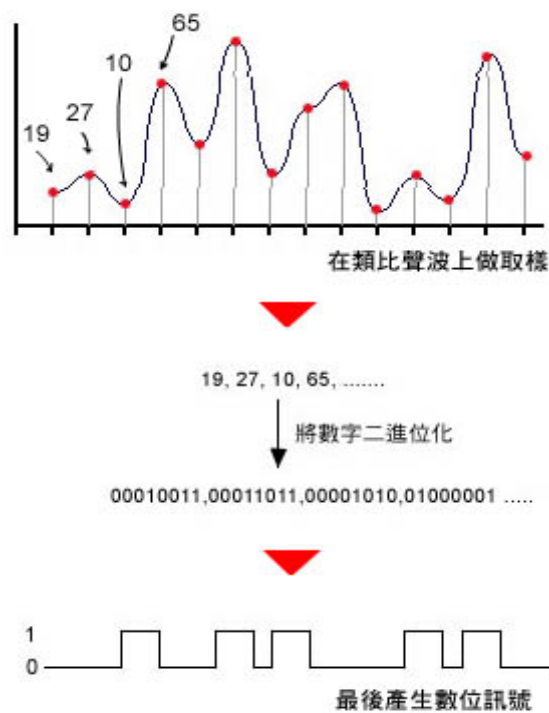


圖 2.1：類比訊號轉換成數位訊號圖

清華大學電機工程系 王小川 教授在「語音訊號處理」一書中針對語音辨識有著以下定義，凡是使用人類發聲器官產生帶有語音訊息的聲音，我們就稱為語音。以電腦技術處理語音訊號，就叫做語音訊號處理 (Speech signal processing)，或叫做語音處理 (Speech processing)，它也有另一個名詞，就是所謂的口述語言處理 (Spoken language processing)，它可以是與語言本身有關，也可能與語言沒有關係。例如語音編碼 (speech code)，其處理方式不會因為語言的不同而有所差異，又如語音增強 (Speech enhancement)，目的在於去除語音中的夾雜的噪音 (noise)，這種語音處理也不會因為語言的不同而有所

不同。有些語音處理就是與語言絕對相關，例如語音辨識(Speech recognition)，與語音合成(Speech synthesis)。因為不同語言有其特定的一組音位，所以就得需要定義特殊的一組語音單位(Speech unit)，作為語音辨識與語音合成的基本單位，同時也要考慮到其特殊的韻律(prosody)、構詞語文法[1]。

以目前有的語音辨識引擎而言，開發系統時候比較常用到的有 Microsoft Speech SDK[21]、IBM Via Voice、VoiceXML[24]與飛利浦的辨識引擎，通常會視需求來選擇應用的辨識引擎，因為採用 Visual Basic 來開發系統，而 Speech SDK 也有支援 Visual Basic 語法，因此採用兩者互相搭配的做法。將語音辨識之技術加入觸控式 Kiosk 導覽系統，實做一個新的使用者介面，讓使用者透過聲音來控制並取得所需要的服務資訊，但由於系統處理使用者的自然語言發音(NLP，Natural Language Processing)所需要的語料庫過於龐大，會造成辨識率(Recognition Rate)大幅下降，因此需定義其辨識的文法(XML Grammar)，將可能出現的關鍵字(Key words)輸入 XML Grammar 中，提升語音辨識過程中的辨識率。

現有大部分的語音辨識系統是採用隱藏式馬可夫模型(HMM，Hidden Markov model)方式來做為依據，它是以機率模型來描述發音

的現象，也就是說，語音辨識並不是電腦真的聽的懂人們所說的語句意思，而是透過統計方式來判斷哪一組文法對應到使用者發出的聲音機率最高，進而取那一段文法當作使用者本身所要表達的語句，本系統使用的 Speech SDK 也是利用 HMM 的模型方式辨識。針對語音辨識方面，研究使用 Visual Basic 配合 Microsoft Speech SDK 6.1 語音辨識引擎來撰寫，以現有的博物館導覽-科技台灣驚嘆號[7]為主題，加入語音辨識的技術提供使用者一個新的使用者介面。

根據 廖峻峰先生在“語音技術與 VoiceXML 應用”[2]一文中提到語音辨識(ASR)的處理，包含：

- 1、使用者輸入(User Input)：將使用者發出的聲音收集起來。
- 2、訊號處理(Digitization)：將類比訊號轉成數位訊號。
- 3、訊號轉換(Phonetic Breakdown)：將訊號轉換成音素。
- 4、統計模型(Statistical Modeling)：根據隱藏式馬可夫模型(HMM) 將音素對應到音標符號 (Phonetic Representation)。對應程式(Matching)：ASR 對應到程式開發者所寫的程式。

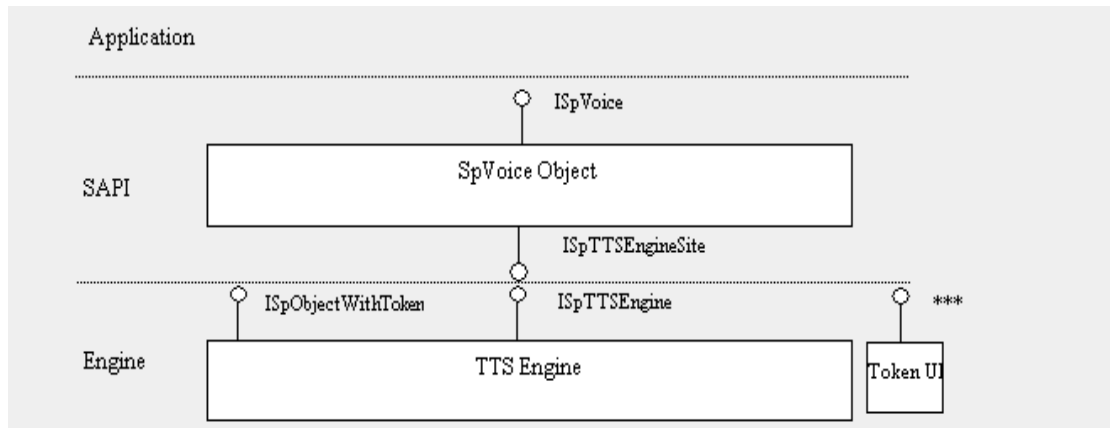


圖 2.2：SAPI 架構圖[1]

所以，根據以上列表並利用 SpeechSDK 進行系統開發，如上圖 2.2 所示，當使用者說出關鍵字之後，Application 端會將關鍵字等傳送到 SAPI 端作處理，如果需要電腦合成語音的話，TTS Engine 就會作動，將合成之後的聲音傳送到 SAPI 等待被使用，依照這個循環，程式與 SpeechSDK 引擎互相配合而完成現有的導覽系統。

2.2 語音合成發音

語音合成又名文句轉語音(TTS，Text-To-Speech)，是指將輸入的文字或儲存於電腦中的文件模擬人聲發出語音的技術，語音合成的技術發展時間比語音辨識早，但應用層面大多仍在閱讀電腦螢幕上的文章，語音指引，互動回饋，或輔助說明。

根據 Dirk Schnell et al.在 Audio Navigation Patterns [16] 文中指

出，語音使用者介面回覆使用者答覆的時候可以有兩種方法，一是使用預錄的聲音(Prerecorded Sound)、另一個是利用語音合成技術(TTS, Text To Speech)。使用預錄聲音可以預先錄製真人的發音，讓使用者介面回應時不會有呆版的機器聲，不過缺點是不能靈活應用，因為只要服務內容一改變，錄製的聲音檔案便不能使用，如果要再錄製的話，所花費的時間、人力、物力都會相當的多，所以本系統採用讓電腦自己合成聲音檔的方式，利用語音合成方法來發聲，雖然聲音聽起來沒有預錄的聲音優美，但是在符合經濟成本的考量下，可以隨時更改的語音合成比較適合。

表 2.1：語音合成與預錄聲音之比較

	語音合成發音	預錄聲音檔
優點	錄製成本較低	發聲較優美
缺點	機器聲發音	錄製成本較高

另外，語音合成技術還有以下優點：

- 1、發音的自然度(清晰、流暢)。
- 2、破音字的處理。
- 3、即時處理的能力。

所以透過語音合成技術的支援，可以減低系統開發的複雜度，而語音合成發音的做法有以下：[1]

1、頻譜參數合成方法(Articulatory Synthesis)：

如 Holmes 的並聯共振峰合成器（1973）和 Klatt 的串、並聯共振峰(Formant)（1980）合成、基於 LPC 等聲學參數的合成系統，但要合成出清晰的語音需要準確的設定參數，使用困難，且合成出的語音仍不夠自然。

2、波形拼接法(Formant Synthesis):如基頻同步累加法(PSOLA)

（1990）在語音波形上做時域(time domain)的韻律修正來合成語音，就可以產生出具有韻律的合成語音。 PSOLA 的設計重點，在改良頻域(frequency domain)耗時，以及在時域(time domain)接合效果太差的情形，其合成的語音在音色與自然度都大大的提升，且架構較簡單，容易實作。

3、串接合成法(Concatenated Synthesis)：

以一個錄好聲音的語料庫來當作比對的標的，從語料庫中抓出相對應的聲音單元，一些在 rule-based 與 knowledge-based 方法下需要做細節的聲韻調整也因此減少了許多，如此簡化了計算拼接與口音等複雜的計算，也特別適合在少量字彙的輸出時使用。

對於 TTS 系統而言，無論接受的是一段文字的輸入或是一篇文章，這些文字本身並沒有包含任何聲學特性（說話的聲調，停頓方式，發音長短等韻律），只有語言學的特性，所以必須透過自動預測的機制來產生這些文字的可能的聲學特性 (acoustic feature) 而所謂自動預測的機制，一般有 rule-based 跟 knowledge based 兩種方法，但是這兩種方法不但合成的聲音平淡又缺乏吸引力且遇到連續發音或要保留語者音色時表現都不好，因此近來串接合成法很受歡迎。

另外，目前在語音科技技術的應用上，語音合成技術通常被應用在以下多處：

- 1、個人電腦輔助：朗讀電子書、電子郵件、文章，或搭配手寫版、文書處理軟體、語音辨識軟體發音。
- 2、數位內容聲音快速生產：免去進錄音室煩惱，可以短時間將大量數位內容聲音化。
- 3、無障礙網頁發音：讓無障礙網頁立刻變成聲音。
- 4、手機、PDA 朗讀即時資訊：針對行動裝置提供語音資訊或語音互動操作介面。
- 5、電話查詢系統：可以朗讀個人帳戶資訊、股票評語、交通路況、氣象報告，其他諸如電話系統、CTI 系統、24

小時無人自動電話語音查詢/預約系統。

6、電話通報：提供電話外撥提醒重要訊息。

7、語音叫號：用來座診所叫名、法庭叫名等。

8、生產操作提示：生產線作業指導提示。

9、收銀盤點：結帳、盤點貨品時，可以根據條碼自動朗讀貨品名稱、數量等資訊。

10、家電、玩具：語音互動人性化介面。

11、導覽系統：電子播報系統、氣象查詢系統、電子地圖導覽系統、鐵路訂票系統等。

故本系統在這將語音合成技術運用在語音導覽上面，讓使用者在瀏覽系統時候，透過 Kiosk 平台操作，系統會透過發聲裝置將語音合成的聲音表現出來，之後有需要修改瀏覽內容時候，只需要改變 TTS 的文字內容即可，不需要再大費周章錄製聲音。另外，我們在 2.3 的 XML 文法設計裡面會提到 XML 設計原理，透過 XML 文法的設計來達到辨識的效果與提升辨識引擎的辨識率，並提供系統使用者在瀏覽博物館資訊時候，可以有一個親切且容易使用的介面來操作，讓使用者可以有一個不同於以往的操作介面。

2.3 XML 文法設計

因為語音辨識之語料庫過於龐大，因此我們透過撰寫 XML 語法可以提升語音辨識的辨識能力，如下圖 2.3，將導覽所需用到的字彙統整並加以處理，寫成 XML 可執行的語法，並利用 SAPI 語法去呼叫此 XML 文法，XML 被呼叫後會產生一個“.cfg”的暫存檔案，經過定義過的編碼，可以讓辨識率之增加。

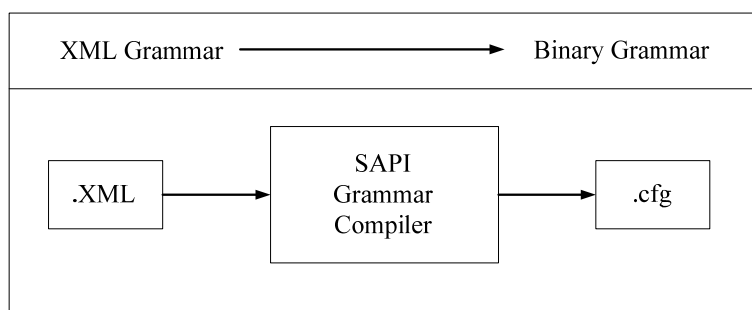


圖 2.3：XML 與 SAPI 之間的關係[21]

根據 Hirohiko Sagawa et al.在 Correction Grammars for Error Handling in a Speech Dialog System[17]一文有提到如何規劃一個完善的 VoiceXML Grammar，如下圖 2.4，當使用者說”I said”，之後可以說出”Tokyo”、”tomorrow”與”Tokyo tomorrow”等，也就是說衍生出來的文法有：

- 1、I said Tokyo。
- 2、I said tomorrow。
- 3、I said Tokyo tomorrow。

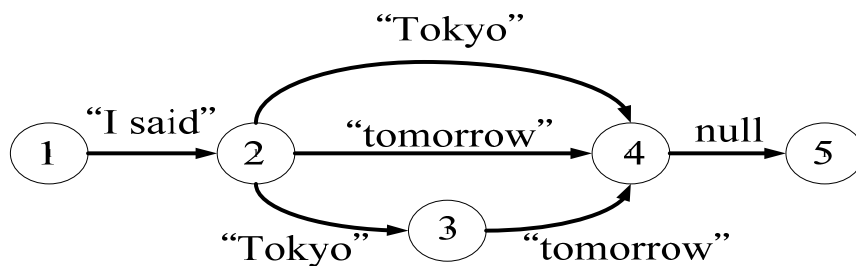


圖 2.4：Grammar 設計圖[17]

使用者只要說出上列三種之一的片語，VoiceXML 語言皆可判斷成想問東京明天的天氣，透過上述文法的組合，即可變化成多種組合以供系統呈現。

在本研究中我們採用的是 XML 的文法設計，XML 的設計裡面與 VocieXML 大同小異，彼此之間的差異只在定義上面不同，還有 XML 文法本身並沒有執行的能力，必須配合主程式去呼叫才可以作用，以下圖 2.5 XML 文法為例：


```

<GRAMMAR LANGID="404">

  <define>
    <id name="selection" val="1"/>
  </define>

  <RULE NAME="selection" TOPLEVEL="ACTIVE">
    <P PROPNAME="人" VAL="selection"><o>我</o></P>
      <ruleref name = "want"/>
    <ruleref name = "action"/>
      <o><ruleref name="製造王國"/></o>
    <o><ruleref name="立足台灣"/></o>
    <o><ruleref name="迎向挑戰"/></o>
    <o><ruleref name="帝國邊緣"/></o>
    <o><ruleref name="科技政策"/></o>
    <o><ruleref name="縱橫世界"/></o>

  </RULE>

  <rule name = "want">
    <L propname="name">
      <p val="1"><o>想</o></p>
      <p val="2"><o>要</o></p>
    </L>
  </rule>

```

```

    <p val="3"><o>希望</o></p>

    <p val="4"><o>要求</o></p>

    <p val="5"><o>想要</o></p>

    </L>
</rule>

<rule name = "action">
<L proptime="name1">
    <p val="1"><o>查詢</o></p>

    <p val="2"><o>找</o></p>

    <p val="3"><o>搜索</o></p>

    <p val="4"><o>搜尋</o></p>

    </L>
</rule>

```

圖 2.5：XML 文法設計圖

在我們所設計的語音使用者介面，我們也利用了 XML 程式的撰寫來增加辨識率，以上圖 2.5 為例，<GRAMMAR LANGID="404">表示支援中文編碼(409 表示支援英文)，<ruleref name = "want"/>表示跑到這一階段時候去呼叫<rule name = "want">這一段程式，而<rule name = "want">裡面的<p val="1">值表示可以搜尋到的關鍵字(Key

words)，簡單的說，當使用者說出“想、要、希望、要求、想要”，等字眼接可以被接受，另外，<rule name = "action">也是如此，說出“查詢、找、搜尋、搜索”等字眼接可以被接受，把兩個整合起來，所衍生出來的選擇有二十種之多。

2.4 語音使用者介面

根據 Dirk Schnelle et al.在 Audio Navigation Patterns [16] 文中指出，聲音(Speech)有著以下的限制：

- 1、聲音是短暫的，加上使用者的短暫記憶(STM, Short Term Memory)，當你聽到時候，它已經消失了。
- 2、聲音是一種看不見的東西，並沒有畫面的呈現。
- 3、聲音有著不勻稱的特性，說話比打字還快速，但是聆聽又比閱讀慢。

因此，為了解決聲音在先天上的缺失，我們在系統中加入一個使用者介面，畫面來源為行政院國家科學委員會[7]，我們嘗試在原本的 Kiosk 系統上加入了語音辨識的功能，除了聲音之外，還有一個可以供使用者觀看的畫面，改善聲音的限制與讓使用者可以依循畫面的指引完成所需要的服務，使用者在使用語音使用者介面導覽的時候可

以看到如下圖 2.6 的畫面。

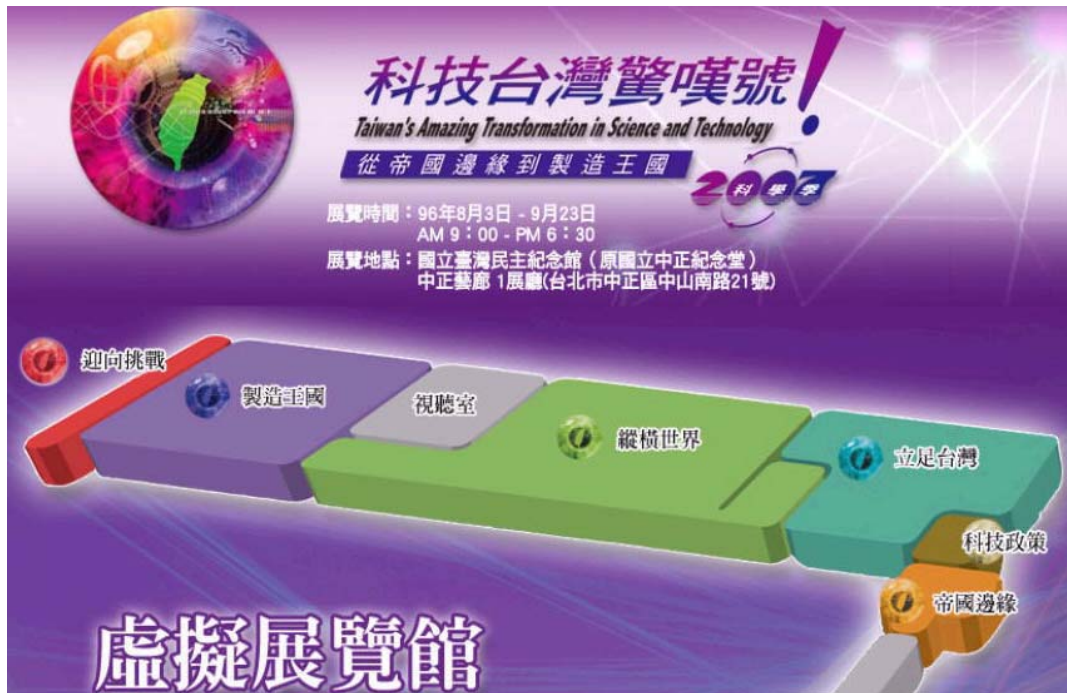


圖 2.6：系統介面導覽圖

Dirk Schnelle et al.又指出，一個語音使用者介面(Voice User Interface)需要有以下幾個條件，才可以稱作好的語音使用者介面[16]:

- 1、提供選單(Menu)，讓使用者可以透過畫面來改善短暫記憶的問題，但是僅僅有畫面的系統還是不夠的，提供選單可以讓使用者有個依據利用聲音控制選項。
- 2、提供求救功能(Help)，使用者在瀏覽系統期間，可能會迷失方向，所以提供功能讓使用者可以隨時隨地說出 HELP

指令，提醒使用者在這麼階段可以做什麼事，可以說出什麼指令來與語音使用者介面互動。如下圖 2.7，使用者在讀取 Info1 與 Info2 時候可以隨時使用 Help 指令。

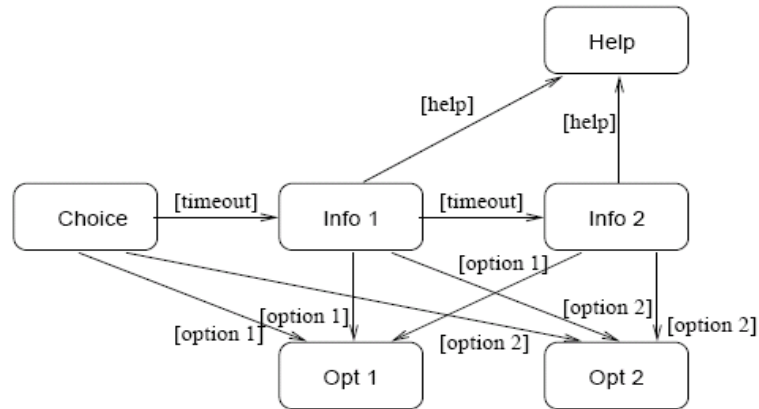


圖 2.7：使用者使用 Help 指令圖[16]

線性選項(Linear Choose) [16]，當使用者對於選取所需要的服務選項時，可以快速的取得需要的資訊，也可以透過下一步(Next)或是上一步(Previous)來控制，如下圖 2.8，使用者在 Choice 選取 Opt1、Opt2、Opt3 時候，可以直覺性的選取需要的服務。

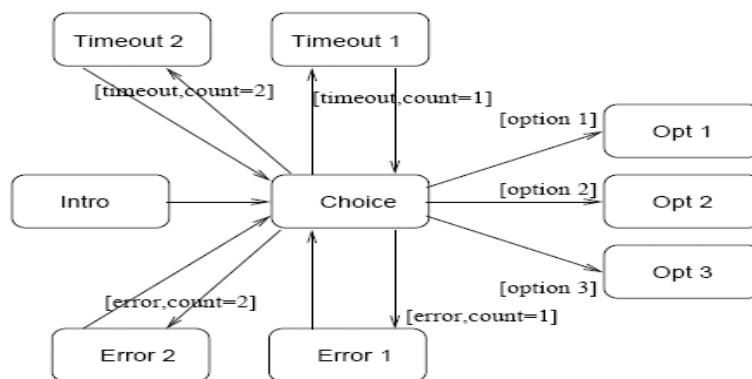


圖 2.8：使用者使用線性選項

還有，在設計一個使用者介面時候，也需要考慮到一些例外狀況，Dirk Schnelle et al.在文中也有提到，因為靠聲音來控制系統，所以使用者在使用系統途中可能會有迷失方向、語音辨識錯誤、或是使用到一半使用者中途離開等狀況發生，所以需要加入一些例外狀況來防止系統長時間等待或是出現錯誤狀況，分別為以下五點[16]：

- 1、當使用者沒有注意聆聽系統說明或是不知道該發出什麼聲音時候，等待時間(timeout)的指令這時候就可以發揮效用，經過一段時間系統沒有收到任何指令或是聲音時候，系統會參照等待時間的控制程序進行反覆說明(repeat)或是執行下一個動作功能。
- 2、為了防止系統出現錯誤情況(error)發生，有經驗的介面設計者通常會強迫使用者聽完冗長的說明，當使用者聽完說明之後會比沒有聽過說明的使用者更能熟悉系統的操作過程。
- 3、使用者在使用系統期間，可能會發生不知道該如何使用或是不知道該說些什麼關鍵字來控制系統，這時候系統提示詳細的說明(more detail explanations)就可以幫助使用者在這個階段完成取得需要的服務。

4、不同的錯誤情況與不同的等待時間有著不同的重要性。

5、過低的辨識成效(Low recognition performance)與背景噪音(Background noise)容易造成錯誤的發生。

所以我們設計這個使用者介面的時候，有將以上五點建議考慮進去，在我們系統使用當中，系統會以跑馬燈的形式提醒使用者可以在這個階段做些什麼事情，以防止使用者會有不知所措的情況發生；我們也將系統設計成強迫使用者聽完使用說明或是提示之後才可以坐下一步的動作，以加深使用者對系統的認知；另外，為了防止系統辨識率過低的問題，我們著手設計了 XML 文法，藉由 XML 文法與 SAPI 的搭配組合，可以讓辨識率大大的提升。

另外，Daniela Oria et al.在 Automatic Generation of Speech Interfaces for Web-based Application[15]一篇中也有提到，在語音使用者介面下需要有以下步驟，可以幫助使用者瀏覽系統：

- 1、幫助(Help)。
- 2、列表(List commands/options)。
- 3、主選單(Main menu)。
- 4、退回上一頁(Go back)。
- 5、重覆(Repeat)。

6、待命(Standby)。

7、運算(Operator)。

所以在設計語音導覽系統時候，我們將「幫助」、「主選單」、「退回上一頁」等功能加入，在系統中「幫助」的功能我們將提示以跑馬燈的形式存在，當使用者瀏覽系統時候，如果遇到不知道該如何瀏覽時，可以透過跑馬燈的提示完成導覽；而「主選單」的功能是提供使用者在系統剛開始接觸到系統時候來選擇需要的服務，透過主選單的功能可以選擇觸控式或是語音導覽服務；而「退回上一頁」是使用者在子服務選項的時候，可以透過退回上一頁回到父服務選項，將以上條件考慮進去，進而完成現在的語音 Kiosk 導覽系統，系統剛完成時候一定會有欠缺考慮的細項，所以完成系統會請研究室同學幫忙進行第一次的測試，透過旁觀者的角度來改善系統的流暢度與親切度。

2.5 互動式多媒體資訊站

Kiosk 的原義是指售票亭，在本研究內可以解釋為資訊服務站，它通常是由工業級的電腦組成，加上可以隨使用環境不同而變化的軟體程式，將可以提供在不同環境下達到多元的服務，如：在機場利用 Kiosk 來訂機票與劃位、量飯店的商品查詢服務、醫療院所的掛號服

務、旅遊地區的導覽服務等，皆可以利用 Kiosk 來取代人力，達到降低人力成本與資訊化服務的概念，Frost & Sullivan Research 針對 Kiosk 市場長期觀察，針對其產業內容與應用領域主要分為以下五類，本實驗將 Kiosk 加入語音辨識功能並運用在博物館導覽系統。

- 1、零售業。
- 2、政府機構。
- 3、金融產業。
- 4、旅遊運輸。
- 5、娛樂產業與通訊互動 Kiosk。

相對於傳統的販賣機，Kiosk 可以依照內部所寫程式來執行，可提供的服務相對較多，可以應內部程式的修改而變化，達到多元化的服務，本篇研究的主題放在原有的 Kiosk 系統與加入語音辨識後的 Kiosk 系統比較，使用者對於兩種系統的使用感覺是如何的？對於語音輸入的使用者介面是否會感覺到好用？透過相互比較的實驗，進而改善系統達到使用者可以接受新的語音輸入介面。

由於 Kiosk 的應用方面很廣，這本篇研究裡面我們將 Kiosk 應用在博物館資訊的導覽，當使用者進入博物館時候，可以透過我們的

Kiosk 系統做預先取得基本導覽資訊的服務，例如，使用者想知道半導體資訊的展區，在 Kiosk 導覽裡面我們提供使用者樓層資訊，或是，使用者可以在 Kiosk 平台上預先取得接下來要參觀的展區的資訊，透過 Kiosk 可以知道展區特展的內容，因為 Kiosk 建置簡單、可以節省許多人事成本，加上可以因應使用環境來做細部調整，所以我們嘗試利用 SWOT 來做加入語音使用者介面後的 Kiosk 導覽之分析。

2.5.1 Kiosk 的 SWOT 分析

一、優勢(Strength)：

- 1、提供使用者預先取得導覽資訊的服務，使用者透過 Kiosk 平台取得預導覽資料的簡介。
- 2、透過 Kiosk 導覽減低博物館人事方面的開銷，Kiosk 可以取代導覽人員的講解，讓使用者透過系統就可以了解資訊。
- 3、利用語音控制 Kiosk 比較有趣，讓使用者可以有一種新的使用者介面的選擇，透過語音來控制導覽過程是一種全新的體驗。

二、劣勢(Weakness)：

- 1、利用 Kiosk 的導覽方式，必須在流程控制上面有很好的效

能才可以讓使用者有想用這系統的意願。

- 2、語音辨識的效能，加入語音辨識的 Kiosk，在博物館等公開且噪音影響場合，其辨識能力是否可以適用於每一位使用者。
- 3、Kiosk 系統提供的資訊有限，因為要提供使用者易用的環境，故在系統設計理念中，我們並不會將所有資訊都輸入進去 Kiosk 平台。

三、機會(Opportunity)：

- 1、國內市場上尚未見到如此類型的 Kiosk 導覽系統，如果率先導入語音使用者介面的操作模式的話且競爭者繼續使用傳統觸控式，可以提前讓使用者體驗到新的操作介面，相對於使用觸控式 Kiosk 機台，語音辨識輸入的 Kiosk 機台可以吸引使用者注意。
- 2、語音 Kiosk 介面可以隨時增加服務項目，如將購票等服務也一併加入服務當中，也可以減少節日時候售票處人潮壅塞問題，提高博物館的使用效率。

四、威脅(Threaten)：

- 1、架設 Kiosk 機器需要大筆的資金投入，才能在市場上面行

成通路，讓使用者可以習慣 Kiosk 的操作模式，以吸引商家導入 Kiosk 導覽系統。

- 2、由於資料需要逐漸更新，所以在資料的傳輸與更新上，對於駭客的防治機制變的很重要。

2.5.2 Kiosk 介面需要的功能

而 L. Lamel et al.在 User evaluation of the MASK Kiosk[20]一文中提出，一個直覺反映的 Kiosk 使用者介面，需要以下功能：

- 1、容易在任務(task)間作切換。
- 2、圖片的切換，讓使用者知道系統目前狀況。
- 3、兩階段的協助(help)指令，當系統待命超過一段時間就發出指令，確保系統不會因為待機而當機。

所以我們在設計語音 Kiosk 系統時候，也加入 Help 指令，不過在本研究中，系統的 Help 指令是引導使用者在這個階段可以做些什麼事情，例如：回到上一頁、往下一頁...等資訊，當系統確定收到使用者發出的指令之後，會將瀏覽資訊以網頁型式呈現出來，讓使用者可以知道系統確實收到指定並作出動作，另外，系統也有防止當機的機制，當 Kiosk 系統都無人使用時候，系統經過一段等待時間，會透過

程式控制來結束瀏覽畫面，此時，系統會跳到一開始的待機畫面等待下一個使用者來臨。

本研究的語音使用者介面根據 2.4 與 2.5 所收集的資料來設計，系統設計的本意在於讓使用者可以簡單的透過本研究系統快速且容易的取得相關資訊，故本研究著重於使用者介面的設計與結合 Kiosk 系統的應用，在國外相關文獻中，有類似的設計團隊在探討如何將語音辨識功能與 Kiosk 系統做結合[20]，不過國內目前並沒有將語音辨識與 Kiosk 應用在博物館導覽的服務，透過我們所設計的系統讓使用者可以瀏覽博物館資訊，並探討如何可以透過系統服務取得相關資訊，我們將在下面 2.6 線性與非線性任務中會說明任務的進行方式。

2.6 線性與非線性任務

所謂線性與非線性任務，在這邊我們以瀏覽系統的程序來為定義，以下面系統的流程圖來做解釋。

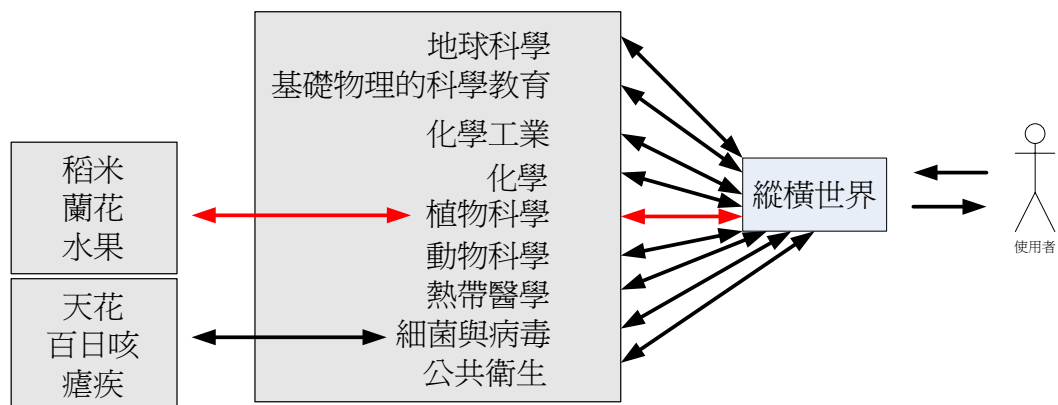


圖 2.9：線性任務下的選擇模式

以上圖 2.9 為例，線性任務是指，當使用者從主選單進入縱橫世界的時候，畫面顯示的選項有『地球科學』、『基礎物理的科學教育』、『化學工業』、『化學』、『植物科學』、『動物科學』、『熱帶醫學』、『細菌與病毒』與『公共衛生』...等服務，之後使用者可以再由『細菌與病毒』進去選擇『天花』、『百日咳』與『瘧疾』的選項，也就是說，使用者能依循畫面的指示來一步一步選取服務內容，可是當服務內容的層級越多，或是使用者一次需要選取多個服務內容時候，線性的任務便不能滿足使用者的需求，所以才會有非線性任務的衍生。

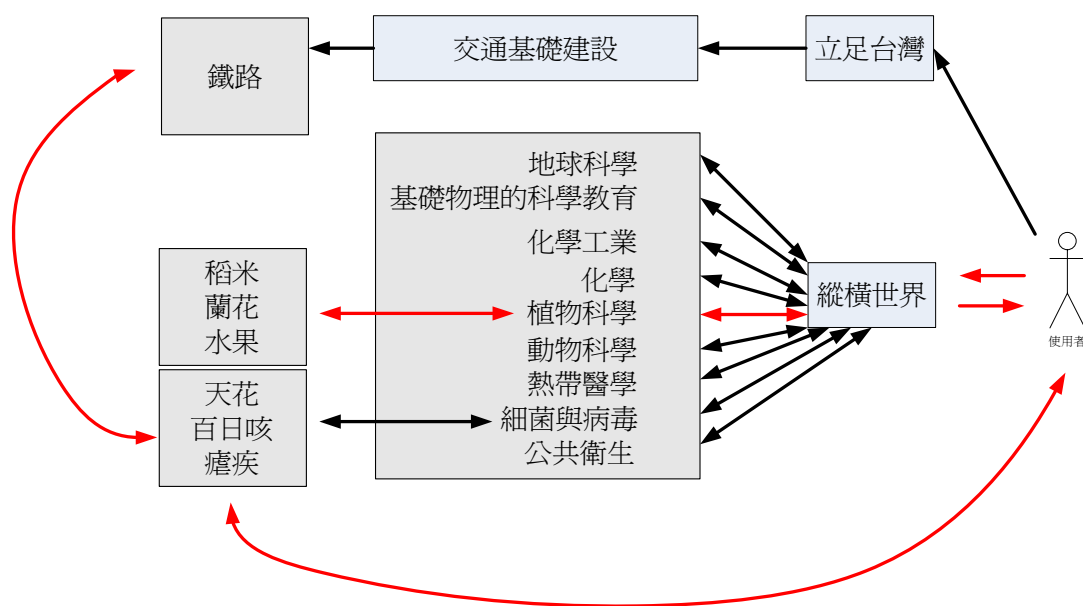


圖 2.10：非線性任務下的選擇模式

透過非線性任務的規劃，使用者可以隨時隨地切換需要的服務內容，使用者如果一開始就知道需要的服務內容選項，即可透過非線性

任務的服務來直接選取目的地服務內容，如上圖 2.10，使用者想直接知道『瘧疾』相關的資訊，直接在系統開始時候說：「幫我找一下瘧疾相關的內容」，系統就會導覽他們到所需要的服務，而看完導覽內容之後如果還需要繼續閱讀接下來的資訊，可以直接跳到需要的地方，如系統導覽完瘧疾內容之後，使用者需要觀看交通基礎建設裡面的『鐵路』相關資訊，直接可以說：「幫我找鐵路資料」，系統會引導使用者到達鐵路相關資訊，不需要像線性任務規劃一樣，必須跳回到主選單選項之後再分別從立足台灣到交通基礎建設，之後才能找到鐵路相關的訊息，所以在本研究當中，我們在語音控制的導覽系統之中加入非線性任務的規劃，目的是要讓使用者可以利用語音輸入來取得資訊，且更方便且更省時的取得需要的服務資訊。

透過語音輸入的 Kiosk 平台，我們提供使用者一個新的使用者介面的服務，與觸控式導覽系統不同的是，語音使用者介面的導覽系統可以提供線性與非線性的導覽服務，在語音系統中，使用非線性來選取服務的效能會比線性的選取更佳，我們在之後的實驗驗證裡面會去驗證這一點。

2.7 使用者滿意度

使用者滿意度的測量發展最早在 1970 年代已經開始了，而 Pearson 在 1977 年針對使用者滿意度提出定義，當時他對使用者滿意度的定義為「對於組織提供以電腦為基礎的資訊與服務所產生的情感與態度的綜合體」，另一學者 Kim 在 1989 年提出，將使用者滿意度的測量分為三大類，第一類是態度所獲得的使用者滿意度，這類主要是由態度量表來衡量使用者對於資訊系統的滿意程度；第二類為資訊品質所獲得的滿意度，這一類主要是以資訊系統的資訊品質的資訊輸出來衡量使用者對資訊系統的滿意度；第三類是由有用性來獲得的滿意度，另外根據管理資訊系統(MIS)的定義，使用者滿意度應該是使用者對於系統正面與負面態度的總合，所以當系統完成之後，使用者對於系統的功能、取得資訊的服務...等，都會讓使用者對系統心理層面的感覺。我們參照了黃瓊慧女士在「台灣企業資料倉儲滿意度評估」[9]與 Pearson 與 Bailey 在 1983 年提出 39 項問卷設計的指標以用來衡量使用者對於電腦的使用滿意度，如下表 2.2

表 2.2：Pearson 與 Bailey 提出的 39 個問卷問項

管理階層的 參與程度	EDP 組資資 源的競合	資源優先配 置的決定	服務性支付 的收回方式
與 EDP 人 員的關係	與 EDP 人 員的溝通	EDP 部門 人員的技 術能力	EDP 人員 的態度
產品與服務 的安排	新系統發 展時間上 的要求	現有系統要求改 變的處理流程	供應商 的支援
系統回 應時間	EDP 生產 力的工具	容易使用	正確性
適時性	精準性	可靠度	幣別
完整性	輸出格式	電腦使用語言	輸出的數量
相關性	錯誤的修復	資料的安全	系統文件
電腦支援的期望	系統了解	有用性	對系統的信任性
使用者參與程度	內控能力	使用者的 教育訓練	工作的影響
Function EDP 組織部門定位	系統彈性	系統整合	

我們根據上表整理的問項再結合我們的設計的系統環境，分別以系統回應時間、容易使用、正確性與系統彈性等四個問項為基礎製作問卷。

- 1、「系統回應時間」：系統回應使用者需求所需要執行的時間。
- 2、「容易使用」：使用者對於介面的使用感覺。
- 3、「正確性」：利用此介面是否可以正確到達使用者需要的服務項目。
- 4、「系統彈性」：系統介面是否符合一般使用者可以接受的流程。

問卷並採用李克特五點(非常同意、同意、普通、不同意、非常不同意)量表分析(Likert-type scale)來制定各個問題的配分，如：非常同意給予五分、非常不同意給予一分，並在取得問卷資料後進行分析，蒐集資料後根據以上四點的分析評估我們可以知道使用者對於系統使用的觀感，這可以讓我們知道使用者使用系統的感覺並加以改善系統缺點。

第三章、研究方法

本研究主要是設計一個新的語音使用者介面來瀏覽博物館資訊，透過此語音辨識系統，使用者可以簡單的說出需要的服務項目、關鍵字，即可將資訊等服務呈現在使用者面前，而本章分為四小節，分別為系統設計與組成架構、可延伸標記語言的設計、開發工具的介紹與研究設計的過程。

3.1 系統設計與組成架構

本研究著重於使用者介面的設計，在原有的觸控式 Kiosk 導覽系統上再增加一個語音使用者介面，如下圖 3.1，使用者透過 Kiosk 裝置說出需要的服務，系統會將使用者說出的語句傳送到語音辨識引擎 (Speech Engine)，之後透過自動語音辨識來 (ASR, Automatic Speech Recognition) 處理，ASR 會尋找 XML 的語料庫文法 (XML Grammar) 比對語句，從語料庫裡面抓到對的關鍵字 (Keyword) 傳回系統，作處理，進而將使用者需要的資料以網頁方式傳回到 Kiosk 平台來呈現。

加入新的語音使用者介面可以讓使用者利用自然語音輸入取代傳統觸控，使用者也可以透過螢幕的畫面呈現、網頁的提示、語音提示...等，可以知道自己在這個階段可以做些什麼，在系統中，使用者

透過 Kiosk 平台觀看需要的資訊服務，網頁下面會有跑馬燈(紅線標示)提示使用者可以說出什麼關鍵字來驅動下一個服務，透過語音導覽介面，使用者可以依照網頁下面的跑馬燈提示，來完成需要的使用者的資訊取得。

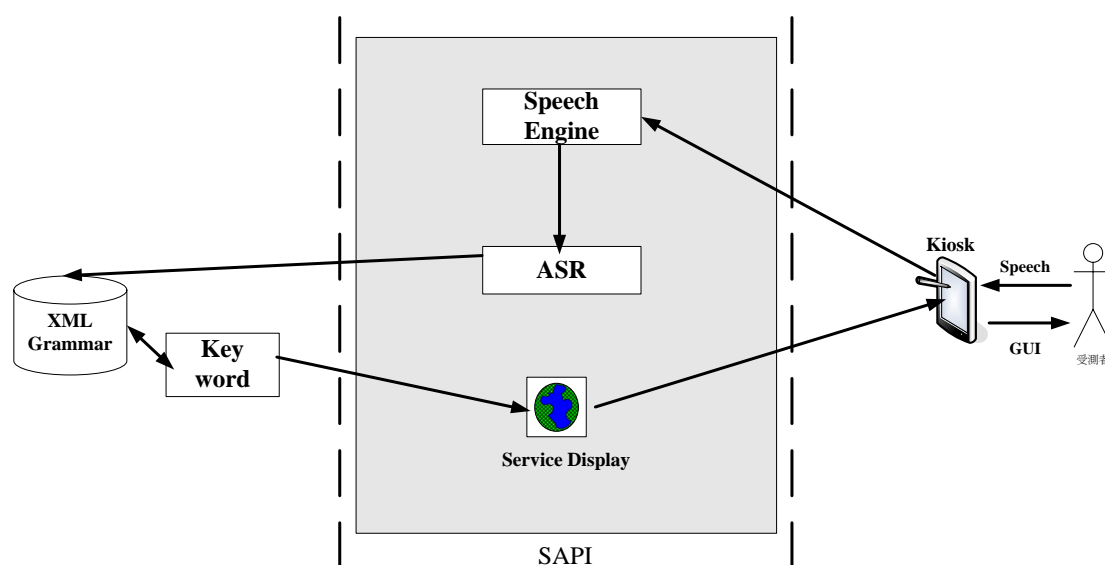


圖 3.1：使用者資料流向圖

如下圖 3.2，跑馬燈出現「在這邊你由左上角的選項說出你要的服務，如：半導體..等」，表示使用者可以透過以上提示說出需要的服務，避免使用者在瀏覽系統時不知道該說出什麼關鍵字而發生卡住的狀況。

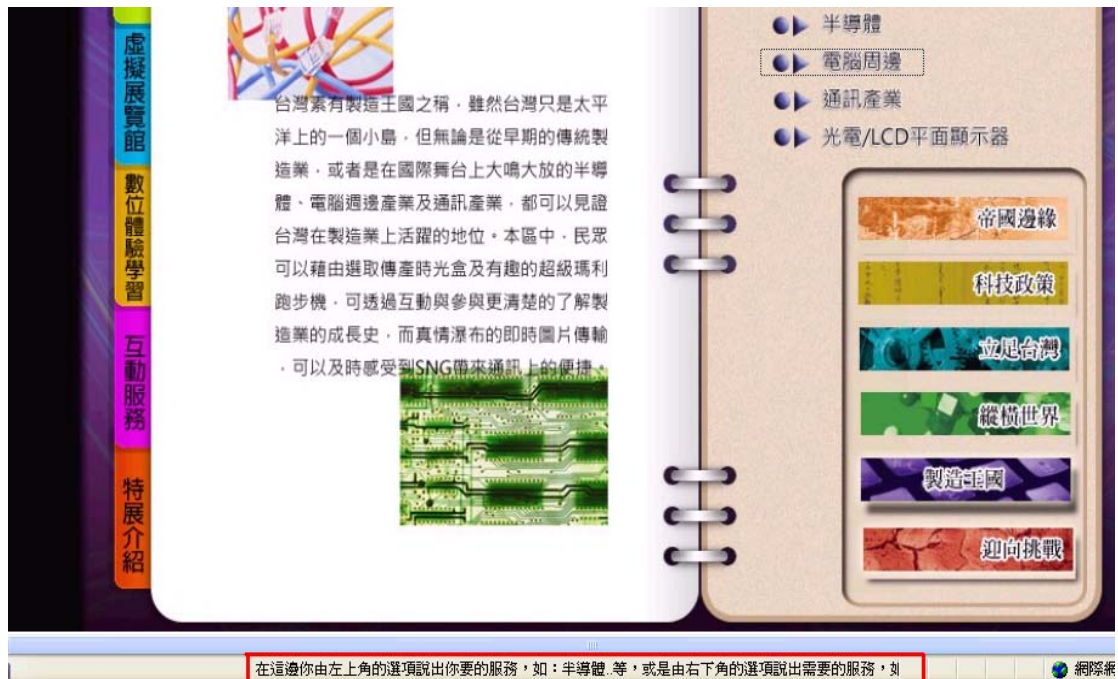


圖 3.2：網頁使用者提示圖[21]

另外，我們在系統中間也增加了等待時間，如下圖 3.3，當等待時間大於 18 秒時候，系統會去呼叫 Timer1_Timer()，系統會用語音合成出「對不起，我沒收到您的訊息，請再說一次」，當執行三次之後，系統會跳到一開始待機畫面，避免系統因為等待下一個使用者這段時間而一直執行。設計這個指令是為了避免系統在等待過程中一直接收到聲音並重複進行辨識工作，造成系統負荷過重或是造成當機，所以設計等待時間並跳出系統服務是需要的。

```
Try
    t.Interval >= 18000 And strText = ""

Catch ex As Exception

    Call Timer1_Timer()
    b = +1
    t.Stop()

End Try

Private Sub Timer1_Timer()
my_voice.Speak("對不起，我沒有收到你的信息，請在說一次")
End Sub
```

圖 3.3：系統等待時間圖

當使用者使用 Kiosk 系統導覽時候，其目的是為了取得資訊內容的服務，所以下圖 3.4 表示使用者使用系統時候可以取得到的資訊內容，例如使用者想知道半導體的內容，如：『晶圓製造』等資訊內容的話，利用傳統觸控式使用者界面的話，需要從『製造王國』選項到『半導體』才可以取得相關資訊，但是使用語音使用者界面的話，直接在開始畫面說出需要的服務內容，即可以取得資訊，這就是我們設

計這麼系統的主要目的，利用非線性的任務規劃來幫助使用者快速取得需要的服務內容。

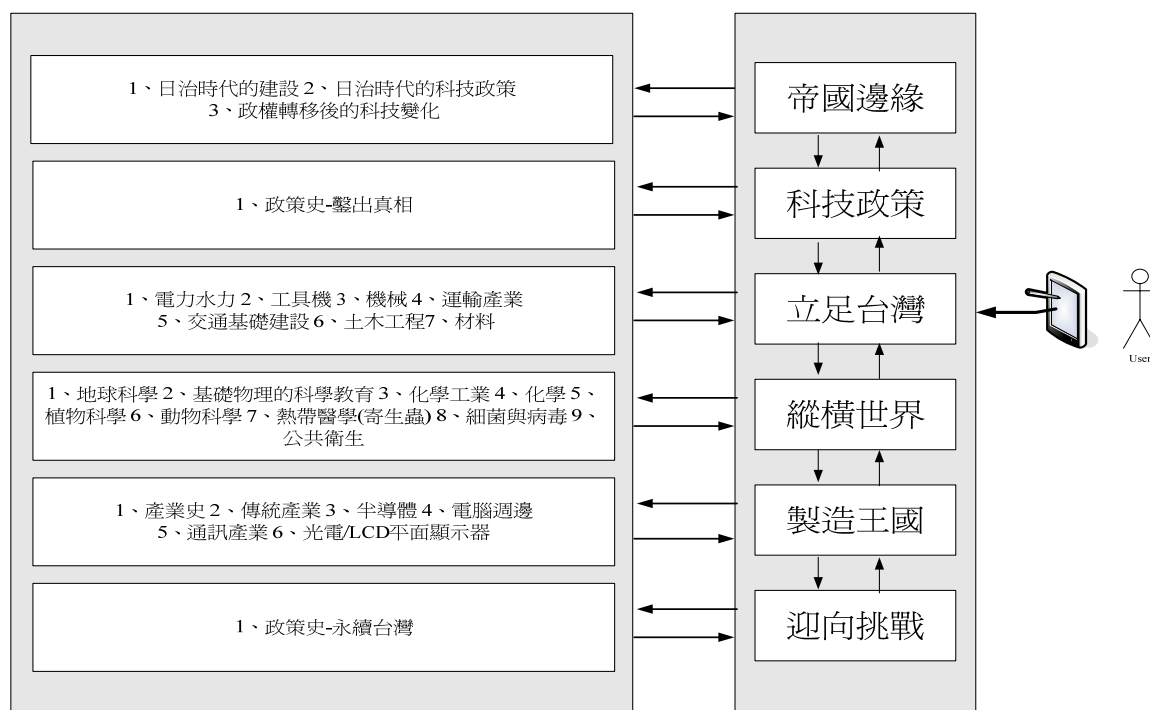


圖 3.4：博物館資訊簡圖

另外，我們在使用者進去瀏覽資訊內容時候，我們也設計了一個按鈕，讓使用者自行決定要不要”聽”系統來解說選取的服務內容，如下圖 3.5，在使用者瀏覽服務內容時候，可以由左上角的「按我說話」來啟動語音導覽的服務，系統會提供使用者語音的導覽服務，使用者只要透過聲音就可以瞭解且取得需要的服務內容；「停止說話」按鈕可以幫助使用者，當系統在敘訴服務的內涵的時候，使用者如果不想聆聽，可以點取此按鈕來取消語音服務。

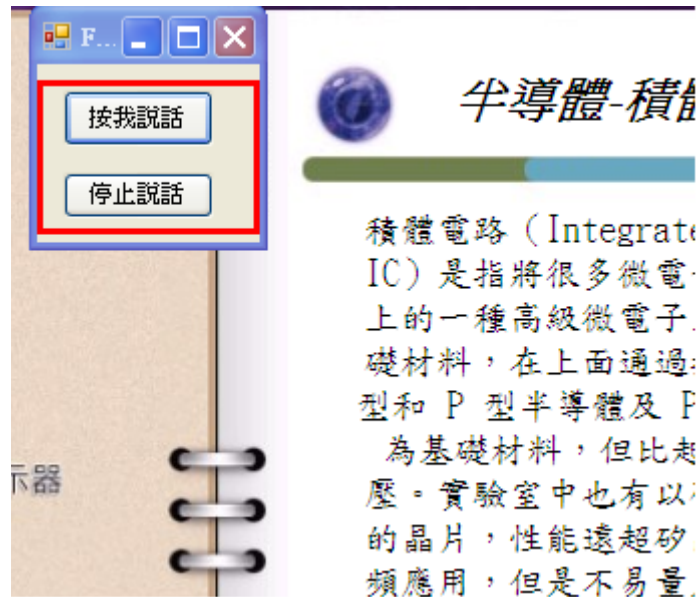


圖 3.5：系統發聲控制圖

3.2 可延伸標記語言 (XML Grammar) 設計

在 XML 的文法設計中，辨識過程其實就是求出音框對狀態關係的過程，最後求出一個最大機率的路徑[1]，如下圖 3.6，如果只是單純的辨識關鍵詞的話，那麼只要遵循狀態與音框關係的時間規則，就能求出一條具有最大機率值的路徑。

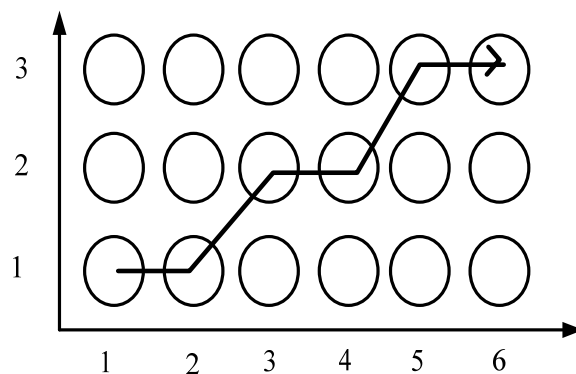


圖 3.6：語音路徑選擇圖一

在單純關鍵詞的辨識過程中，狀態轉移的過程也比較單純，如下圖 3.7，從”I would like to drive from Seattle to New York”文法中可以看出眾多選擇，例如：Method 的子選項中可以有 Fly 與 Drive 兩種選擇，Direction 則有 From 與 To 選項，City 的子選項有 Seattle、New York、Los Angeles 與 Albuquerque 四種選擇，依照機率值來選擇最佳路徑，所以系統辨識出 I would like to drive from Seattle to New York 的選項機率值為 $\frac{1}{2 \times 2 \times 4 \times 2 \times 4} = \frac{1}{128}$ 。而 XML 文法設計範圍越大，則有越多的選項以供選擇。

I would like to drive from Seattle to New York

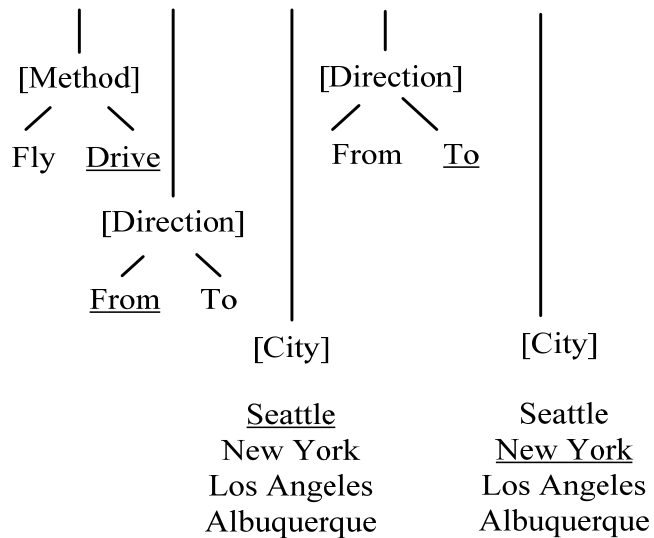


圖 3.7：語音路徑選擇圖二[21]

3.3 開發工具

在系統的開發過程中，我們運用語音辨識技術來開發使用者介面，並利用觸控式面板的電腦搭配來使用，已完成語音使用者介面的開發，硬體與軟體部分如下：

3.3.1 硬體：

- 1、電腦主機：Asus A3H 筆記型電腦一台
- 2、測試平台：HP 平板電腦(模擬 Kiosk 系統)
- 3、高感應頭戴式麥克風

3.3.2 軟體：

- 1、Microsoft Speech SDK 6.1
- 2、Microsoft Visual Studio 2005
- 3、XML Grammar Compliance Tool
- 4、SQL Server 2005 Express
- 5、Dream wave 2004

3.4 研究設計

本實驗總共分為兩個階段來測試，第一階段是語音瀏覽程式剛完成時候，讓使用者使用語音 Kiosk 系統來做測試，透過使用者實驗過程中，設計者利用放聲思考法來針對系統本身流程與優缺點等問題提

問，讓使用者說出自己對系統的看法，設計者再透過收集使用者想法的資料來當做參考，進而更改語音程式內容或是流程，以符合貼近容易使用與親切的使用者介面系統。

另外，第二階段的研究是正式測試，經過第一階段測試的改進之後，系統比一開始的瀏覽介面更能貼近使用者的需求，之後進行第二階段的測試，在第二階段我們比較語音使用者介面系統與觸控式系統的時間效能，分別針對線性任務與非線性任務底下完成一次瀏覽所需要的時間做相關比較，分別取得時間數值並做統計 Z 檢定分析，以便研究語音使用者介面在線性任務底下或是非線性任務底下的效率較高。

3.4.1 第一階段研究

第一階段研究採用放聲思考法來測試，實驗一開始時候，邀請五位研究生來幫忙測試系統，透過研究生親自的流程實驗，並利用放聲思考法來測試，在測試系統的同時並講出系統的優缺點，以下是針對第一階段實驗者的說明與對話紀錄，另外圖 3.8 表示第一階段實驗之流程規劃。

一、人員：接受放聲思考法的受測者共五名，主測者一名。

二、實驗說明：

- i. 為了避免受測者的學習經驗影響實驗結果，所以我們篩選受測者皆為未使用過語音辨識系統的研究生五位。
- ii. 受測者只能透過聲音來控制 Kiosk 系統，不能用手觸控來當作操作媒介，目的在於讓使用者針對語音輸入的方式提出他們的看法。
- iii. 以放聲思考法(Thinking aloud)進行第一階段實驗，除了一般思考過程之外，特別要求受測者在進行瀏覽資訊時，說出如何知道自己現在所在位置，以及如何前往下一個頁面資訊等問題，主測者不會去干預或提示受測者的測試過程與問題。

三、方法與程序：

- i. 步驟一：如下圖 3.9，使用者設定訓練檔，透過 Speech SDK 設定的語音訓練，經過訓練後可以達到較高的辨識準確率，而本次測試為求公平性，故測試之前讓每位使用者錄製訓練檔，讓每位使用

者皆有自己的辨識依據的設定檔，透過一次的訓練並儲存設定檔以供辨識引擎作辨識依據，當使用者在使用系統時候，我們選取使用者自己的辨識設定檔以供系統辨識時候作為依據。

第一次實驗-使用放聲思考法

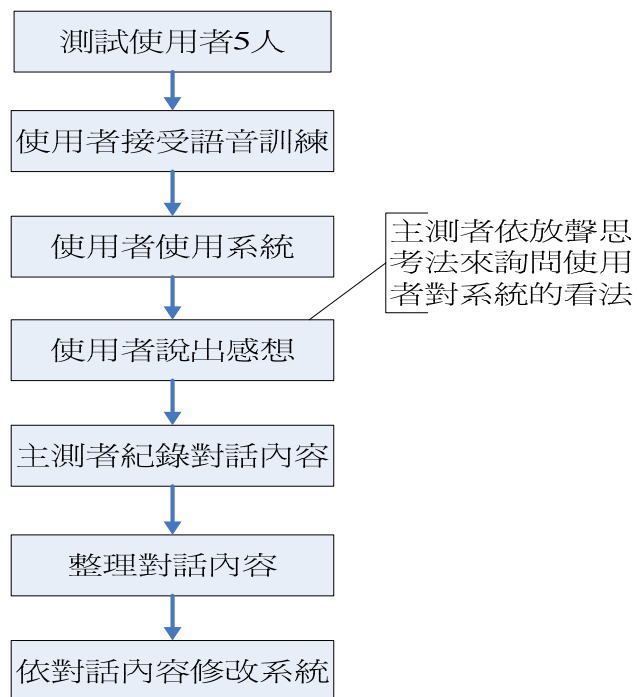


圖 3.8：第一階段實驗流程圖

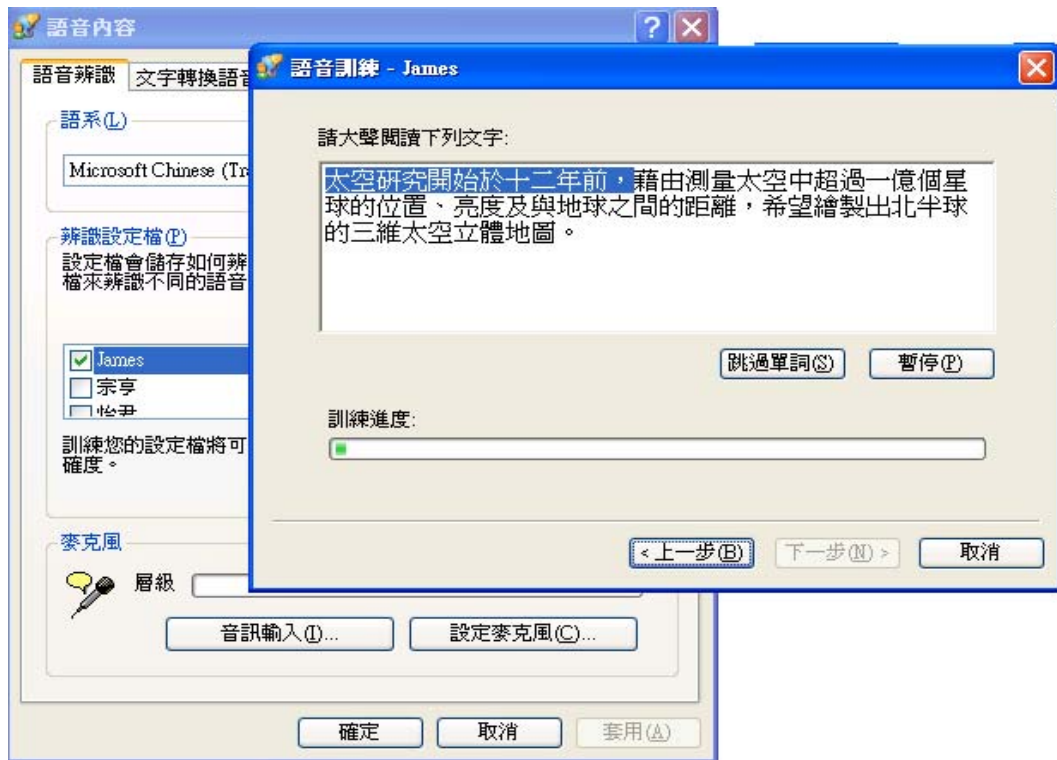


圖 3.9：使用者語音訓練圖

- ii. 步驟二：通過訓練之後，使用者接著測試系統，透過使用者的測試來了解系統的優缺點。使用者從一開始的選單畫面選擇導覽方式(選擇觸控式導覽或是語音導覽)，接著從子選項到子選項的瀏覽，受測者不需要主測者的協助，透過畫面與語音的指導瀏覽系統服務，經過測試之後使用者告知主測者在瀏覽期間是否有遇到困難或是提示的地方是否清楚等相關問題，主測者會利用錄音方式將對話過程紀錄下來，並將結果整理成表 3.1。

- iii. 步驟三：透過放聲思考法提出問題並將對話過程紀錄下來，如下表 3.1，並參考使用者意見紀錄表的意見來改善系統。



表 3.1：使用者放聲思考法紀錄表

使用者一對話紀錄：

使用者一：我對這系統的感受，一開始這系統對我說還蠻新鮮的，那我覺得便利性還蠻多的，而且阿，語音辨識的能力也蠻高的，像之前我用比較高或是較興奮的語調，它還是辨識的出來，那有一點讓我蠻驚訝的是，我連用有點台語的口音來講，它也是可以辨識的。

主測者：那在瀏覽的過程中，你覺得瀏覽的順暢度如何？

使用者一：關於那瀏覽的順暢度方面，我覺得盡量看能不能不要等到”咚”一聲之後，才可以輸入指令，也就是說，當機械語音在講的時候，這時候輸入指令還是可以切換服務的功能，那流暢度的話基本上是可以的。

主測者：那你覺得這個系統除了你所說的插話功能外，還有什麼需要改進的嗎？

使用者一：改進方面的話，其實我也是第一次接觸到這個語音導覽系統，那我是覺得，這樣的機能已經很足夠的，謝謝。

使用者二對話紀錄：

主測者：在測試的過程中，你覺得這系統在瀏覽過程中順暢度如

何？

使用者二：順暢度歐，還可以阿。

主測者：那你在使用完這系統之後，你覺得這系統有什麼需要改進的地方呢？

使用者二：使用完之後歐，覺得系統的辨識率與回應速度的問題。

主測者：那你是覺得系統的回應速度太慢與辨識率不高是最大的兩個問題？

使用者二：對阿，就是你回應速度要快的話，辨識能力也要加強。

使用者三對話紀錄：

使用者三：對這系統的看法，我覺得蠻新鮮的，因為之前沒用過類似的系統利用說話的方式來操作且導覽，蠻好玩的

主測者：那你覺得這個系統在瀏覽時候順暢度如何呢？

使用者三：順暢度歐，還不錯啦，瀏覽時候畫面蠻清楚的，每一個步驟前後關係簡單，還不錯啦。

主測者：那你在使用完這系統之後，你覺得這系統有什麼需要改進的地方呢？

使用者三：我覺得系統在開啟服務時候，有時候會停頓一下，說話的時候，有時候說第一次時候沒反應，在講一次時，會連續跑兩次的服務，應該是停頓的問題啦。

主測者：那就是說，系統最大的問題在於停頓嗎？

使用者三：對的，改善後應該可以更好。

使用者四對話紀錄：

主測者：在測試的過程中，你覺得這系統在瀏覽過程中順暢度如何？

使用者四：還不錯，系統流暢度可能還需要再修改一下。

主測者：能否提一些可以改進的意見以供參考？

使用者四：網頁服務開啟的時候好像有點慢。

主測者：就是服務開啟時候可以再快一點？

使用者四：恩，對，還有有時候選項好像會讀不太出來。

主測者：也就是說，辨識率的問題？

使用者四：對，大概就是這兩個問題。

使用者五對話紀錄：

使用者五：在試用完這個系統之後，我覺得這個方法還蠻不錯的
因為之前去便利商店使用過類似的系統，不過它是用
按的，不能像現在這個系統這樣，透過聲音來到達需
要的地方，對於我來說是一個不錯的選擇。

主測者：在測試的過程中，你覺得這系統在瀏覽過程中的順暢度如何？

使用者五：順暢度歐，恩，還不錯，不過是電腦的關係吧，系統開啟網頁的時候感覺有點慢，還有進去到像半導體的時候，如果有聲音或是一個按鈕讓可以讓使用者自行決定，要不要更進一步的取得導覽的詳細內容，畢竟一開始如果給使用者太多他不要的資料，他可能會不理會系統吧。

主測者：所以你希望系統可以依照使用者的本身需求來取得更深一層的服務內容資訊？

使用者五：對的，另外，看能不能結合更多的聲光效果，這樣應該可以更吸引人。

iv. 步驟四：將表 3.1 的對話內容簡化，並將對話內容整理如下表 3.2，之後在參照整理好的意見表來修改系統，以達符合使用者需求之系統。

表 3.2：使用者意見整理表

使用者一：希望加入可以隨時插話的機制。
使用者二：希望回應速度與辨識能力再加強。
使用者三：希望服務開啟時候可以更快。
使用者四：希望服務開啟可以快一點、辨識率有點不高。
使用者五：希望能有更多的服務內容，且可以自行選擇要不要取得進一步資訊。

- v. 步驟五：修改程式內容並參照使用者意見更改成更能符合使用者需求之系統。

3.4.2 第二階段研究

第二階段我們採取使用者正式測試，透過二十位使用者實際的任務測試來取得實驗數據，一開始二十位使用者經過隨機分配到 1、觸控式導覽 2、語音控制導覽兩組，並進行「線性任務實驗」與「非線性任務實驗」，透過兩組的實驗互相比較其時間效能，也就是說，當完成一個任務時候所需要的時間之比較，透過實驗取得時間值，之後利用統計方法 Z 檢定來分析兩者之間的效能比較，並利用問卷方式讓使用者填寫，回收後作相關分析，其效能比較之 Z 檢定分析會在第四章做詳細介紹。

第四章、系統實作與結果分析

在本章節裡面，我們將會呈現系統的細部說明與分析後的結果，並與之前所做的三個假設做相互對應，映證我們的假設是正確的，使用者在語音使用者介面的使用狀況確是比觸控式使用者介面還要好。

4.1 系統實作

在系統實作方面，我們以行政院國家科學委員會的科技台灣驚嘆號-2007 科學季導覽網頁為例[7]，我們嘗試將導覽內容加入 Kiosk 系統裡面，並增加了語音辨識的功能，並以博物館導覽的資訊為服務內容，下圖 4.1 為使用者可瀏覽的博物館導覽資訊圖，當使用者使用 Kiosk 導覽系統實後，只需要說出想進入的項目，例如使用者想知道『製造王國』資訊或者是利用系統設計的非線性任務的特性直接取得『鍵盤』資訊，此時我們所設計的語音使用者介面就會透過 SAPI 進行語句的分析並判斷使用者所說出的語句，之後從 Web Server 取得相對應的網頁資訊，如下圖 4.2。我們將科技台灣這資訊導入語音使用者介面，並針對這麼議題設計出它專有的 XML Grammar，所以本研究系統在科技台灣這議題的導覽資訊中，辨識率可以達到一定標準，若將其應用在其它導覽內容可能導致辨識率不高的問題。

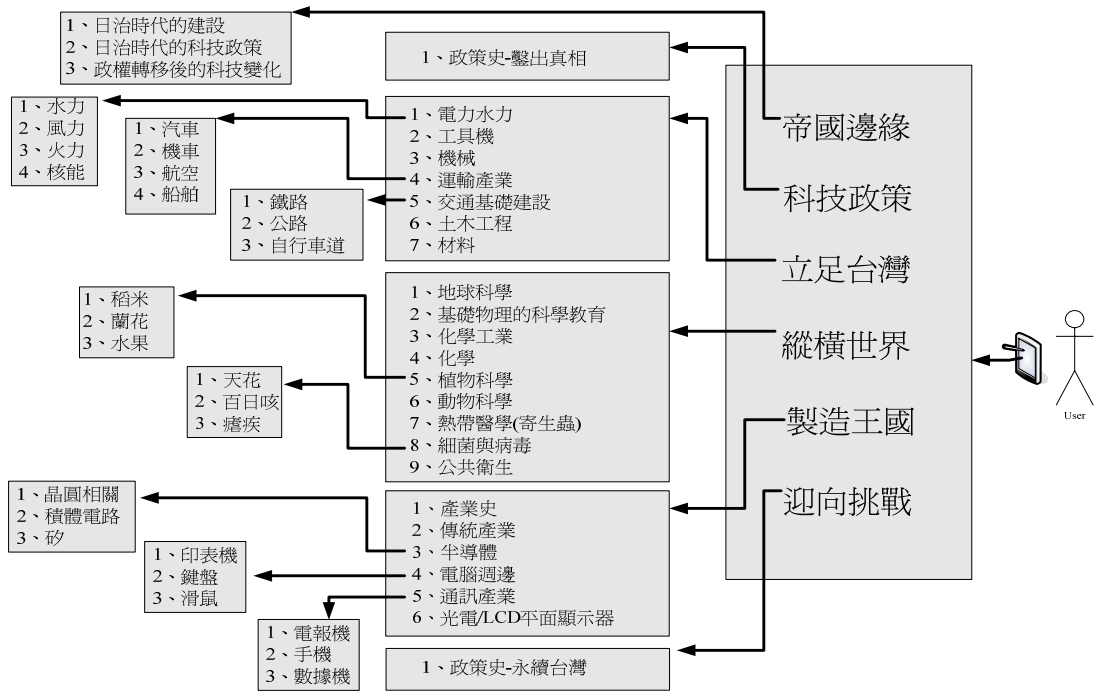


圖 4.1：博物館導覽資訊圖

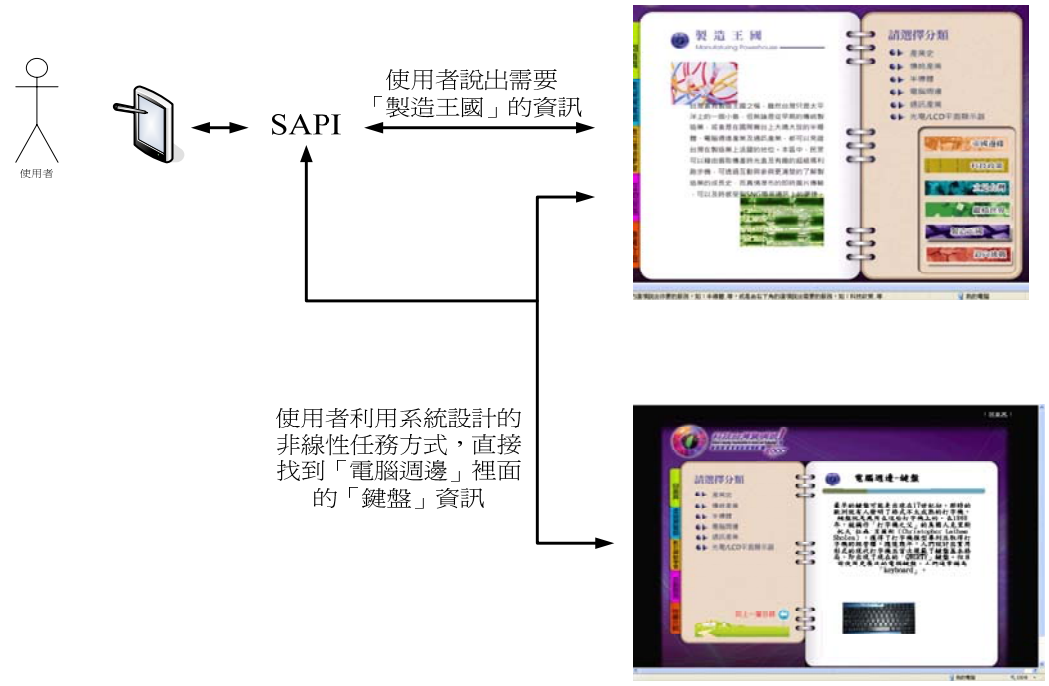


圖 4.2：任務選擇圖

另外，我們也進行了第二階段測試，一開始，我們系統根據第一階段實驗的放聲思考法來進行改善，針對之前系統的缺點提出以下改善，在服務內容方面，我們增加了第四層的子選項，舉例來說，使用者可以從製造王國的電腦週邊裡面查詢到鍵盤的相關資料，提供使用者可以查詢更多資訊；另外，在系統開啟服務方面，我們嘗試修改系統程式內容且增加了相關硬體設施來加速開啟瀏覽畫面；另外辨識率不高問題，我們也修改相關 XML Grammar 檔案，讓辨識率與系統靈活度可以更高，透過以上修改希望可以讓使用者在系統使用上可以更加親切且提供更完善的服務。

在系統使用上，使用者說出「我想查詢運輸產業」，如下圖 4.3 所示，SAPI 會去讀取 XML Grammar 檔案，分別在”want”、”selection”與”place”的 XML Grammar 項目裡面分別去選取「我想」、「查詢」與「運輸產業」的字彙，系統將字彙組成一段文字再根據 SAPI 程式裡面的定義去辨別所需要做的事情，在「我想查詢運輸產業」文字中，系統會將運輸產業的資訊內容以 Web Browser 的方式呈現，使用者便可以取得所需要的服務資訊。

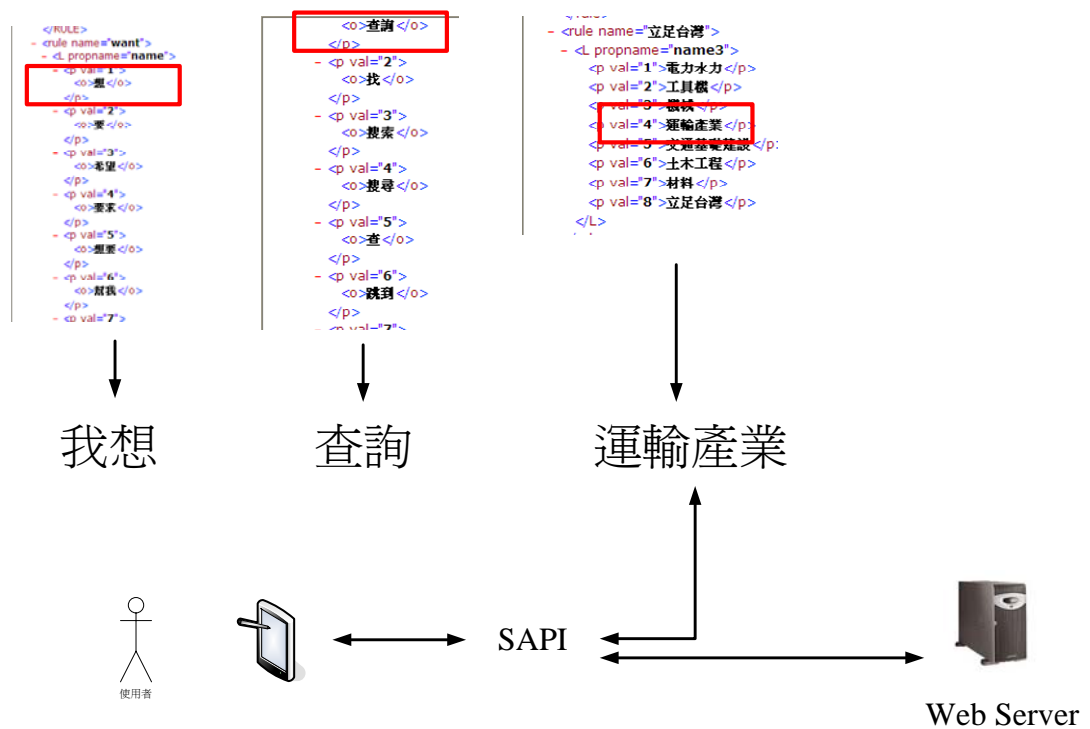


圖 4.3：XML 文法選擇圖

4.2 時間效能檢定

在時間效能檢定方面，我們進行了第二階段實驗，第二階段實驗內我們邀請二十位受測者來進行改良後系統的實驗，一開始我們將受測者採取隨機抽樣的方式來分組，分別為語音使用者介面組與觸控式使用者介面組，而每一組人數為 10 人，分組之後，第一組語音使用者介面組 10 位受測者分別測試語音線性任務與語音非線性任務；第二組觸控式使用者介面組 10 人分別測試觸控式線性任務與觸控式非線性任務，當線性任務與非線性任務皆完成之後，主測者會給予受測

者問卷填寫，並收集使用系統的時間與問卷滿意度調查以做為統計分析用，如下圖 4.4 所示。因為觸控式實用者介面沒有提示功能，為了避免實驗的不公平性，因此在實驗與因使用者介面的時候，我們將 TTS 提示功能拿掉，以符合兩個實驗的公平。

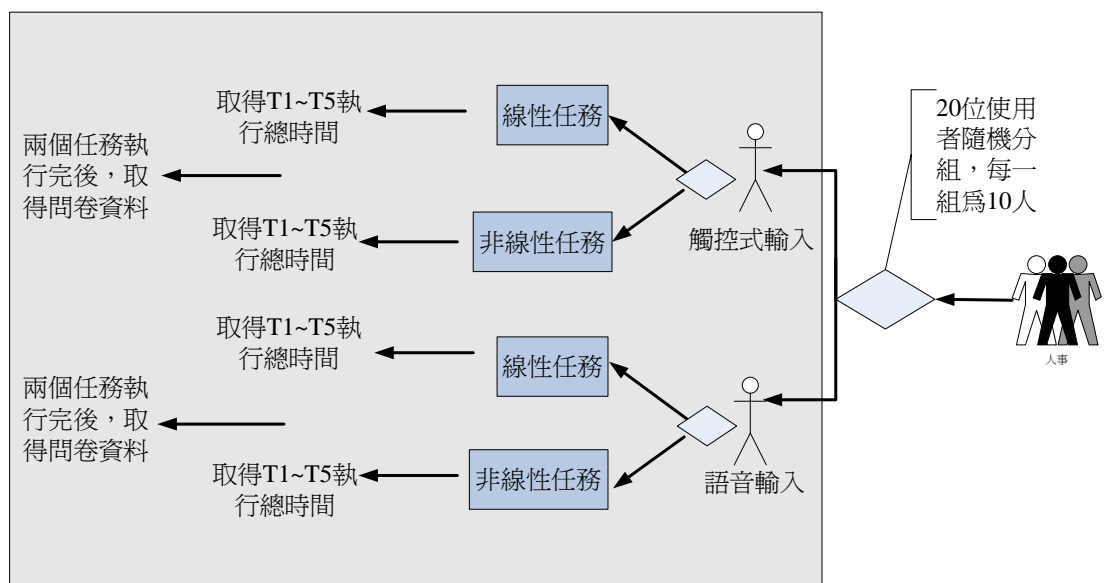


圖 4.4：使用者分組與執行任務圖

我們在表 4.1 會介紹到任務的規劃，系統設計者規劃任務的路徑，透過使用者的參與使用並進行線性與非線性的實驗測試，實驗一開始我們告知使用者這次任務我們所要找的資料為何，讓使用者們自行去判斷該如何透過系統平台快速的取得所要的資訊服務，系統設計者在旁邊紀錄使用者達到任務需求資訊所需要花的總時間。

表 4.1：任務規劃表

實驗一 觸控式使用者介面線性任務的規劃如下：在觸控式-線性任務規劃裡面，我們要求使用者透過觸控式系統平台來取得鍵盤資訊、鐵路資訊、動物科學、日治時代的建設與政策史-鑿出真相等資料內容，系統設計者並在旁邊紀錄達成 $T1+T2+T3+T4+T5$ 任務所需要的總時間。

T1：主選單→製造王國→電腦週邊→鍵盤資訊。

T2：主選單→立足台灣→交通的基礎建設→鐵路。

T3：主選單→縱橫世界→動物科學。

T4：主選單→帝國邊緣→日治時代的建設。

T5：主選單→科技政策→政策史-鑿出真相。

實驗二 觸控式使用者介面非線性任務的規劃如下：在觸控式-非線性任務規劃裡面，我們設計了不同的任務需求，使用者需要從觸控式系統平台之中取得 T1 到 T5 等相關資料，系統設計者會在旁邊紀錄達成任務 $T1+T2+T3+T4+T5$ 的時間。

T1：找尋印表機資訊，接著搜尋 T2

T2：找尋積體電路資訊，接著搜尋 T3

T3：找尋地球科學資訊，接著搜尋 T4

T4：找尋航空資訊，接著搜尋 T5

T5：找尋化學工業資訊。

實驗三 語音使用者介面線性任務的規劃如下：在語音使用者介面-線性任務規劃裡面，我們要求使用者透過語音系統平台來取得晶圓相關資訊、運輸產業、百日咳資訊、日治時代的科技政策與政策史-鑿出真相等資料內容，我們設計不同於觸控式的線性任務，在同階層不同任務底下測試任務所需要完成的時間，系統設計者並在旁邊紀錄達成任務 T1+T2+T3+T4+T5 所需要的總時間。

T1：主選單→製造王國→半導體→晶圓相關。

T2：主選單→立足台灣→運輸產業。

T3：主選單→縱橫世界→細菌與病毒→百日咳。

T4：主選單→帝國邊緣→日治時代的科技政策。

T5：主選單→科技政策→政策史-鑿出真相。

實驗四 語音使用者介面非線性任務的規劃如下：在語音使用者介面-非線性任務規劃裡面，我們設計了不同的任務需求，使用者需要從語音系統平台之中取得 T1 到 T5 等相關資料，系統設計者會在旁邊紀錄達成任務 T1+T2+T3+T4+T5 的時間。

T1：找尋鐵路資訊，接著搜尋 T2

T2：找尋蘭花資訊，接著搜尋 T3

T3：找尋化學工業資訊，接著搜尋 T4

T4：找尋數據機資訊，接著搜尋 T5

T5：找尋土木工程資訊。

上表 4.1 的線性任務規劃表示 T1 到 T5 的時間總和，觸控式使用者介面的線性任務 T1 與 T2 表示使用者需要選取到第四階層的服務內容才算任務結束，其它 T3 到 T5 表示使用者只需要選取到第三階層的服務內容即可，而語音使用者介面下的線性任務也是相同意思；另外，在觸控式非線性的任務規劃，使用者在執行完 T1 的時候需要緊接著進行 T2 的任務，計算使用者執行完 T1 到 T5 所需要的總時間。之後我們針對二十位受測者分別進行隨機分組與進行測試，所得到的時間數據如下表 4.2 到表 4.5，我們依據此任務所時間的統整表來驗證我們之前所作的假設。

表 4.2：時間整理表(秒)

使用者	所需時間			
	線性觸控 式任務	非線性觸 控式任務	線性語 音任務	非線性 語音任務
一	35	95	36	33
二	42	112	34	31
三	38	127	23	36
四	36	103	28	28
五	42	116	22	31
六	43	124	43	49
七	41	101	24	27
八	31	132	31	39
九	42	119	22	36
十	36	98	24	23
平均數	38.6	112.7	28.7	33.3

表 4.2 顯示，語音使用者介面底下的線性任務(28.7 秒)與非線性任務(33.3 秒)分別比觸控式的線性任務(38.6 秒)與非線性任務(112.7 秒)

平均完成任務所需要的秒數更低，驗證了我們的假設 1，語音的使用者介面可以比觸控式使用者介面更加節省時間，也就是語音使用者介面系統筆觸控式使用者介面系統更加有效率；另外在非線性任務底下，非線性語音任務(33.3 秒)也比非線性觸控式任務(112.7 秒)所花費的時間更少，所以假設 2，在非線性任務的安排底下，利用語音使用者介面瀏覽系統會比利用觸控式使用者介面瀏覽系統更加有效率在這邊也是成立的。

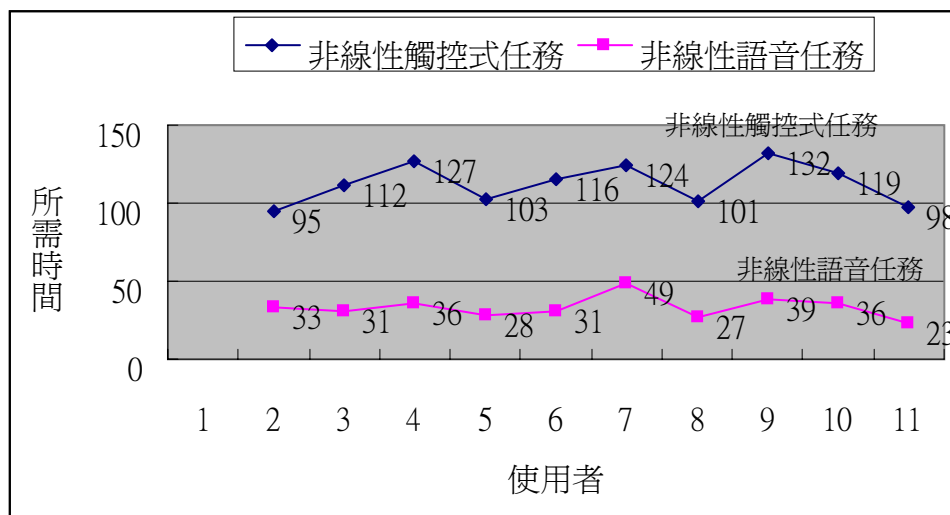


圖 4.5：非線性任務所需時間圖

另外我們利用 Z 檢定來分析是否假設 1 與假設 2 是否也是正確的，我們令

$\mu 1$ ：線性觸控式任務平均數

$\mu 2$ ：非線性觸控式任務平均數

$\mu 3$ ：線性語音任務平均數

$\mu 4$ ：非線性語音任務平均數

$\mu 5$ ：觸控式任務輸入平均數

$\mu 6$ ：語音任務輸入平均數

由於我們認為在非線性任務底下，語音使用者介面會比觸控式使用者介面更加有效率，所以採單尾檢定，我們的虛無假設與對立假設如下：

$$H_0 : \mu 4 - \mu 2 \geq 0$$

$$H_1 : \mu 4 - \mu 2 < 0$$

表 4.3：非線性任務介面比較表

z 檢定：兩個母體平均數差異檢 定		
	非線性語音任務	非線性觸控式任務
平均數	33.3	112.7
已知的變異數	53.1222	168.4556
觀察值個數	10	10
假設的均數差	0	
z	-16.86775869	
臨界值：單尾	1.644853627*	
臨界值：雙尾	1.959963985*	
臨界值：單尾	2.326347874**	
臨界值：雙尾	2.575829304**	
* p<0.05 , ** p<0.01		

由上表 4.3，我們可以看到當在 95%與 99%的信賴區間底下，Z 值為-16.86775869，小於臨界值，因此我們拒絕虛無假設 H_0 ，所以 $\mu_1 - \mu_2 < 0$ 成立，也就是說，非線性語音控制底下所需要花費的時間平均比非線性觸控式所需要的時間還要少，也證明了我們提出的假設 2，.在非線性任務的安排底下，利用語音使用者介面瀏覽系統會比利

用觸控式使用者介面瀏覽系統更加有效率。

另外，在語音控制方面與觸控式控制方面所花費的時間平均數為以下表 4.4。

表 4.4：兩種介面所需平均時間(秒)

使用者	所需時間(秒)					語音平均
	線性觸控式任務	非線性觸控式任務	觸控式平均	線性語音任務	非線性語音任務	
一	35	95	65	36	33	34.5
二	42	112	77	34	31	32.5
三	38	127	82.5	23	36	29.5
四	36	103	69.5	28	28	28
五	42	116	79	22	31	26.5
六	43	124	83.5	43	49	46
七	41	101	71	24	27	25.5
八	31	132	81.5	31	39	35
九	42	119	80.5	22	36	29
十	36	98	67	24	23	23.5
平均			75.65			31

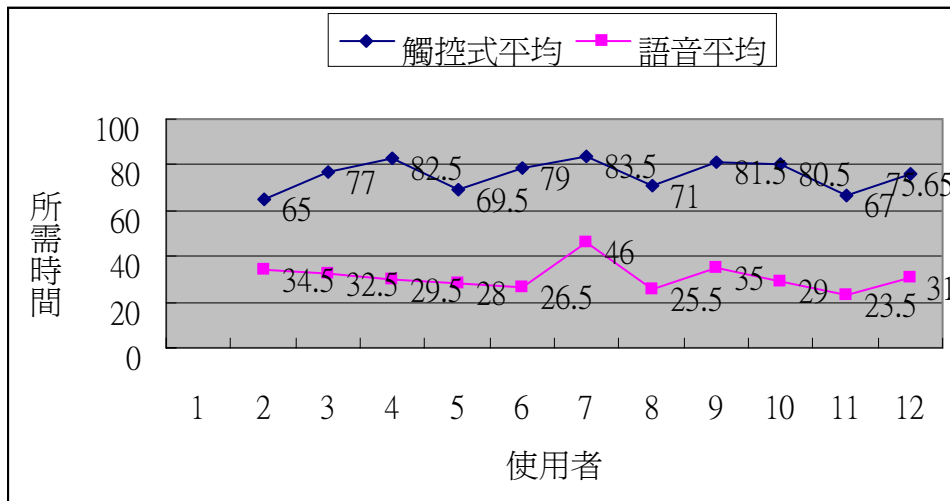


圖 4.6：觸控式輸入與語音輸入的平均花費時間圖

以上表 4.4 所表示，在觸控式使用者介面輸入的平均為 75.65 秒，而語音使用者介面輸入的平均為 31 秒，這跟我們的假設 1，透過語句來控制瀏覽流程，語音的控制介面可以比觸控式使用者介面更節省取得資訊的時間不謀而合。

另外，我們也利用 Z 檢定來分析觸控式使用者介面所花費的時間是否真的會比語音使用者介面所花費的時間更多，我們訂定 μ_5 為觸控式任務輸入時間平均數如下； μ_6 為語音任務輸入時間平均數，如下：

$$\langle \text{線性觸控式輸入時間} + \text{非線性觸控式輸入時間} \rangle \div 2$$

$$\langle \text{線性語音輸入時間} + \text{非線性語音輸入時間} \rangle \div 2$$

所以根據上面所示，我們設定虛無假設與對立假設為以下：

$$H_0 : \mu_6 - \mu_5 \geq 0$$

$$H_1 : \mu_6 - \mu_5 < 0$$

表 4.5：觸控式輸入與語音輸入的 Z 檢定分析

z 檢定：兩個母體平均數差異檢定		
	語音平均	觸控式平均
平均數	31	75.65
已知的變異數	47.44722222	41.94444444
觀察值個數	10	10
假設的均數差	0	
z	-14.93388993	
臨界值：單尾	1.644853627*	
臨界值：雙尾	1.959963985*	
臨界值：單尾	2.326347874**	
臨界值：雙尾	2.575829304**	
* p<0.05 , ** p<0.01		

由上表 4.5 可以知道，Z 值為 -14.93388993，小於 95% 與 99% 信賴區間的臨界值，所以拒絕 H0 接受 H1，也就是說觸控式輸入明顯大於語音輸入，利用 Z 檢定分析證明了我們的假設 1，透過語句來控制瀏覽流程，語音的控制介面可以比觸控式使用者介面更節省取得資訊的時間。

4.3 問卷分析

在問卷設計方面，我們根據黃瓊慧女士在「台灣企業資料倉儲滿意度評估」一文中所用到的 Pearson 與 Bailey 設計的 39 個問卷問項[9]並依照本系統使用環境下設計了以下問卷，以「系統回應時間」、「容易使用」、「正確性」與「系統彈性」四個問項為主軸，並參考賴鼎陞先生等人在「博物館數位導覽系統建置與使用者評估」[12]一文中提到的問卷並加以改良並利用李克特的五點量表分析(Likert-type scale)，針對系統本身滿意度的項目做調查。並根據 Kwan Min Lee et al. 在 Speech Versus Touch[19]一文中，對於使用者介面的滿意度分析，如「系統特性」、「使用者介面的特性」、「使用者介面的操控性」與「使用者介面的使用感覺」等各個項目，調查使用者對於觸控式使用者介面與語音使用者介面的表現滿意度，而項目分數分為非常同意為 5 分、同意為 4 分、普通為 3 分、不同意為 2 分與非常不同意為 1 分，

完成下表 4.3，其中，第 5 題、第 11 題、12 題、14 題與 15 題為反向題，也就是說，如果受測者勾選非常同意的話，得分為 1 分，勾選非常不同意的話，得分為 5 分。問卷設計後我們讓二十位受測者在進行完線性任務與非線性任務實驗之後填寫一份問卷，問卷共十七題，受測者需全數填寫完畢，以便評估使用者對我們系統介面的滿意度調查。

使用者在測試完系統之後，透過附錄的題目來填寫下表的滿意度，我們收集這些數據來做統計分析，以驗證我們之前所做的假設 3，以證明使用者在語音使用者介面上是否有較高的滿意度。我們將附錄所表示的「系統特性」、「使用者介面的特性」、「使用者介面的操控性」與「使用者介面的使用感覺」等主要問項作分數統整，並利用李克特五點量表分析(Likert-type scale)，將其數據整理到下表 4.6，而其中，反向的題目分數的分數，5 分改為 1 分、4 分改為 2 分、3 分則維持 3 分、2 分的改成 4 分、1 分的改為 5 分，在下表我們將更正後的數據以粗體框線顯示。

表 4.6：問卷各個問項的平均值

使用者 題號	一	二	三	四	五	六	七	八	九	十	平均
觸控式使用者介面-系統特性											2.925
1	3	3	4	3	2	3	3	2	3	3	2.9
2	4	3	3	2	3	3	3	4	3	3	3.1
3	4	3	3	4	2	3	3	4	3	2	3.1
4	2	3	2	2	2	3	3	2	4	3	2.6
觸控式使用者介面-使用者介面的特性											2.12
5	2	2	1	2	3	2	2	1	2	3	2
6	2	3	2	3	2	2	2	3	1	2	2.2
7	1	2	2	1	3	2	1	2	2	3	1.9
8	2	2	1	2	3	2	3	1	2	2	2
9	3	3	2	3	2	2	3	3	2	2	2.5
觸控式使用者介面-使用者介面的操控性											2.825
10	4	3	3	4	3	4	3	3	3	3	3.3
11	3	3	3	3	3	4	3	3	3	3	3.1
12	2	3	2	3	3	3	2	4	3	3	2.8
13	2	2	1	2	2	3	2	3	2	2	2.1
觸控式使用者介面-使用者介面的使用感覺											2.55
14	2	3	3	2	3	2	3	1	2	2	2.3
15	3	3	2	3	2	3	2	3	3	3	2.7
16	3	3	3	3	2	3	2	3	3	3	2.8

17	2	3	2	2	2	3	3	2	3	2	2.4	
題號	使用者	一	二	三	四	五	六	七	八	九	十	
語音使用者介面-系統特性											3.75	
1	3	3	4	4	4	4	4	4	4	3	4	3.7
2	4	4	4	4	4	4	5	4	4	4	4	4.1
3	4	3	4	3	3	3	5	3	4	4	4	3.7
4	3	3	3	4	3	3	3	4	4	4	4	3.5
語音使用者介面-使用者介面的特性											4.04	
5	2	4	4	5	5	4	4	3	5	4	4	4
6	5	3	4	5	4	4	4	5	5	3	4	4.2
7	3	4	4	4	4	3	5	4	4	4	3	3.8
8	4	4	4	4	4	3	4	4	5	4	3	3.9
9	4	4	4	4	4	4	4	5	5	4	5	4.3
語音使用者介面-使用者介面的操控性											3.925	
10	4	3	4	4	4	4	5	4	5	4	5	4.2
11	4	3	4	4	4	5	5	4	3	4	4	4
12	4	4	4	4	4	4	2	4	3	4	4	3.7
13	4	3	4	4	4	3	5	4	3	4	4	3.8
語音使用者介面-使用者介面的使用感覺											3.65	
14	3	3	3	3	3	3	4	4	3	4	3	3.3
15	4	4	3	3	3	3	4	4	3	4	3	3.5
16	4	3	4	4	4	4	4	4	4	4	4	3.9

17	4	3	4	4	4	4	4	4	4	4	3.9
----	---	---	---	---	---	---	---	---	---	---	-----

根據上表 4.6，我們可以知道在系統特性方面，語音使用者界面的平均取得分數為 3.75 分，觸控式使用者界面的平均分數為 2.925 分，所以我們可以說，在語音使用者界面底下使用者對於系統特性的滿意度大於觸控式使用者界面；而在使用者界面特性上，語音使用者界面的平均分數為 4.04，而觸控式使用者界面分數為 2.12，故在語音使用者界面底下使用者界面特性的滿意度大於觸控式使用者界面；在使用者界面操控性上，語音使用者界面的平均分數為 3.925，而觸控式使用者界面分數為 2.825，故在語音使用者界面底下使用者界面的操控性滿意度大於觸控式使用者界面；在使用者界面使用感覺上，語音使用者界面的平均分數為 3.65，而觸控式使用者界面分數為 2.55，故在語音使用者界面底下使用者界面的使用感覺滿意度大於觸控式使用者界面。這可以間接驗證我們的假設 3：使用者使用完系統之後，對於語音使用者界面會比觸控式使用者界面有較佳的滿意度，經過上表的平均數驗證，可以證明我們所提出的假設 3 是正確的。

另外，我們也以統計 Z 檢定來分別比較觸控式使用者介面與語音使用者介面之間的差別。在驗證觸控式輸入與語音輸入的滿意度調查之前，我們做了下假設：

$\mu 7$ ：觸控式使用者介面的系統特性平均數

$\mu 8$ ：語音使用者介面的系統特性平均數

$\mu 9$ ：觸控式使用者介面的特性平均數

$\mu 10$ ：語音使用者介面的特性平均數

$\mu 11$ ：觸控式使用者介面的操控性平均數

$\mu 12$ ：語音使用者介面的操控性平均數

$\mu 13$ ：觸控式使用者介面的使用感覺平均數

$\mu 14$ ：語音使用者介面的使用感覺平均數

在系統特性方面，我們分別測試 95% 與 99% 的信賴區間，接受域與拒絕域假設如下：

$$H_0 : \mu 7 - \mu 8 \geq 0$$

$$H_1 : \mu 7 - \mu 8 < 0$$

表 4.7：系統特性檢定表

Z 檢定：兩個母體平均數差異檢定		
	觸控式	語音
平均數	2.925	3.75
已知的變異數	0.055833	0.063333
觀察值個數	4	4
假設的均數差	0	
Z	-4.77978	
臨界值：單尾	1.644854*	
臨界值：雙尾	1.959964*	
臨界值：單尾	2.326348**	
臨界值：雙尾	2.575829**	
* $p < 0.05$, ** $p < 0.01$		

由以上表 4.7 得知，Z 值皆小於 95% 與 99% 的臨界值，故拒絕 H_0 ，接受 H_1 ，相對驗證在系統特性比較下，觸控式使用者介面在滿意度的得分比語音使用者介面來的低，也就說，使用者在系統特性的比較之下，語音使用者介面較優。

在使用者介面特性方面，我們分別測試 95% 與 99% 的信賴區間，
接受域與拒絕域假設如下：

$$H_0 : \mu_9 - \mu_{10} \geq 0$$

$$H_1 : \mu_9 - \mu_{10} < 0$$

由下表 4.8 可以知道，Z 值皆小於 95% 與 99% 的臨界值，故拒絕 H_0 ，接受 H_1 ，証明了在使用者介面特性比較下，觸控式使用者介面在滿意度的得分比語音使用者介面來的低，也就說，使用者在使用者介面特性的比較之下，語音使用者介面較優。

表 4.8：使用者介面檢定表

Z 檢定：兩個母體平均數差異檢定		
	觸控式	語音
平均數	2.12	4.04
已知的變異數	0.057	0.043
觀察值個數	5	5
假設的均數差	0	
Z	-13.5765	
臨界值：單尾	1.644854*	
臨界值：雙尾	1.959964*	
臨界值：單尾	2.326348**	
臨界值：雙尾	2.575829**	
* p<0.05 , ** p<0.01		

在使用者介面操控性方面，我們分別測試 95% 與 99% 的信賴區間，接受域與拒絕域假設如下：

$$H_0 : \mu_{11} - \mu_{12} \geq 0$$

$$H_1 : \mu_{11} - \mu_{12} < 0$$

表 4.9：使用者介面的操控性檢定表

Z 檢定：兩個母體平均數差異檢定		
	觸控式	語音
平均數	2.825	3.925
已知的變異數	0.275833	0.049167
觀察值個數	4	4
假設的均數差	0	
Z	-3.85906	
臨界值：單尾	1.644854*	
臨界值：雙尾	1.959964*	
臨界值：單尾	2.326348**	
臨界值：雙尾	2.575829**	
* $p < 0.05$, ** $p < 0.01$		

由上表 4.9 可以知道，Z 值皆小於 95% 與 99% 臨界值，故拒絕 H_0 ，接受 H_1 ，証明了在使用者介面操控性比較下，觸控式使用者介面在滿意度的得分比語音使用者介面來的低，也就說，使用者在使用者介面操控性的比較之下，語音使用者介面較優。

在使用者介面使用感覺方面，我們分別測試 95% 與 99% 的信賴區間，接受域與拒絕域假設如下：

$$H_0 : \mu_{11} - \mu_{12} \geq 0$$

$$H_1 : \mu_{11} - \mu_{12} < 0$$

表 4.10：使用者介面的使用感覺檢定表

Z 檢定：兩個母體平均數差異檢定		
	觸控式	語音
平均數	2.55	3.65
已知的變異數	0.056667	0.09
觀察值個數	4	4
假設的均數差	0	
Z	-5.74456	
臨界值：單尾	1.644854*	
臨界值：雙尾	1.959964*	
臨界值：單尾	2.326348**	
臨界值：雙尾	2.575829**	
* p<0.05 , ** p<0.01		

由上表 4.10 可以知道 Z 值皆小於 95% 與 99% 臨界值，故拒絕 H_0 ，接受 H_1 ，証明了在使用者介面使用感覺比較下，觸控式使用者介面在滿意度的得分比語音使用者介面來的低，也就說，使用者在使用者介面使用感覺的比較之下，語音使用者介面較優。

藉由以上四個問項的比較，我們可以知道語音使用者介面在使用者的心中表現比觸控式使用者介面還要好，跟我們所提出的假設 3：使用者使用完系統之後，對於語音使用者介面會比觸控式使用者介面有較佳的滿意度相符，故驗證了我們的假設。經過一開始的時間檢定驗證與問卷的檢定驗證，皆符合我們提出的假設 1、假設 2 與假設 3，證明了我們所設計的語音使用者介面的確是比現有觸控式使用者介面還要好。

第五章、結論

在語音使用者介面的開發過程，本系統經過兩次的實驗，第一次為利用放聲思考法來改善系統缺點，根據使用者對於系統的意見作為參考並修改系統服務內容，希冀可以讓語音辨識的服務真正加入 Kiosk 系統之中，這樣使用者可以透過新的服務方式來認識並取得需要的資訊，透過此系統的呈現，讓使用者利用語音來控制系統，並在使用者測試後參照使用者意見加以修改系統，以便更能符合需求；而第二次的實驗中，我們分別測試了觸控式使用者介面與語音使用者介面與在非線性任務底下，觸控式使用者介面與語音使用者介面之間的差別，透過統計 Z 檢定的分析，我們可以輕易的看出語音使用者介面的加入，真的可以讓 Kiosk 導覽系統更加方便使用與更有效率。

為了改善觸控互動式資訊站介面的缺失，我們著手設計了這一個語音互動式導覽系統介面，一開始我們希望系統可以符合使用者使用上的需求，聆聽使用者對於系統的意見是非常重要的，所以我們利用放聲思考法收集了使用者對於系統的意見，使用者對於系統的看法不外乎有，希望可以加入插話機制、希望回應速度可以更加快速、希望服務的內容可以在增多一點...等，另外，使用者對於系統的辨識率也

有所稱讚，所以，我們針對以上使用者所提供的意見內容修改系統內容，完成修改之後，使用者對於系統改善後的滿意度也有所增加。

在完成此系統的設計與驗證我們所做的三個假設，並與現有觸控式使用者介面相比較，根據我們的假設，語音使用者介面有較佳的使用效率，這與我們的假設一相符，而在非線性任務的處理下，經實驗證明我們所設計的語音輸入比觸控式輸入更能節省時間，這與我們提出的假設二相符，而在滿意度方面，經過問卷分析我們得知語音輸入的滿意度大於觸控式輸入，我們推測可能是因為語音使用者介面可能給予使用者較新奇的使用感覺，所以使用者在滿意度方面給予語音輸入較高的分數，這也與我們的假設三相符，證明我們所設計的語音使用者介面比傳統觸控式的輸入方式相比較為較佳。

我們發現語音使用者介面可以提供系統使用者更加方便的搜尋效率，與更佳的使用滿意度，現有觸控式系統所能提供的服務在我們新設計的語音使用者介面裡面都有提供，透過我們所設計的語音使用者介面，使用者可以一進入博物館之後並透過 Kiosk 使用本系統，可以利用口語命令來取得所需要的服務資訊，我們實作這個系統也是要让使用者可以透過另一種輸入方式輕鬆且省時的取得需要的資訊，也測試了解使用者對於語音辨識系統建置在 Kiosk 平台上面的滿意度的

高低。經過一開始的介面設計的相關資料收集與兩次實驗後的資料與數據收集並完成此系統，也許這個系統還有一些考慮不周的地方，我們希望後來的研究者可以收集更多使用者對於系統的看法並加以改善，希冀這一個語音使用者介面可以真正達到使用者所想要的需求。

第六章、問題與討論

在 Kiosk 系統發展過程中，語音技術的相關應用是相當稀少，以現行的 Kiosk 應用來說，ATM 提款機為最常見也是最成功的案例，不過 ATM 使用到的語音技術為 TTS 功能加上觸控式控制，而台灣鐵路的 Kiosk 訂票系統也是利用觸控式取票，而語音 TTS 功能以只是輔助讓取票功能不至於太單調，我們在這次研究裡面加入一個語音使用者介面，讓 Kiosk 系統除了可以說出語音的 Output 功能，有可以讓使用者利用口語 Input 輸入需要的服務項目，進而達到人機互動的理念。

那為什麼語音使用者介面的 Kiosk 導覽系統不常見呢？我們推論出幾個原因如下：

- 1、Kiosk 建置的成本過高問題。
- 2、語音辨識時的背景噪音問題。
- 3、使用者的使用習慣。

在問題一成本過高的問題，我們收集過相關資料，在微軟出版的 ATM 與 Kiosk 白皮書一文中指出[13]，建置一套最基礎型的 Kiosk 系統大約需要 3000 到 15000 美元之間，比較專業的型號可能就需要 20000 到 40000 美元之間，如下表 6.1 粗框線所示：

表 6.1：建置 Kiosk 需要的成本表[13]

機台特色	基礎功能 ATMs	普通功能 ATMs	全功能 ATMs
所需花費	3000~15000 美元	12000~20000 美元	20000~40000 美元
平台	Kiosk 平台	PC 平台	PC 平台
放款功能	可以	可以	可以
存款功能	否	否	可以
與公司產品整合	否	可以	可以
功能擴充	否	有限制的	可以
作業系統	Windows CE.NET, Windows XP Embedded, 或 Windows XP Professional	Windows XP Embedded 或 Windows XP Professional	Windows XP Embedded 或 Windows XP Professional

一般除非說真正需要 Kiosk 系統，如銀行 ATM 提款機，企業才會投入大量成本去購買，所以有關成本問題，我們建議有賴於大量採購或是一些專案的執行才可能解決成本建置過高的問題。在本實驗測試的時候，我們的測試環境為南華大學研究所實驗室，而在博物館內放置 Kiosk 導覽系統的話，我們必須考慮到使用者在開放式的環境底下背景噪音的問題，在本研究裡面我們著重在討論語音使用者介面的設計，所以不考慮背景噪音是否會影響到使用者進行語音辨識時候的辨識問題，如果要將背景噪音問題考慮進去，我們建議想進行語音研究的先進們，可以將噪音問題考慮進去。在使用者習慣問題的解決方案，我們認為可以對使用者進行簡單的教育訓練或是在 Kiosk 平台上面放置使用教學，讓有興趣使用 Kiosk 導覽系統的使用者，可以透過事前的簡單教育訓練與使用教學讓使用者可以輕易的使用 Kiosk 導覽系統。

另外，如果有後續的研究者想繼續研究語音辨識技術與 Kiosk 平台的結合，我們會建議研究者可以將導覽服務的資訊內容建置在行動式的裝置上面，如行動式 PDA、智慧型手機、其他行動式裝置...等，這樣使用者就可以隨時隨地的接收到想要得知的資訊，因為 Kiosk 建置需除了成本過高的問題之外，還有一點為 Kiosk 機台過於龐大，使用者需要到特定的地點才可以使用的機器，然而，若是使用行

動裝置的話，現在幾乎人手一機的情況底下，可以提供使用者更加方便的查詢方式，試想，透過手機與語音辨識技術的結合，使用者對手機講出需要的資訊服務，系統會將資訊傳送到使用者的行動裝置，並將導覽的資訊內容說出來，這相對於 Kiosk 系統而言，建置的費用大大減低，也可以提供每位使用者都可以使用到語音服務的系統。

最後，本系統在開發過程中使用的 TTS 引擎目前只有支援簡體中文發音，也就是說 TTS 發聲引擎不認得繁體中文，我們建議之後的研究者可以使用微軟的 Speech Server 或是 SpeechSDK 6.1 以上更高的版本來作為開發引擎，不過也相對的需要去支付微軟辨識引擎的費用，如果作為學術用途，也可以使用 VoiceXML Gateway[24]來作為開發平台，可以降低系統開發所需要的成本花費並將平台運用在 IP-based PSTN 上面，也就是說使用者可望透過網路電話方式撥打進去取得相關瀏覽資訊。

參考文獻

一、中文部份

1. 王小川，「語音訊號處理」，全華科技圖書股份有限公司，初版，民 93 年。
2. 廖峻峰，「語音技術與 VoiceXML 應用」，NCCU Computer Center，June 22 2003
3. 朱育民，「應用語音技術建置一個國小英語教學之學習護照系統」，全國碩博士論文網，國立中央大學網路學習科技研究所。
4. 朱孝國的筆記本，<http://irw.ncit.edu.tw/peterju/>，參照時間：2007 年 10 月。
5. 曾鈺涓，「經驗無線-博物館數位導覽系統案例研究」，國立臺中技術學院主辦之 2005《數位設計研討會》，民 94 年，台中。
6. 林士翔，陳柏琳，「現階段強健性語音辨識技術研究之簡介」，國立台灣師範大學資訊工程學系，Web2.0 技術與應用研討會，參照時間：2007 年 12 月。
7. 科技台灣驚嘆號-2007 科學季
<http://www.2007science.com.tw/vm/Default.html>。
8. 黃光男，「博物館行銷策略」，台北市：藝術家出版社。1998，10 月再版。
9. 黃瓊慧，「台灣企業資料倉儲滿意度評估」，國立中央大學資訊管理學系，參照時間：2008 年 2 月。
10. 維基百科 <http://zh.wikipedia.org/>。
11. 謝金順、張武成，「線上遊戲軟體設計因素與使用者滿意度關聯之研究」，資訊管理展望，第五卷，第 1 期，民國九十二年三月。
12. 賴鼎陞¹、高淑惠¹、黃雅惠²「博物館數位導覽系統建置與使用者評估」，國立故宮博物院資訊中心¹、宏碁數位藝術中心²，國科會專案計畫 (NSC92-3114-S-136-001)。
13. ATM 與 Kiosk 白皮書
<http://www.microsoft.com/taiwan/windows/embedded/vertical/kioskatm.aspx>，參照時間：2008 年 2 月。

二、西文部份

14. Adam Steele, Kandarp Mehta, "Design and Evaluation of Voice User Interface", School of CTI 243 S. Wabash Avenue, Chicago, IL 60604.
15. Daniela Oria, Akos Vetek, "Automatic Generation of Speech Interfaces for Web-based Application", Nokia Research Center, Helsinki, Finland.
16. Dirk Schnelle, Fernando Lyardet and Tao Wei, Telecooperation Group Darmstadt, 2005, "Audio Navigation Patterns", Darmstadt University of Technology.
17. Hirohiko Sagawa, Teruko Mitamura, Eric Nyberg, "Correction Grammars for Error Handling in a Speech Dialog System", Language Technologies Institute,

Carnegie Mellon University.

18. James A. Rodger , Parag C. Pendharkar ,“A field study of the impact of gender and user’s technical experience on the performance of voice-activated medical tracking application”, a Eberly College of Business & Information Technology, Indiana University of Pennsylvania, MIS and Decision Sciences, Indiana, PA 15705, USA.
19. Kwan Min Lee , Jennifer Lai , “Speech Versus Touch : A Comparative Study of the Use of Speech and DTMF keypad for Navigation” , Journal of Human Computer Interaction IJHCI, Vol. 19, No. 3, 2005.
20. L. Lamel , S. Bannacef , J.L. Gauvain, H. Dartigues , J.N. Temem ”, User evaluation of the MASK Kiosk”, 2002, Speech Communication Volume 38, Issues 1-2, September 2002, Pages 131-139.
21. Microsoft Speech SDK <http://www.microsoft.com/speech/download/sdk51/>.
22. Min-Jen Tsai , “VoiceXML dialog system of the multimodal IP-Telephone – The application for voice ordering service”, Institute of Information Management, National Chiao Tung University, 1001 Ta-Hsueh Road, Hsin-Chu, 300, Taiwan, ROC.
23. Nobuo Hataoka , Yasunari Obuchi , Teruko Mitamura , Eric Nyberg , “Robust Speech Dialog Interface for Car Telematics Service”, Consumer Communications and Networking Conference, 2004. CCNC 2004. First IEEE.
24. W3C VoiceXML <http://www.voicexml.org>.
25. WEI ZHANG , VINCENT G. DUFFY , RICHARD LINN , “Voice Recognition Based Human-Computer Interface Design”, Department of Industrial Engineering and Engineering Management, Hong Kong University of Science and Technology , Clear Water Bay, Kowloon , Hong Kong.
26. Yandong Fan , Elizabeth A. Kendall ,“A Hybrid Dialogue Strategy for Speech-enabled Mobile Commerce”, Proceedings of the 2005 IEEE international Conference on e-Business Engineering(ICEBE’05).

附錄

附錄一：語音(觸控式)使用者滿意度調查之問卷題目

這是一份有關「觸控式與語音使用者介面的相關比較」的學術問卷，主要目的是在瞭解您使用過本研究設計的系統的一些經驗與看法，故本問卷填答者須具備使用過本研究所設計的系統之經驗。本問卷內容僅供學術分析使用，不涉及任何商業行為，敬請安心填答。您的意見對本研究相當寶貴，您所填寫的資料僅供本研究學術使用，保證絕對不對外公開，謝謝您的大力合作。在此感謝您熱心的參與。

敬祝您 身體健康 萬事如意

系統特性	非常同意 (5)	同意 (4)	普通 (3)	不同意 (2)	非常不同意 (1)
1. 利用此語音(觸控式)系統來取得資料時，對於回應速度感覺滿意？					
2. 使用語音(觸控式)系統瀏覽資訊，可以容易取得需要的服務？					
3. 使用語音(觸控式)系統可以正確取得需要的服務資訊？					
4. 使用語音(觸控式)系統，可以有彈性且不需要依照畫面指示來選取服務？					
使用者介面的特性	非常同意 (5)	同意 (4)	普通 (3)	不同意 (2)	非常不同意 (1)
5. 使用語音(觸控式)使用者介面瀏覽，覺得此介面是無聊的？					
6. 使用語音(觸控式)使用者介面瀏覽，覺得此介面蠻新奇的？					
7. 使用語音(觸控式)使用者介面瀏覽，覺得使用起來具娛樂性的？					
8. 使用語音(觸控式)使用者介面瀏覽，覺得此介面是好玩的？					
9. 使用語音(觸控式)使用者介					

面瀏覽，對此介面是非常感到興趣的？					
使用者介面的操控性	非常同意 (5)	同意 (4)	普通 (3)	不同意 (2)	非常不同意 (1)
10. 對於語音(觸控式)使用者介面的使用感覺容易？					
11. 對於語音(觸控式)使用者介面，覺得使用起來很複雜？					
12. 使用語音(觸控式)使用者介面瀏覽，感覺容易混淆？					
13. 對於語音(觸控式)使用者介面，使用起來可以很直覺？					
使用者介面的使用感覺	非常同意 (5)	同意 (4)	普通 (3)	不同意 (2)	非常不同意 (1)
14. 總體而言，對於語音(觸控式)使用者介面的使用感到疲累？					
15. 總體而言，對於語音(觸控式)使用者介面的使用有挫敗感？					
16. 總體而言，對於語音(觸控式)使用者介面感覺到滿意？					
17. 總體而言，對於語音(觸控式)使用者介面感到舒適？					