

# 媒介空間<sup>1</sup>中的凝視、語言和意義： 以 Skype 視訊為例

黃齡儀

國立政治大學新聞研究所

95451024@nccu.edu.tw

## 摘要

新傳播科技和社會的關係為何？不同於鉅觀觀點，本文欲以微觀觀點描繪新傳播科技所形塑的人與人互動形式，從改變的互動形式中，希望能回饋到人機互動設計中。

本文以新傳播科技—Webcam 在 Skype 上的應用為例，以機緣理論——檢視其生物、文化面向的限制和創意，檢視的重點則放在非口語傳播元素—視覺。應用 Webcam 這項新科技如何改變了人們互動時的凝視規則，視覺線索等等，人們藉著體現互動(embodied interaction)又如何從中產生創意，創造出新的凝視規則，將是本文探討的重點。

資料分析則藉著對話分析和凝視過錄厚描在 Webcam 下的互動形式。研究發現 Webcam 的媒介空間所產生的特殊空間感，的確影響了視覺的感知，並影響了人們溝通的內容和方式。另外，也由於 Webcam 中介的溝通，使的凝視規則改變，看與不看、如何看都值得人機互動設計者深思。

最後，回到傳播的本質也就是「意義從何而來」的問題。新傳播科技中的意義不再獨尊語言，意義是由各個元素（口語和非口語）如同交響樂團般互相演奏、配合交織而成，並沒有何種元素比哪種元素更為重要。除了呼應實踐語言的意涵外，也再一次強調了各種非口語傳播元素在人機互動設計中的重要性。

**關鍵詞：**人機互動、機緣理論、非口語傳播、實踐語言

[收稿]2007/07/31; [接受刊登] 2007/12/25

---

<sup>1</sup> 引用自Paul Dourish (2004) 的定義，指的是結合音訊、視訊的溝通合作網絡和計算環境。不同於VR，實體和非實體世界間的互動是相互交織影響的。

## 一、前言

### (一) 研究動機

科技和人的關係為何？從麥克魯漢的科技決定論、雷蒙威廉斯的社會決定論和拉圖的行動者網絡理論，都不能滿足我對科技和人關係的想像。沒有那麼大企圖構聯社會和科技的關係，然而，卻深深對機緣理論<sup>2</sup>中所描繪科技所形塑人與人互動的形式深深著迷。

科技改變了人與人互動的形式，科技中特有的機緣，是如何被人們感知、挪用和創造出新奇的互動形式呢？電話的發明，使人們創造了獨一無二的電話開場白。「喂，請問你哪位」這句話雖然簡單，但卻道盡了科技與人之間剪不斷、理還亂的糾葛。種種因電話而改變的人際互動銘刻在我們的日常生活當中。當我們講電話時傾向拉大音量、拉長通話時間以及遠距親密的感受，不正都是機緣理論的最佳實證嗎？

然而，傳播學門的確是獨沽於口語的研究上<sup>3</sup>。科技不僅僅改變了人們口語的互動，更進一步來看，它更不知不覺中形塑了人們非口語的互動，例如：手勢、音調、凝視和臉部表情等等。當我們透過網路攝影機和另一方互動時，最常見的情形是彼此間的注視減少了。不同於面對面互動，我們似乎甚少把心力放在說話的另一方。或者我們傾向做出誇大的表情來吸引對方的注意。以上小小的例子似乎透露了一個被我們慣常忽略的領域，非口語傳播的重要性。

非口語傳播的研究為何在科技與人的關係間佔有一定的重要性？因為非口語傳播的研究可以作為人機互動設計的起點，或者使成爲一個更加完善的傳播系統。例如：當我們了解了手勢在人們認知及互動的重要性，手勢元素的加入可以促進人機之間的溝通，就像Click在微軟系統中的手勢作用。又或者當我們了解更多人們運動中的手勢、身體和動作，就能設計出像Wii<sup>4</sup>一樣，不需閱讀操作手冊就能立

<sup>2</sup> 引用自James J. Gibson以及Hutchby的定義。

<sup>3</sup> 這個可從傳播研究及各大學院校中課程比例中看出，大部分的研究課程都是著重在語言及口語的研究上。

<sup>4</sup> 引用自維基百科：Wii是任天堂公司所推出的第五部家用遊戲主機，GameCube

刻上手的電玩遊戲。

在這篇論文中，筆者希望藉由檢視視訊互動系統中（VMC）人際間的非口語傳播互動來做為未來人機互動設計的參考。然而，囿於能力及篇幅，筆者希望聚焦於凝視（gaze）這項非口語傳播上，尤其是在現有科技介面所賦予的機緣下，人們在其中所產生的互動，並在此建構出的意義，進一步反饋實踐語言<sup>5</sup>的重要性。

## （二）研究背景

既然非口語傳播的研究可以作為人機互動設計的起點。然而，為何筆者特意挑了凝視呢？也就是為什麼要研究凝視？

第一、凝視在人際互動中，尤其是對於Webcam逐漸發展出的親密關係互動佔有重要的地位。<sup>6</sup>首先，Efran(1968)對凝視在人際互動中的解釋是：當我們越喜歡對方，我們越會去凝視他。另外，也有不少學者研究凝視對於親密關係的影響。Argyle & Dean(1965)更發展出了親密關係的公式，認為親密關係是眼神接觸、身體距離、對主題的熟悉程度以及微笑的次數等等的函數。然而，我們都有的經驗是，當使用Webcam和親密的對方對話，有時大部分我們不會使用凝視這個動作，取而代之的是瀏覽其他網頁，甚至盯著螢幕中所呈現的自己，而不是盯著對方的一舉一動。透過Webcam所中介的凝視為什麼和日常親密關係的凝視不同呢？或者說，凝視既然在親密關係互動中佔了這麼重要的位置，但為何在使用了Webcam後的凝視，似乎有某些地方改變了呢？

---

的後繼機種。Wii最與眾不同的特色是它的標準控制器「Wii Remote」。Wii Remote的外型為棒狀，就如同電視遙控器一樣，可單手操作。除了像一般遙控器可以用按鈕來控制，它還有兩項功能：指向定位及動作感應。前者就如同光線槍或滑鼠一般可以控制螢幕上的游標，後者可偵測三維空間當中的移動及旋轉，結合兩者可以達成所謂的「體感操作」。

<sup>5</sup> 實踐語言是由鍾蔚文老師在言說分析課堂上提出，指的是在傳播意義的建構中，空間、工具和其他非口語傳播元素的重要性。可參考Resnick, Lauren B (1997). *Discourse, tools, and reasoning: essays on situated cognition*. New York: Springer.

<sup>6</sup> Webcam從早期的視訊會議到現在的Live Show，逐漸發展出以日常生活和性活動為主的主题。

可參看 *Narratives across New Media* 中對 Webcam 的敘述。

研究凝視的第二個原因是：凝視屬於視覺感知（visual perception），而視覺感知的研究在Gibson的努力下，歷經了典範的轉移。<sup>7</sup>在這個新的典範下，Gibson將視覺感知和空間的關連性緊緊扣合，並以下列三項論點加以論證：環境表面配置的直接感知（the direct perception of surface layout），改變中的環境表面配置的直接感知（the direct perception of changing surface layout），以及自我移動的直接感知（the direct perception of the movements of the self）。<sup>8</sup>因此，由Gibson的取徑來看待媒介空間中的凝視，不同於以往的看法，必須強調媒介空間的配置對視覺感知的影響，也就是強調物、人和空間之間的關係。從這個取徑來看，又更加呼應了實踐語言的概念<sup>9</sup>，也就是在特定空間中的動作、物都是傳播意義建構中所不可或缺的元素。因此，由凝視的概念不僅可以彰顯空間、物的重要性，更可以從中看出人如何在這些機緣下與科技互動，並進一步看出意義如何透過空間、物以及人們的感知而建構出來。

然而，由Gibson的取徑來看凝視，似乎偏重於本質性、生物性機緣的面向，卻忽略了某些文化面向的機緣。<sup>10</sup>然而，媒介空間中什麼樣的視覺感知是屬於Norman所提的範疇呢？媒介空間是由實體和非實體空間交織而成，且之中的互動會相互反饋。因此，互動雙方所凝視的不僅僅是實體空間中的攝影機、電腦，還包含非實體空間中所再現出的影像，包括對方的影像、自己的影像以及出現在所有電腦螢幕上的顯示。以Skype來說，出現在螢幕上的影像可以包括自己的影像、他人的影像和文字介面的顯示。而自己的影像或他人的影像皆可任意調整大小。<sup>11</sup>因此，螢幕上所呈現的各種影像，包含攝影機的角度、

---

<sup>7</sup> 引自Gordon(1989). Theories of visual perception. St. Edmundsbury Press, Great Britain. Pp.146

<sup>8</sup> 引自Gibson(1986). The ecological approach to visual perception. Lawrence Erlbaum Associates, Inc., USA.

<sup>9</sup> Hutchins& Palen (1997). Constructing Meaning from Space, Gesture, and Speech. In Lauren, Roger& Clotilde(Eds.), Discourse, Tools, and Reasoning(pp22-39).Springer, Germany.

<sup>10</sup> Norman對Affordance做的修正，人們是帶著自己的目標、價值、信念和使用的過去經驗來感知，因此設計者必須學習了解使用者約定成俗的使用方式。

<sup>11</sup> 比較特別的是，對方的顯示可以調整到全螢幕，然而自己的確沒有辦法調整至全螢幕。另外，自己的影像也往往被顯示在右下方。

影像顯示在螢幕中的配置以及大小等，皆可以屬於文化面向機緣的展現方式。

另外，Goodwin(2001)<sup>12</sup>提及互動中的影像呈現（image in interaction），包含了在活動中不斷變化、部署甚至彼此影響的各種符號場域（multiple semiotic fields），而這各種符號場域指的就是在不同媒介上呈現的符號（signs-in-their-media）。包括科學家在互動過程中所使用的圖表，法官在質問犯人所使用的錄影帶證據等，都是不同符號場域的例子。而這些不同場域中的影像互動，面對面的、圖表的或影像的，皆在整個活動過程中不停交織，並生產相關的對話及傳播意義。由Norman及Goodwin來補充Gibson所沒有強調的人造物視覺感知，也將視覺感知的討論擴充到再現的部份。因此，本研究除了觀察實體空間中配置對視覺感知的影響，也會觀察在非實體空間中所呈現影像所帶來的文化機緣。

### （三）研究問題

在面對面互動中，不僅人際間的對話存在著規則，就連如何看對方，看與不看之間似乎都存在著一種不言自明的默契。當我們走在路上，感知到一個陌生人的凝視，會有某種不舒服的感覺，而這種感覺使我們避開和他可能的眼神交會。注視與否、如何注視以及注視所造成的人際效果成爲了非口語傳播研究所重視的主題。

然而，在認知心理學中，視覺線索所帶來的空間感更是一個重要的課題。在不同空間中凝視會帶來不同的感知。因此，認知心理學更將凝視和空間緊密的結合在一起。

媒介空間是本研究的主題。在媒介空間中，人們藉由視訊和音訊擴展了自己的空間，而實體與非實體的空間卻是彼此交織，不斷反饋的互動。人們透過網路攝影機看彼此，人們也藉由網路攝影機之間的影像進行互動。和面對面互動不同，人們不是直接注視著對方，人們是透過科技的小孔注視或不注視。顯然的，媒介空間中的關於「看」

---

<sup>12</sup> Goodwin(2001).Practices of seeing Visual analysis: an Ethnomethodological Approach. In Leeuwen&Jewitt, *Handbook of Visual Analysis*. (pp.157-181).London, SAGE.

的學問比想像中複雜許多。注視，這項非口語傳播元素，實踐語言之一，在這媒介空間中扮演了什麼樣的角色，是不是能顛覆我們平常對非口語語言從屬地位的想像呢？

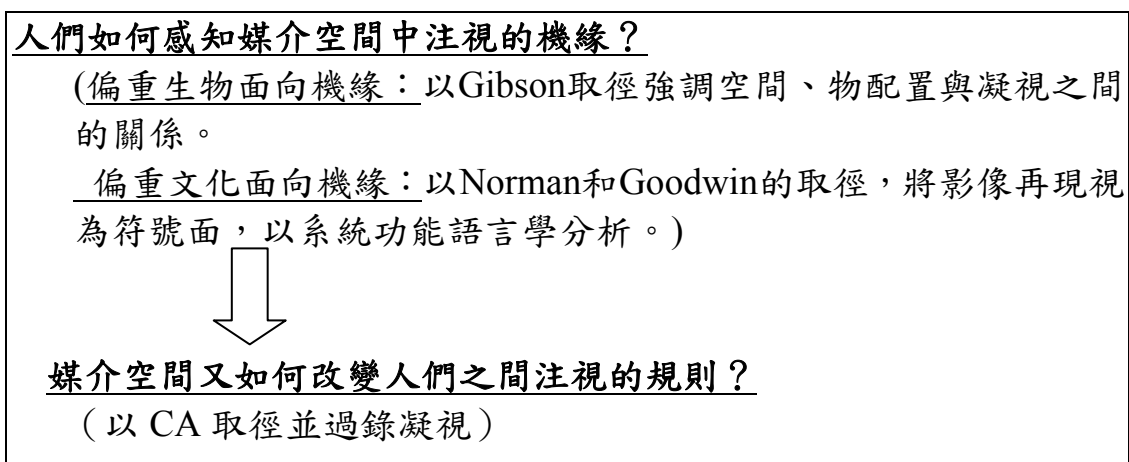
因此，本研究問題為：人們如何感知媒介空間中「注視」的機緣，而媒介空間又如何改變人們之間「注視」的規則。「注視」、「媒介空間」在整個意義建構中，又扮演了什麼樣的角色？本研究希望藉由 Skype 視訊的檢視，回答上列問題，並進一步回饋 HCI 中關於「注視」的相關設計，以及傳播研究中實踐語言的重要性。

#### (四) 研究架構

以下為研究架構圖：

「注視」、「媒介空間」、「語言」在整個意義建構中，扮演了什麼樣的角色？

表 1 研究架構圖



## 二、文獻探討

### (一) Gibson 的機緣理論 (Affordance Theory) 和視覺感知 (Visual Perception)

機緣是 Gibson 用來解釋生物和環境之間關聯的理論。機緣不僅僅和物的物理特性有關，例如它呈現水平、垂直或形狀，更是相對於生物而言。Gibson 舉出一個例子來說明相對性：當一個物體高於地面

且到膝蓋高的距離，那麼這個物體提供了（afford）「坐」的機緣。然而，對於不同身高的人們（例如：大人和小孩），小孩則無法感知到「坐」的機緣，對這個小孩而言，這個物體所提供的有可能是「爬」的機緣。因此，Gibson 援引生態學者對利基（niche）的解釋，認為利基就是一連串的機緣（A niche is a set of affordances）。不同動物有不同的利基，而不同的動物也暗示著不同的利基。機緣既不是全然客觀的物體特性，也不全然是主觀的感知，機緣打破了傳統人物二分的特性，它可以說是兩者皆含，也可以說介於兩者之間。機緣同時指向著環境和它的觀察者。因此機緣理論開創了一個人和物之間互動的新視野。

然而，Gibson 舉出了一個郵筒的例子，卻令人懷疑此「物」非彼「物」。Gibson 舉出了郵筒

提供了一個寫信的人寄信的機緣。然而，此物為人造物，而非自然物。一個人造物的機緣難道不需文化、歷史面向的討論就能如此推斷？當一個寫信的人知道可以使用郵筒寄信，豈和小時候媽媽曾教過他，或和郵筒上面的標示、語言無關？延續著 Gibson 對機緣的討論，當代的學者 Norman 即認為人造物的機緣必須著重文化面向的討論。而這在下節會探討。

Gibson 另一個開創的典範即是視覺感知中對空間的重視。<sup>13</sup>Gibson 在第二次世界大戰中負責駕駛員的訓練工作。然而，他發現在心理學家利用景深等畫畫的理論，根本無法用在飛行員實際起飛和降落的情況。他發現飛行中不斷移動的空間感才是真正的訓練重點。也因此他在 1950 年出版的「The perception of the visual world」成為了空間感感知的重要經典。

Gibson 曾舉出下列四個視線（vision）來說明視覺感知和空間感、移動感的關聯。這四種分別為快照視線（snap shot vision）、縫隙中的視線（aperture vision）、環繞的視線（ambient vision）和可流動的視線（ambulatory vision）。Gibson 認為早期實驗室對視覺的研究屬於快

---

<sup>13</sup> 著名的視覺感知研究學者 Ian E. Gordon 稱 Gibson 的視覺感知與空間感為新典範。

照視線。也就是將眼睛固定於一點上，並將之暴露於在固定點附近的刺激，以得到實驗結果。如果暴露的時間較長且連續，則為縫隙中的視線。另外一種則是戴上頭套以避免受試者自由的走動或轉動頭部，這樣的視線感知雖然不被固著於一點，但仍屬於缺乏空間感的視覺感知。最後，可流動的視線則是最趨近於日常生活的視覺感知，頭、身體皆可任意的移動，感知周圍的空間。

爲什麼 Gibson 視覺感知與空間感、移動感的理論可以被拿來使用在 Webcam 的討論呢？由於 Webcam 所有的特殊機緣，使的人們的視覺感知不同於一般的日常生活，更明確的來說，Webcam 就如同縫隙中的視線以及環繞的視線，人們只能由對方所允許配置的 Webcam 小孔中看事物，同樣的，人們透過 Webcam 的中介下，不管如何轉動頭、身體，也不能如日常生活般的感知空間感。因此，在這樣特殊的空間配置之下，人們的視覺感知會如何的變化？而這特殊的視覺感知又如何影響了人們的互動以及彼此的意義建構呢？這是本文嘗試要連結的部份。

## (二) Norman 文化機緣以及 Goodwin 的視覺符號面

在前面 Gibson 對機緣的解釋中，他強調不僅僅和環境的物體特性有關，而且是相對於生物本身，指的是相對於生物的特性而非生物的經驗（Affordances are properties of things taken with reference to an observer but not properties of the experiences of the behavior）。然而，從郵筒的例子來看，很難用來解釋人造物的機緣卻不涉及經驗的部份。

因此，Norman 建議以設計者的角度出發，認為設計者應該考量約定成俗的文化、習俗等等來邀請（afford）使用者，而這也正是本文所指涉的文化機緣面向討論。Norman 也舉出了何謂好的設計與壞的設計，以及壞的設計對使用者所造成的影響，並建議以使用者爲中心的設計（user-centered design）。甚至在涉及經驗的部份，Norman 更出版了「Emotional design」，將設計的層面分爲三個層次：出自內心的（visceral），行爲的（behavioral）和反思的（reflective），試圖以這三種設計的層次進入使用者的生活脈絡經驗中。



從上述機緣理論的發展來看，人造物可以說既有生物面的機緣，也有文化面的機緣。生物面向的機緣可以從空間感觀察出來，然而，文化面的機緣呢？在 Webcam 的中介中，設計者如何呈現對方和自己的影像就是本文欲集中觀察的焦點。然而，設計者所呈現的對方和自己影像的呈現卻又必須考慮媒介空間的特性，媒介空間為一實體和非實體空間交織而成的空間，設計者所設計的影像呈現雖然屬於非實體空間的影像呈現，然而，實體空間（包含兩人互動，攝影機的配置等）卻又不斷反饋到非實體空間的影像呈現之中。因此，要探討 Webcam 文化機緣面向的討論，還必須考慮到實體空間依情境脈絡而產生的對話及所產生的影像。

因此，文化機緣面向的討論必須同時兼顧實體和非實體空間中的影像互動。而 Goodwin 就曾對影像互動 (image in interaction) 作過精緻的分析。Goodwin(2001) 提及互動中的影像呈現 (image in interaction)，包含了在活動中不斷變化、部署甚至彼此影響的各種符號場域 (multiple semiotic fields)，而這各種符號場域指的就是在不同媒介上呈現的符號 (signs-in-their-media)。包括科學家在互動過程中所使用的圖表，法官在質問犯人所使用的錄影帶證據等，都是不同符號場域的例子。而這些不同場域中的影像互動，面對面的、圖表的或影像的，皆在整個活動過程中不停交織，並生產相關的對話及傳播意義。透過 Webcam 中介的影像互動相當複雜。然而，一方面我們可以將 Webcam 中介的各種影像互動，不管是實體或非實體的影像，皆視為是一種符號資源，而不同的即為這些符號資源在不同媒介上所呈現的方式。從 Goodwin 的角度來看，實體或非實體中的影像，皆屬於符號資源，而影像的特性則因著在不同媒介上的呈現而有所不同。因此，當兩人互動影像的呈現不是透過眼睛的感知媒介，而是透過設計者的使用介面而呈現時，兩人的互動、對話和語言，甚至凝視的規則又會有何改變呢？這將是分析的重點。

### (三) Kress & van Leeuwen 影像設計的系統功能語言學文法

Kress & van Leeuwen(1996)提出將影像以系統功能語言學的語法分析，認為影像如同語言般具有下列三種功能 (Halliday, 1975)：

1. 語言的認知功能 (ideational function)，語言在此的功能可再現世界。例如：「我吃了蘋果」和「蘋果被我吃了」就代表著兩種不同的認知功能，對前者而言，主詞為言者，而對後者而言，主詞為蘋果，因此代表了兩種不同認知世界的角度。同樣的，影像中如何敘事，其手法為何？也可以說明設計者認知世界的角度。
2. 語言的人際關係功能 (interpersonal function)，從語言中可以看出言者和聽者的關係，而在影像中則可以看出設計者和預設使用者之間的關係。例如：「拿茶杯給我」和「請拿茶杯給我」，這兩種言說方式及代表著兩種不同的人際關係。同樣的，影像拍攝的角度，呈現的方式也都隱含了不同人際關係的假設。
3. 語言的文本功能 (textual function)，指的是語言和外在社會結構之間的關係。語言如何組織，排列皆會影響語言的意義。例如：外籍新娘的新聞中，若旁邊加上的字幕是「深情守候丈夫」或是「狠心拋妻棄子」這兩種組合，就會對閱聽者的觀感有著不同的影響。同樣的，影像的排列組合也會造成同樣的效果。而 Kress & van Leeuwen 即細緻的分析影像排列組合背後所代表的意義。

而 Webcam 的文化機緣面向，即是由設計者的角度切入，檢視設計的文法，並從中觀察何為約定成俗的規則，而這些約定成俗或稱為以使用者為中心的設計，又如何影響了以 Webcam 中介雙方的互動呢？由於 Webcam 上的敘事較為瑣碎，因此本文的分析將著重在人際關係功能和文本功能，以檢視 Webcam 不同攝影的角度以及影像呈現背後所代表的意義。以下簡略整理 Kress & van Leeuwen 所提出的影像文法：

### 表 2 影像中的人際關係功能分析

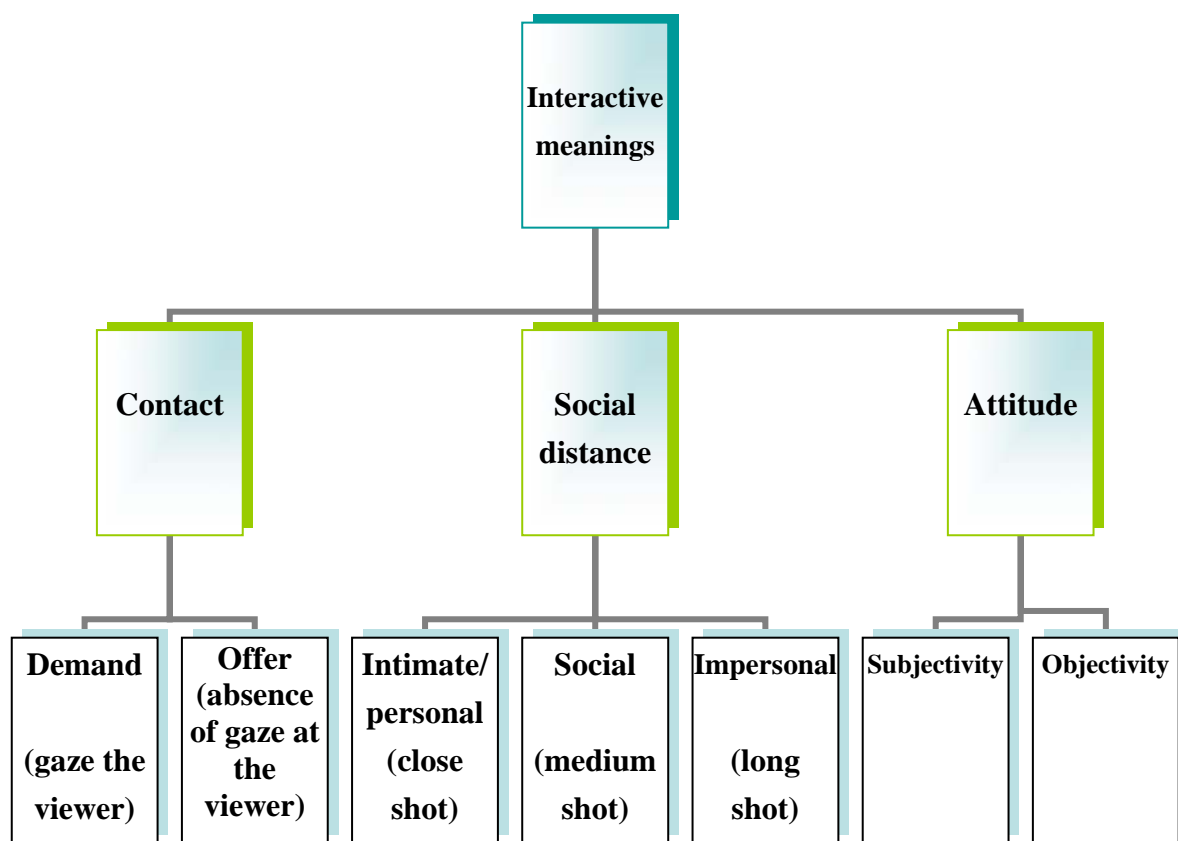


表 3 延續上表

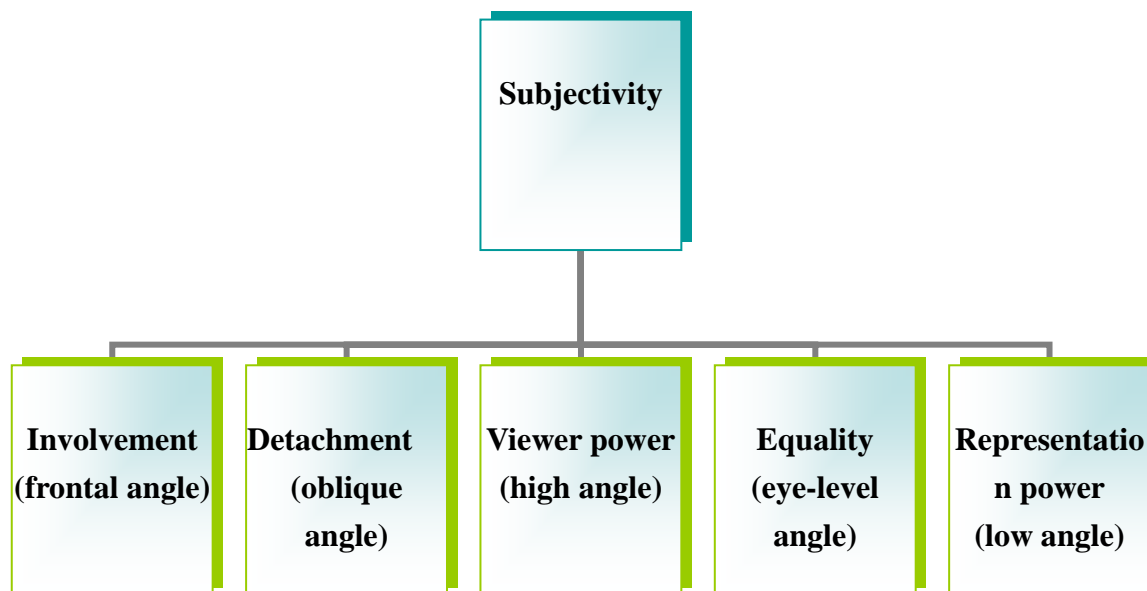
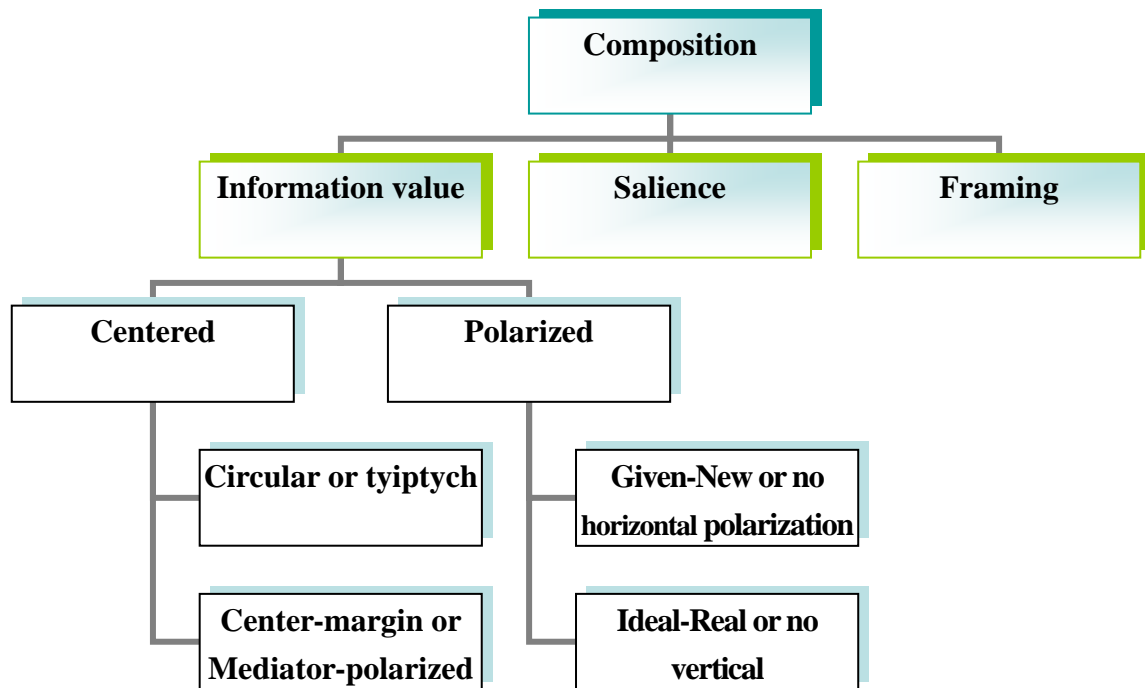


表 4 分析影像中的文本功能



在影像的人際關係功能當中，Kress & Leeuwen 將其分為接觸（contact）、社會距離（social distance）和態度（attitude）三大面向。而在影像中的呈現則是：接觸面又區分為命令和給予，而這是以「凝視聽者」的有無來判斷。當設計者或互動者所呈現的影像是「凝視對方」或「缺乏凝視」，很有可能互動雙方有著命令以及給予的人際角色功能。另外，在社會距離的呈現上，又區分為親密的（personal），社交的（social）以及距離疏遠的（impersonal）。在影像上則是以近拍或遠拍來呈現。當鏡頭要表現出特寫鏡頭，那麼影像設計者即試圖拉近觀者和表演者之間的距離，同樣的，當 Webcam 互動雙方所呈現出的特寫為何，也可以意味著雙方的親疏遠近關係。

最後一個態度的部份，又分為主觀和客觀。在主觀方面，又可區分為涉入與否以及權力的面向。如果，一個鏡頭欲表現涉入的情感，則會表現出正面的角度；反之，將會表現出傾斜的角度。另外，從角度的高低也可以表現出言者和聽者之間的權力關係。高角度代表著觀者的權力高於另一方；反之，則意味著權力較低。

在影像的文本功能中，主要區分為三大項：資訊價值（information value），顯著性（salience）

和框架 (framing)。首先來看資訊價值的部份，這裡指的是影像排列：上下以及左右，甚至是圓圈的排列都預設了設計者的價值觀。當影像擺在左邊，和右邊比起來，左邊的影像代表著較為穩定、抽象以及不可變動的結構，右邊則代表可爭議的焦點。例如在電視新聞的訪談中我們常常可以看見主持人坐在左邊，而被訪問者坐在右邊，這之中也有相同的邏輯在背後運行。同樣的，圖像在上或在下也代表了不同的意義。上面的圖像通常是較為鉅結構的描述，而下方的圖像則為複雜的描述。例如在一般的電視新聞中，可以發現新聞跑馬燈通常至於下方，較為瑣碎細節的資訊並不會擺置在上方。

簡單的回顧 Kress & Leeuwen 關於影像設計語法的探討，本文希望藉由 Webcam 介面設計面向來看 Webcam 有著什麼樣的文化機緣？在其中約定成俗的文化面向有哪些？而這些面向又如何邀請了使用者，並影響了使用者之間的互動呢？

#### (四) 實踐語言

面對面的傳播中 (FTF)，我們通常都會將口語 (spoken language) 視為最重要的溝通工具，而其他非口語傳播元素 (例如：手勢、凝視) 以及週遭的環境、物等視為脈絡線索 (context)，而非共同建造意義的傳播工具。然而，實踐語言的重點卻是：意義是由各個元素如同交響樂團般互相演奏、配合交織而成，並沒有何種元素比哪種元素更為重要。

Hutchins & Palen (1997) 即藉著觀察波音七四七駕駛模擬機艙中機師們的談話，探討空間、手勢和語言中的關係。這觀察之中，Hutchins & Palen 發現，在機艙狹小的空間中，由於機師們無法自由的轉動頭和身體以表達意義，因此，他們特殊的手勢在這空間中就成為最重要的非口語傳播元素之一。而在一個漏油的情境之中，第二機師指向油槽顯示板並將手指向上，只說了：「現在」(right now)。然而，對於外行人的我們是無法了解這手勢所代表的意義，若光聽到 right now，是無法拼湊出完整的意義。唯有了解這項手勢在這活動中所代表的意涵，才能完整的建造出意義。同樣重要的，光是了解了手勢在這項活動中的意涵仍是不夠的，它必須和週遭的空間和物發生關係，

我們才能一層一層的建築出它的意義來。正如在這個例子中，手勢若不在機艙的情境中和油槽顯示板發生關聯，那麼意義就會有所不同。

再回到本文中對 Webcam 的討論。Webcam 是一個特殊的媒介空間，不僅人和人產生互動，人和機器之間也不時的產生互動。在這樣的空間之中，語言不是單一重要的元素，其他非口語傳播元素也不斷的和物、人交織而產生意義。Webcam 中最顯著改變的即是人們的視覺線索，在這樣的空間中，語言不再視唯一主宰的符號資源 (mode)，而是人們身體和物，視覺和 Webcam 等體現的互動 (embodied interaction)。因此，本文所探討的媒介空間、凝視和語言之間的關係，可以放在更大的脈絡之下，也就是人機互動中的體現互動。然而，囿於篇幅和能力，本文挑選了在媒介空間中改變最大，影響最深遠的「凝視」(gaze) 來討論，並進一步希望將這項非口語傳播元素回饋到人機互動的設計上。

### 三、資料分析

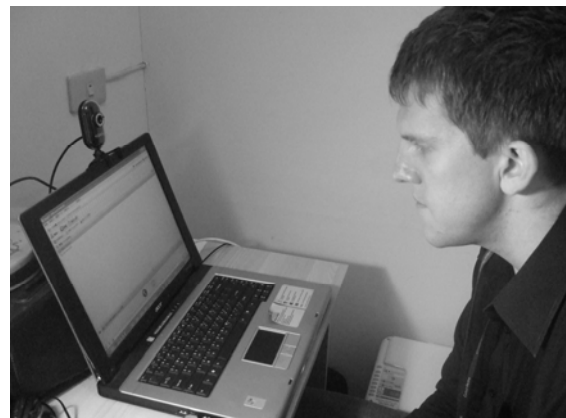
#### (一) 生物機緣面分析

##### 1. 空間配置分析

圖 1：Webcam 和電腦分離的裝置



圖 2：Webcam 裝置在電腦上



上圖顯示了兩種不同的 Webcam 空間配置，而這兩種不同的擺設方式，也間接影響了雙方的互動。拿 S 和 E 為例：S 所使用的為左圖的配置，E 所使用的為右圖的配置。以下為 S 和 E 的對話過錄：

## 2. 因空間配置不同而產生的對話

Turn	Speaker	
1	E	你可以彼此看彼此嗎？做得到嗎？
2	S	...彼此看彼此，什麼意思？這樣喔。
NV1	S	[ 將頭轉向 Webcam，螢幕上顯示 S 正在凝視 (gaze)的效果。 ]
NV2	E	[ 笑 ]
3	E	就是你同時看著我，我同時看著你。
4	S	我只能看自己的鏡頭，因為我的鏡頭和你的又不是平行的。
5	E	這樣彼此互相看好像有點困難喔。
NV3	E	[ 眼珠上下轉動，以同時看鏡頭和螢幕。 ]
6	S	除非你的鏡頭是在正上方。
7	E	對呀，我是啊，你呢？
8	S	我不是啊。
9	E	為什麼？你的 pc 不可以，對不對？
10	S	因為我的夾子夾在旁邊。
NV4	S	[ 調整 Webcam 的位置，使其於 PC 的正上方。 ]
NV5	E	[ 張著嘴巴看螢幕的移動情形。 ] 這樣很怪耶。
11	E	你把它放在上面了。
12	S	對啊。
NV6	S	[ 又調整了一下 Webcam 的位置 ]

由 S 和 E 的對話中可以發現，E 由於使用右圖的裝置，因此使的 Gaze 的動作較為方便，在 NV3 中 E 眼珠上下轉動，以同時看鏡頭和螢幕，正是因為 Webcam 在電腦螢幕的上方，因此 E 可以藉著眼珠的轉動卻不需身體的移動來達成 Gaze 的目標。然而，S 執行 Gaze 的動作有所困難，因為 Webcam 在電腦的旁邊，S 必須移動身體將 Webcam 放置在電腦的正上方，才能比較順利的看著螢幕和鏡頭兩方。

## 3. 缺乏空間感的視覺線索

圖 3: 透過小孔呈現的影像

圖 4: Skype 中的影像



Skype 影像中的呈現是透過 Webcam 小孔所呈現，而這小孔卻是被固定住的，無法跟隨著人的移動而改變方向。



因此 Skype 中所呈現的影像多半呈現出使用者的頭部和頸部，無法呈現出其他背景線索。

由上面兩圖，可以發現到透過 Webcam 所中介的視覺線索，不同於面對面時的互動。當我們是面對面互動時，我們可以藉著頭的轉動、身體的移動來了解整個空間線索，以了解當時對話脈絡。但是當我們使用 Webcam 對話時，我們的視覺線索依賴的是由對方 Webcam 固定小孔中所呈現的影像。也就是對方的 Webcam 小孔變成了我們的眼睛，替我們決定了該看哪些東西，而這被決定該看的東西也不能藉著移動來幫助我們的理解。

#### 4. 缺乏空間感的視覺線索對話

Turn	Speaker	
1	E	你可以 Show 一下你那個視訊的位置到底在哪裡嗎？
2	S	什麼 Show 一下？
3	E	就是給我看一下你的機器 webcam 是擺在哪裡。
4	S	... 這裡啊。
NV1	S	[ 手指頭指向 webcam，在螢幕上顯示出手指頭指向 E 的效果 ]。
5	E	喔，你指向我，我怎麼知道？ <sup>14</sup>
6	S	我在用 Webcam，怎麼讓你知道 Webcam 在哪裡啊？
7	E	喔...那算了。...可我看不到，我只能看到你的頭而已，不能看到你在哪裡，你的背景。
8	S	背景。...
NV2	S	[ 調整 Webcam 角度，使其指向房間。 ] 這是床...窗簾...。

<sup>14</sup> 請參照附錄中的 Webcam 擷取三。

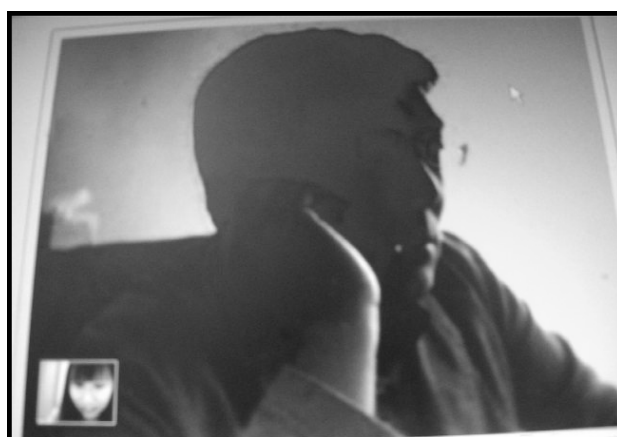


從上述E的要求中，可以看出E渴望得出一些環境的線索。正如Gibson所說的，眼睛位於頭中，而頭又在地面支持的身體上。<sup>15</sup>我們人對視覺的渴望不是只有點的，而是整個面。藉著頭的轉動、眼珠的移動等身體的調配，來感知整個環境。很可惜的，在這個SKYPE環境中，就算E運用了所有感知能力，仍無法理解整個環境。因為她所能看的就是透過S的Webcam中所能呈現出來的影像。

有趣的是，當S試著解決這個難題時（告訴E他的Webcam位置），竟然不是說Webcam在我的某某方，而是以手勢（gesture）直接指向Webcam，而這樣平常在FTF中可以被理解的手勢，卻在SKYPE中呈現一種詭異的形式。因此，這就是一種不對稱性。手勢的理解在SKYPE環境中，還需要多方位的思考。<sup>16</sup>

最後，embodied interaction 著重在人們「not only act on technology, but how to act through technology」<sup>17</sup>的部份，也就是，人們在這其中所表現出來的創意。S反而拿起Webcam當作是錄影機的功能，在房間四處錄製，而這影像反而取代了手勢，能夠幫助E了解環境線索。

## （二）文化機緣面分析



### 1.分析Skype中的影像

從上面的圖表來分析左邊 Skype 最典型出現的影像。先從人際互動關係來看：

圖 5:Skype 中的影像

<sup>15</sup> 原文是 the eyes in the head on a body supported by the ground

<sup>16</sup> Heath, Luff& Sellen(1997).Reconfiguring Media Space: Supporting Collaborative Work. In Finn,Sellen& Wilbur(Eds.), Video-Mediated Communication(pp335-346).Lawrence erlbaum asociaties.

<sup>17</sup> Paul Dourish(2004).Where the action is :the foundations of embodied interaction.Cambridge,Massachusetts.pp154

### (1) Demand/offer.

在這張圖中可以看到 E 的臉屬於正面角度，並隱約凝視 S。然而，S 的臉為斜角度，和 E 並沒有凝視的關係。根據 Kress & Leween 的歸類，此時的對話可能有 E Demand 而 S Offer 的關係。

### (2) Social distance

Skype 中所呈現的影像，大部分屬於頭及肩膀的呈現。細看此影像中，E 只有呈現頭部，而 S 呈現頭部及肩膀，根據上表可以發現 E 所呈現的社會距離可能 S 來的親近。

### (3) Attitude

首先是看角度，前面有提到 E 的角度屬於正臉的角度，而 S 屬於傾斜的角度，因此 E 屬於 Involvement，而 S 相較下則屬於 Detachment。另外，從角度與權力關係來看，E 和 S 所呈現的角度和眼睛同高，顯示較為平等的關係。

另外，從整個版面配置來看，可以發現自己的影像(E)會配置在左下方，而對方(S)的影像會配置在右方。而且佔了顯目的大片。從 Given-New 來看，左邊通常代表已知，不須重視的配置，右邊代表需要重視或者是需要討論的議題。因此，Skype 的影像配置也有相似的邏輯，將自己視為不重要的，而將對方視為焦點。另外，從 Ideal-Real 的配置來看，自己的影像在下方，對方的影像在上方，說明了自己屬於較細節的部份，也就是其實我們的注意力會放在自己瑣碎的表情、動作，而傾向將對方視為背景。另一個值得重視的地方是 Salience，這裡指的式版面配置的比例。Skype 中可以將對方調整至全螢幕，然而，卻無法將自己調整至全螢幕。從這樣的比例關係看來，Skype 的確預設了對方比自己重要的設計邏輯。

## 2. 由CA看文化面的機緣

### (1) 因影像配置而產生的對話

Turn Speaker

NV1 兩人 [正在調整網路攝影機的角度及對照螢幕]

- 1 E 我的那麼小格喔？<sup>18</sup>
- 2 S 自己的當然小格啊。
- 3 E 我的可不可以變大一點。
- 4 S ...恩，不行吧。
- 5 E 我看看。
- NV2 E [8秒] 在桌上移動滑鼠，螢幕上的游標在視訊的位置游移。
- 6 E 視訊...視訊設定...影像功能...噢，為什麼不行？我的怎麼那麼小格啊？
- 7 S == 自己的本來就小格啊。
- 8 E 等我一下啊。
- NV3 E [8秒] 站起來將門關起來，人從視訊中消失，但 E 又保持傾斜的姿勢，以便自己還保持一半入鏡的影像。
- 9 S 你把那個視訊的螢幕拉大。
- 10 E 什麼？[笑]
- 11 S 你把那個 SKYPE 的畫面拉框拉大，讓那個視訊可以是全螢幕。有比較大一點嗎？
- 12 E ...是你大，但是我不大啊。

## (2) Offer/Demand的關係

從上述影像分析中，可以大略理解 E 主要扮演 Demand 者的角色，而 S 主要扮演 Offer 的角色。以下是 S 和 E 對話中的 Mood 分析。

表 5 對話分析

和上圖相符的 1min transcription	S	E
輪番次數(number of turns)	5	7
句子總數(numbers of clauses)	5	8
未完成之句子(incomplete clauses)	0	1
敘述句(declarative)	5	3
完整的(full)	5	2
省略的(elliptical)	0	1
‘為何’開頭的質詢句 (wh-interrogative)	0	5
完整的(full)	0	4
省略的(elliptical)	0	1

<sup>18</sup> 請參照附錄中的Webcam擷取一。

從上述簡單的對話分析看來，E 的輪番次數略比 S 多，且 E 多採取為何開頭的質詢句，顯示出 E 在雙方互動中扮演較主動積極的角色。甚至 E 扮演質詢者的角色，而 S 扮演回應者。這樣的對話和影像中的 Demand& Offer 關係剛好呈現不謀而合的關係。

### (三) 注視規則如何改變—以 CA 取徑並過錄凝視

#### 1. Skype的開場白

Turn	Speaker	
1	S	怎樣。
NV1		[開啟視訊畫面，看著螢幕中的自己，到 Turn4 前，持續八秒]。
2	E	什麼，為什麼不能叫你先開？
3	S	開了啊。
4	E	你現在在看哪裡。
NV2	S	[轉動頭，兩邊擺動，注視著螢幕和 Webcam 的小孔]。
5	E	[笑]，你為什麼頭兩邊轉？
NV3	S	[看著螢幕中的自己整理頭髮，到 Turn6 結束持續 12 秒]。
6	E	你在整理什麼？...好，我看到了。[開啟視訊的程式]。看到我了嗎？
7	S	還沒。
8	E	[4 秒] [調整 Skype 視訊的功能]。看到了嗎？
NV4	S	[4 秒] [認真注視著螢幕]。好糊喔。

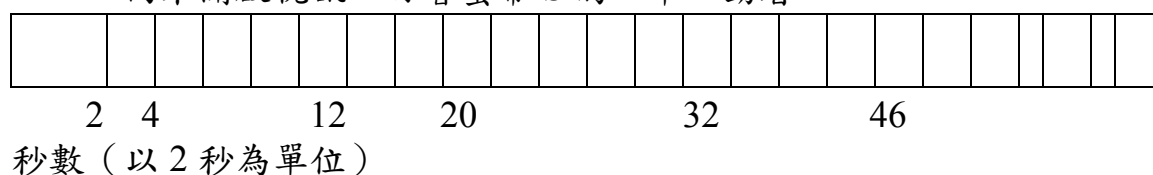
#### (1) Skype開場白的凝視規則

表 6 Skype 開場白之凝視過錄

Shown				
	看著螢幕中的自己	一下看螢幕， 一下看攝影機 www Webcam	對著螢幕整理 頭髮和做表情	認真盯著螢幕看
兩人的螢幕（包含實體和非實體空間）				

## Evelyn

→尚未開啟視訊，盯著螢幕 S 的一舉一動看



從上述簡短約一分鐘的 Skype 開場白，可以窺知人際互動中凝視規則的改變。首先，Skype 視訊中的開與不開，誰先開以及影像如何呈現，就是一個重要的議題。因此，互動雙方花在確認彼此能不能看見彼此的空間上，可以說是比一般面對面互動 (FTF) 情境來的長。當然，在 FTF 中，人們無需花時間在確認對方是否看的見自己，更無須擔心自己影像呈現的問題。在 Turn2 中 E 說了：什麼，為什麼不能叫你先開？充分表現了 Skype 開與不開中微妙的人際權力關係。E 在 46 秒後才開啓視訊，然而，在 46 秒前，E 以凝視之姿和 S 互動，緊盯著 S 的一舉一動，但是 S 卻無從得知 E 的動作。因此，E 的這句話，不僅僅是一句簡單的質詢句，它還承載了語言中的人際功能 (interpersonal function)，命令了 S 開啓視訊以暴露在 E 眼前的一舉一動。從這裡，可以看到語言和凝視之間的關係。語言和凝視在這裡是互補的，沒有誰先誰後的問題。也就是因為這樣的言說，造成了視訊的開與不開，他者的可見與自身的不可見。然而，正因為這樣的不可見，更增加了不可見互動者 E 無形的權威感。

另外，從 S 四次的凝視動作來看，凝視螢幕中的自己佔了凝視互動的絕大部分。凝視可以是凝視螢幕中的自己，螢幕中的對方，或者是螢幕中的其他網頁，亦或是凝視 Webcam 中的小孔 (比較不同的是，唯有此項才能造成雙方凝視的效果，然而，這卻又是最不尋常的凝視方式)。S 在 NV1 到 NV4 中，花在凝視自己身上的時間遠超過凝視對方的時間 (只有在 NV4 中注意著 E 的出現與否中)。從這裡可以看出在媒介空間中，人們改變的凝視規則。在媒介空間中，視訊的開與不開，誰先開以及影像如何呈現中是個重要的議題。然而，視訊開了之後的凝視規則的改變更是另一個值得探究的方向。視訊開了之後，人們傾向於花更多的時間凝視螢幕中再現的自我，而非螢幕中的

他者。反而，視訊沒開的 E 才能專注在凝視 S 的影像。這樣的結論似乎下的太快，也許是因為 E 還沒開視訊之前 S 只能盯著自己看，因此，在下一個例子中，我們將看看 E 和 S 雙方都使用視訊中的凝視規則。

從上述一分鐘的語料中，我們還能觀察到什麼呢？視訊開了之後不僅僅改變了凝視的規則，而這樣的凝視規則有可能也改變了其他規則。在 NV3 中，S 看著螢幕中的自己整理頭髮，到 Turn6 結束共持續了 12 秒。我們可以發現，在 FTF 中如果互動雙方沉默了 12 秒，這會是相當不自在的狀況，更別說只有互動的一方專注在撥弄自己的頭髮，觀看自己的影像了。然而，NV3 中，如此的互動竟然可以持續 12 秒，可以解釋的原因之一是，凝視規則的改變導致沉默時間的拉長，以及表演動作的誇大。當 E 和 S 不對稱的凝視互動時，S 感受到 E 的凝視，卻因為 E 視訊沒開而無法得知 E 的一舉一動。在這樣的情境下，S 的表現變的誇大而具表演性質。S 也許假設 E 正在凝視，因此故意做出誇大的動作來吸引 E 的注意力。也許 S 看不到 E 凝視而忽略了有人正在觀看，因此自顧自的「照起鏡子」來了。

## 2.Skype 互動中的對話（兩方都已開啟視訊）

Turn	Speaker	
1	E	你現在在看電腦螢幕喔？
2	S	...對啊。
3	E	在看你自己的喔。
4	S	恩，蠻帥的。
5	E	蛤？
6	S	對啊。
7	E	你在看你自己喔。
8	S	對啊。
NV1	E	〔五秒〕〔嘆氣〕。
9	S	我覺得你的臉好腫。
10	E	為什麼，你只看的到我鼻子以上耶。
11	S	對啊，可是臉看起來很腫。
12	E	沒辦法，這是角度的問題，你竟然這樣說我真是太過分了。〔調整 webcam 角度，使整個臉入鏡〕。
NV2	S	〔看著 E 的影像笑〕。

### (1) Skype 互動中的凝視過錄

表 7 Skype 互動中的凝視過錄



從此段約 30 秒的語料中，可以看到雙方互動時，觀看自己螢幕中的影像佔了相當大的比例。這呼應了上段 Skype 開場白中的觀察：視訊開了之後，人們傾向於花更多的時間凝視螢幕中再現的自我，而非螢幕中的他者。反而，視訊沒開才能專注在凝視對方的影像。因此，媒介空間裡的凝視規則改變了，人們傾向於凝視自己的影像，不同於 FTF，人們會凝視對方，人們的注意力集中在自己身上大於別人的身上。

另外，從上段的語料中，可以發現 E 在問了「你在看你自己嗎」之後，S 轉將注意力移至螢幕中 E 影像。雖然，這樣的邏輯相同於 FTF，但對於互動另一方的感受卻是截然不同的。也就是，當 FTF 中，我們試圖引起對方的注意力，而對方也回應了他的注意力，將視線放回互動的另一方，然而，弔詭的是，E 完全感受不到 S 重新將注意力放回到她身上。因此，S 只能藉著「我覺得你的臉好腫」這句話來暗示 E，他的注意力回來了。

當然，這樣的話語對於一個女生來說，自然是要反駁的了。但是，有趣的是，E 在這裡的反駁一點尷尬程度也沒有，反而將一切的原因歸咎於攝影機角度的問題。原本在 FTF 中，當一個男生直接向女生說出「你的臉好腫喔」，這是相當不禮貌，甚至可能會造成極為尷尬的情況。但是，在 Webcam 所中介的互動中，如果對方 S 說了這樣的話，E 立即聯想到這是攝影機角度的關係，而不是本身的臉腫。因此，媒介空間中的對話規則也在某種程度上改變了，尤其在關於臉部、動作等的描述增加了，而一切不美或奇怪的影像呈現，又可以歸咎於攝影機的小孔，因此，人們也不會因為奇怪的影像呈現而感到尷尬。

最後的一個部份，是E突如其來的望向Webcam中的小孔，以造成對S凝視的效果。顯然的，E望向Webcam中的小孔，這是比較不尋常的方式，因為E所凝視的小孔並不能看到任何東西，只能在E心中預期使S看的更清楚，使S知道E正在凝視他，上述的動作E完全不能看到，而是發生在E預設的效果中。從這裡可以看出科技與人微妙的互動。新科技也許有許多的限制，這些限制改變了或造成了人們互動的規則、樣貌，但是，人卻可以not only act on technology, but how to act through technology」<sup>19</sup>，在限制中發現創意，創造可能。當E回應該如何才能表現臉比較不腫的樣子，除了調整Webcam的角度，另一個方法就是給S一個正臉的凝視和微笑。而唯有透過Webcam的小孔凝視，才能達到此種效果，雖然，E完全看不到，但是這卻是E心中的預期，也是E對Webcam小孔所做的假設。

### 3. Skype結束的對話

Turn Speaker

- 1 E 這樣就不會醜了，不是嗎？〔凝視著 Webcam 中的小孔以及螢幕中的影像〕。就是比較正常的樣子...〔再次凝視 Webcam 中的小孔〕。角度抓對的話。
- 2 S ...啊〔凝視 webcam 中的小孔〕，沒有眉毛。〔凝視螢幕中的 E〕。

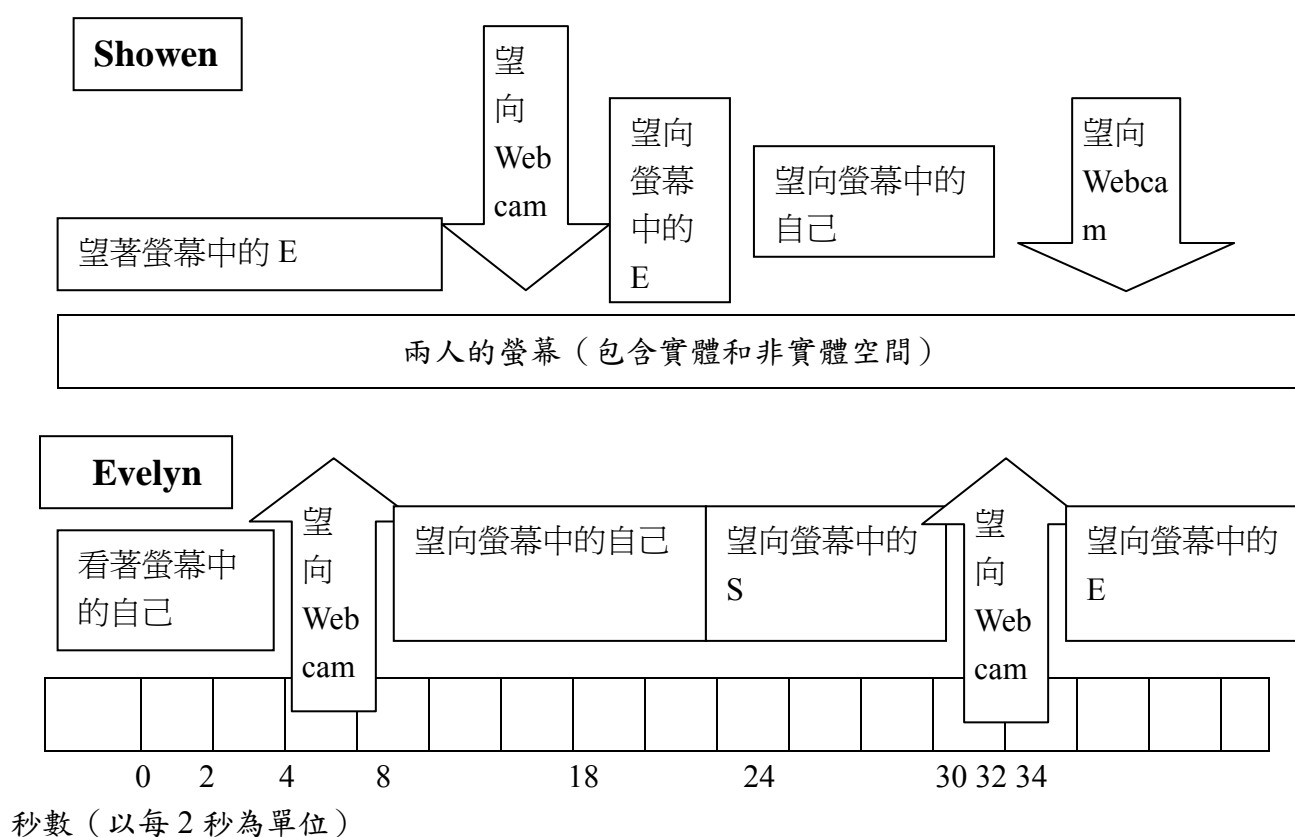
<sup>19</sup> Paul Dourish(2004).Where the action is :the foundations of embodied interaction.Cambridge,Massachusetts.pp154



- 3 E 我哪有沒有眉毛。你想太多。  
 NV1 S [4 秒] [對著螢幕做動作，擺弄頭髮]。  
 4 E 好吧，最後請跟大家說再見。〔凝視 Webcam 中的小孔〕。  
 5 S [轉向 Webcam 中的小孔]，各位同學，台灣的新聞品質就靠你們了。  
 6 E [凝視螢幕中的 S，笑]。好吧，好吧，再見，再見。〔關掉視訊，登出〕。

### (1) Skype 結束的凝視過錄

表 8 Skype 結束的凝視過錄'



在此段約 30 秒的語料中，可以發現互動雙方藉著 Webcam 小孔互動的次數增加。在 E 詢問「這樣就不會醜了，不是嗎」之後，爲了表示希望得到 S 的回應，E 在 6 到 8 秒間望向 Webcam 希望對 S 造成凝視的效果。另一次 E 望向 Webcam 小孔則是在 E 對 S 做要求「好吧，最後請跟大家說再見」的時候。因此，可以發現 E 通常在詢問、要求之後凝視 Webcam，雖然 E 可能沒有意識到這個動作，但是 E 藉著凝視 Webcam 小孔所造成的影像凝視效果，更增強了詢問、要求和

命令的語氣。

不同於 E 的表現，S 並不利用凝視 Webcam 來達成命令、疑問的效果。反而將 Webcam 利用為即時錄影機的功能。當 S 要和 E 道別時，不同於 FTF 中只是看著對方說再見，S 望向 Webcam 並說出再見，是以極為戲劇化的方式呈現。也就是，S 雖然知道望向 Webcam 會對對方造成凝視的效果，但是 S 將他挪用為在攝影前展演的一種工具、概念。他的用意不在於說服、加強和輔助語言，反而藉由 Webcam 中的凝視而充分表現出他的表演慾望。也因此，從 S 和 E 的互動對談中，可以看出兩種截然不同的人和科技的互動。S 將 Webcam 視為可展演的工具，而 E 則將 Webcam 視為輔助語言的工具。

因此，從這段 Skype 結尾的語料中可以窺知凝視規則的改變，而這凝視規則的改變又和前述所說的生物機緣和文化機緣有關。不同的視覺空間會造成不同的機緣，而這不同的機緣卻因著每個使用著文化背景感知的差異又有所不同，甚至設計者本身的設計就隱含著文化機緣的形式，讓使用者不得不跟隨著指示而感知。正如同 E 和 S 之間不得不使用 Webcam 中的小孔以達成凝視的效果，然而，對於 S 來說，Webcam 所賦予的文化機緣使他可以充分表現原本的表演慾望，對 E 而言，Webcam 的功用在於輔助，尤其是強化命令和質詢語氣的表現。因此，媒介空間人們不同的感知亦會造成不同的凝視規則。

#### 四、結論和建議

本文分析的第一部分由 Gibson 取徑強調空間、物配置與凝視之間的關係；另以 Norman 和 Goodwin 的取徑，將影像再現視為符號面，以系統功能語言學分析。試圖處理為何在媒介空間中人們的感知有所不同。第二部份由對話分析取徑過錄凝視的規則，發現凝視的規則有所不同。由第一部份和第二部份，可以得到在媒介空間中，人們因為生物的、文化面向機緣的感知而改變了凝視的規則。最後部份，筆者希望說明凝視在整個意義的建構上，佔了舉足輕重的位置，不僅是次於語言的位置，在媒介空間中甚至凌駕於語言之上。從第一和第二部

份的分析，可以看出許多因為 Webcam 機緣下所產生的對話。因此本文分別分析了因空間配置和影像設計所產生的對話。例如：你可以彼此看彼此嗎？做得到嗎？／我的那麼小格喔？／喔...那算了...可我看不到，我只能看到你的頭而已，不能看到你在哪裡，你的背景。等等。凝視凌駕語言之上，即是實踐語言的意涵，而這不僅呼應了非口語傳播的重要性，更進一步反饋了人機互動設計中不可或缺的元素。

本文屬於初探性研究，因此所擷取的語料主要來自於資料易取得的週遭朋友，而這也正是本文的限制所在。在 CSCW 中，所研究的語料來源主要來自於同事。然而，熟悉的朋友凝視與陌生人的凝視，或者是在視訊上偶遇的陌生人凝視的規則又是什麼？顯然的，會和本文所討論的兩位熟悉朋友間的凝視規則不同。因此，建議未來的方向可以研究與陌生人的凝視規則研究。尤其是目前 Webcam 最常被拿來凝視視訊女郎。

另外一個限制與建議則是不僅僅凝視，還有許多非口語傳播元素可以研究。例如：手勢的研究。手勢機緣及活動的研究也是人機互動中重要的元素之一，例如：手寫 PDA 的設計，手勢就成了當然的主角。總之，回到本文一開始的興趣——人機互動，前文提過人機互動是一種 Embodied interaction，而這之中要如何 act through 則仍有許多創意等待被發掘。

## 參考書目

- Gibson (1986). *The ecological approach to visual perception*. USA: Lawrence Erlbaum Associates, Inc..
- Goodwin(2001). “Practices of seeing Visual analysis: an Ethnomethodological Approach ,” In Leeuwen&Jewitt (eds.) *Handbook of Visual Analysis*. Pp.157-181. London, SAGE.
- Gordon (1989). *Theories of visual perception*. Pp.146. St. Great Britain: Edmundsbury Press.
- Hutchins& Palen (1997). “Constructing Meaning from Space, Gesture, and Speech ,” In Lauren, Roger& Clotilde (eds.) *Discourse, Tools, and Reasoning* Pp.22-39. Germany: Springer.
- Heath, Luff& Sellen (1997). “Reconfiguring Media Space: Supporting Collaborative Work ,” In Finn,Sellen&Wilbur (eds.) *Video-Mediated Communication* Pp.335-346. Lawrence erlbaum asociaties.
- Ian Hutchby (2001). *Conversation and Technology*. England: Cambridge.
- Kress & Leeuwen (2006). *Reading Images. The Grammar of Visual Design*. NY: Routledge.
- Paul Dourish (2004). *Where the action is :the foundations of embodied interaction*. Cambridge, Massachusetts.

# Meaning Construction, Gaze and Language in Media Space: the Case of Skype

Ling-yi Huang

Department of Journalism  
National Chengchi University

## Abstract

What's the relationship between the new communicative technology and the society? Instead of a macro perspective, this article tries to use a micro perspective to analyze human-computer interaction. By the micro analysis, I answer the question of how to design human-computer interface.

This article focuses on a new technology, the Skype's webcam, and uses affordance theory to examine biological and cultural constraints and creativities. The examination focuses on the nonverbal communicative element - gaze. How do people change their visual perception and the rules of interaction? How do people create embodied interaction and new rules?

The data transcribes the conversation and gaze in the use of Skype. The research finds that special media space influences people's visual perception and changes the ways of communication. The human-computer interaction designers should be concerned about these issues.

Finally, this article tries to answer the basic question of communication, i.e., where the meaning is from. Language is no longer the privileged communicative element in new technology. Verbal and nonverbal elements intertwine with each other in the symphony of meanings. No element is more important than another. This article emphasizes the importance of nonverbal communicative element in human-computer interaction interface design.

**Keywords: Human-computer Interaction Interface,  
Affordance Theory, Nonverbal Communication**