

國立臺灣大學文學院語言學研究所

博士論文

Graduate Institute of Linguistics

College of Liberal Arts

National Taiwan University

Doctoral Dissertation



多模態，隱喻，與詮釋的交會：

以政治漫畫與藝術歌曲為佐證

The Encounter among Multimodality, Metaphor, and
Interpretation: Evidence from Political Cartoons and
Poetry Music

林盈妤

Tiffany Ying-Yu Lin

指導教授：江文瑜博士

Advisor: Wen-yu Chiang, Ph.D.

中華民國 104 年 7 月

July 2015





國立臺灣大學博士學位論文
口試委員會審定書

多模態,隱喻,與詮釋的交會:

以政治漫畫與藝術歌曲為佐證

The Encounter among Multimodality, Metaphor,
and Interpretation: Evidence from Political
Cartoons and Poetry Music

本論文係 林盈妤 君 (學號 d98142002) 在國立臺灣大學語言學
研究所完成之博士學位論文,於民國 104 年 7 月 2 日承下列考試委員
審查通過及口試及格,特此證明

口試委員:

江文瑜

(簽名)

劉月仁 (指導教授)

張亭興

呂佳慧

蘇以文

謝辭



從小立志成為大學教授的我，即將取得博士學位的此時，內心無比翻騰，回首來時路，當年那位初踏入語言所時，眼神中閃耀著熱情與夢想的女孩，如今終於戴上畢業帽，撥開金黃的帽穗，嶄新的未來正向我招手，期盼夢想將如最初那般，熱情洋溢地發光發熱。即將揚帆啟程的我，滿心期待，滿懷感恩，在此誠摯地向所有曾幫忙我為我加油打氣的師長親友深深鞠躬，謝謝你們，謝謝那些給我溫暖的正面力量，讓我堅持至今。

感謝我人生的恩師，江文瑜教授，您完全體現了真善美的價值，有您當我的指導教授，是此生最美麗的福分。謝謝您總是鼓勵我要“aim high”為我訂定最高標準，不斷激發我的潛能，讓我勇於挑戰突破框架，謝謝您總是相信我做得到，這份信念始終支持著我，讓我擁有不輕易妥協的決心，期許自己能做到最好。感謝永遠的模範大師，黃宣範老師，謝謝您總是不吝鼓勵我幫忙我，謝謝您讓我體會語言學性感與感性，讓我看見語言學家如何用專業關懷弱勢，對台灣有所貢獻。由衷地感謝臺大語言所曾勉勵我給我諸多啟發的師長，謝謝張顯達老師，安可思老師，蘇以文老師，宋麗梅老師，馮怡蓁老師，呂佳蓉老師，謝舒凱老師，李佳霖老師。感謝我的口試委員，謝謝鄧育仁老師，張榮興老師，蘇以文老師，呂佳蓉老師，謝謝您們提供我許多寶貴的建議，讓我的博論更精進！感謝臺大求學期間曾鼓勵我的好老師，謝謝王寶祥老師，Tim Casey 老師，張玉玲老師，胥嘉陵老師，楊乃冬老師，朱秋而老師，黃銘傑老師，徐維娟老師，邱福棟老師，謝謝您們曾給我的啟發與鼓舞！感謝 UC Berkeley 訪學期間曾協助我照顧我的師長，謝謝 Professor Terry Regier, Professor George Lakoff, Professor Eve Sweetser, Professor Leonard Talmy！感謝 Professor Charles Forceville 給我許多啟發與鼓勵！感謝臺大語言所敬業的助教美玲，嘉蘭姐，白姐，謝謝您們的幫忙！感謝優秀的前輩學長姐曾經給我的幫助，由衷感謝我的恩師學長歐漢斯老師！謝謝盛秀學姊，惠如學姊，季樺學姊，書萍學姊，嘉馥學姊，正賢學長，維倫學長，Pierre 學長，Ben 學長，乃欣學姊，芷誼學姊，Sally 學姊，宜萱學姊，欣怡學姊，人鳳學姊，謝謝你們曾給我的加油打氣！謝謝我的好夥伴書珮，Mike，Simon，Paul，Sarah，妙茹，晉寧，素素，Angela，夜馨，Echo，雅玲，得心，憶如，馨妍，Amber，Emily，Jenny，Perlin，Nate，珮琪，玥彤，佩玥，Taco，江妍，姿妤...你們的溫暖友情是我在所上最珍貴美好的回憶！

感謝臺大統計中心的統計諮詢師嘉鴻，謝謝您提供專業的統計分析建議，協助本博論呈現出重要的研究成果。

感謝最疼我的爸媽，始終對我充滿信心，在我身兼人妻人母與博士生多職忙得不可開交時，總是盡全力支援我，感謝我親愛的弟弟全家福常幫我加油打氣，讓我有最溫暖的娘家當避風港。感謝疼愛我的公婆，在我最需要專心研究寫論文之際，大力幫忙我們全家，讓我安心無後顧之憂地攻讀博士。感謝我的另一半汝憬，謝謝你讓我相信自己，與我並肩面對所有難關與挑戰，謝謝你陪我一起擁抱家庭與夢想，讓我成為更幸福更好的人。感謝寶貝宥維來到我們的生命，從你出生至今，兩年來，哺育你伴你成長的每分每秒，都是我最珍愛的時光，雖然我常常得趁半夜你熟睡時挑燈夜戰，一邊寫論文一邊迎接日出，但知道你甜甜地睡著，媽媽就充滿力量，不再感到孤單辛苦，因為等你起床還要打起精神照顧你呢！謝謝你讓媽咪的身心變得更勇敢堅韌也更強壯，謝謝你陪媽媽一起完成夢想，未來，將有更多豐富的旅程與夢想等著我們去探索去實現呢！

摘要



近年來認知語言學領域中的隱喻相關研究受到高度重視，語言中的隱喻如何型塑並影響我們的思維與行動，在認知科學領域，心理學，數學，藝術，哲學等領域均獲得廣泛的討論與研究，非語言隱喻也隨之逐漸在圖像領域及多模態/多媒體的研究中受到注意。然而，過去的研究大多著重於語言與其他媒介之間一對一的對應關係以及認知，然而本研究認為多模態與語言之間的關係不只有一對一的關係，應可視為多重組合關係的連續體(continuum)，豐富且複雜，因此本研究提出創新的互動式詮釋理論：多模態融合聚變理論(multimodal fusion model)，提供多模態分析以解釋語言與其他媒介之間一對多的對應關係，探討特定主題及抽象概念如何再現於多模態體裁(multimodal genre)中，以政治漫畫與藝術歌曲為例證，深入了解 (1)『美國牛肉進口』這個熱門的政治議題，如何透過眾多的政治漫畫呈現於不同立場的報紙中？(2) 經典詩作『偶然』又如何透過樂曲中的聽覺與視覺，再現於四個版本的藝術歌曲中？ 本論文進一步結合閱聽者訪談結果，審視閱聽者如何詮釋並理解多模態隱喻於特殊體裁。本文進一步提出顯著的多模態特性與多模態視覺/聽覺/語言提示(visual/aural/verbal cues) 均可協助促進多模態隱喻的運作與詮釋。

本研究以多模態融合聚變理論為基本架構，建構了美牛進口相關政治漫畫的多模態語料庫，歷時收集涵蓋台灣兩大報『自由時報』與『聯合報』共 56 篇政治漫畫。本文主張多模態融合聚變結合視覺，語言，與概念三大模式，衍生自轉喻與隱喻(metonymic-metaphoric)共同組成的兩大網絡：相關化轉喻網路(related metonymic network) 和多元化隱喻網路 (diversified metaphoric network)。本研究證明多模態融合聚變是政治漫畫中重要且重複運用的表現技巧，具有將複雜抽象的政治概念有效率地濃縮封裝(encapsulating)的認知功能，並產生諷刺與幽默的效果。本分析進一步強調轉喻的重要性，證明轉喻與隱喻彼此密切交織(interwoven)於多模態融合聚變中，蘊藏於概念情境(conceptual scenarios) 如政治是遊戲(POLITICS IS WAR)與政治是戰爭(POLITICS IS WAR)的隱喻對應範疇內。本研究發現雖然兩報的批判性訊息與相異的政治立場均浮現自多模態融合聚變理論，卻分別透過顯著的視覺特色與語言情境，更加強化與對比。

本研究以四首為徐志摩詩作『偶然』譜寫的藝術歌曲為例證，提出以隱喻為基礎的譜面分析，進而發現此體裁中特殊的現象：『隱喻之鏈』(chain of metaphor)，即歌曲中的多模態隱喻彼此可前後呼應，互相串聯於整首歌曲中，以維持整體的連貫性。另外，本研究發現藝術歌曲中的多模態隱喻可透過轉喻的方式再現(metonymic representation) 於歌曲中（如：模擬雲的視覺形狀於譜面或模擬海浪的節奏於伴奏中）。作者更進一步提出藝術歌曲中的多模態隱喻兩大面向：具體形象(concrete image)與抽象隱喻延伸(abstract metaphorical extension)，

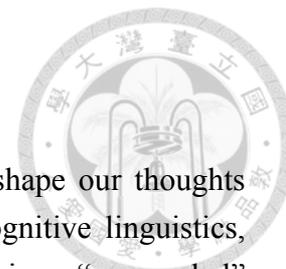
透過分析提出相對應的音樂技巧以及具體抽象光譜(concrete-abstract continuum)為四首歌曲分類。本文並從身體經驗(embodied)為基礎的概念隱喻出發，發掘抽象概念如空間關係和情緒均可對應於音樂中，如透過音量/音域高低，音色明暗等音樂特性呈現。

本研究進一步整合閱聽者反應的量化與質化分析，從閱聽者訪談的研究結果中發現，閱聽者對於政治漫畫的高度理解，來自有效率的人物辨識，背景理解與多模態之間的有效且互動式的結合，而閱聽者對於藝術歌曲的高度理解則來自能分辨藝術歌曲中的具體形象與抽象意涵的詮釋，詮釋過程中所建立的對應關係與閱聽者對於藝術歌曲的喜好度具有正面相關，不同版本的喜好度也間接反應了閱聽者如何感知及詮釋藝術歌曲中語言與聽覺的融合下所產生的音樂效應。

本論文從分析者與閱聽者的視角，探究視覺，聽覺，與概念如何互動,整合,並透過多模態隱喻詮釋，本研究希望拓展以多模態隱喻為主題的跨領域研究，針對特定的多模態體裁深入分析，並結合認知與語用學的觀點，進一步理解多模態體裁中的詮釋與審美經驗歷程。

關鍵字：多模態融合聚變理論、多模態隱喻/轉喻、閱聽者反應、多模態提示、多模態體裁

Abstract



In the past few decades, studies on how verbal metaphors shape our thoughts have been widely discussed in the field of cognitive science, cognitive linguistics, psychology, mathematics, art, and philosophy, etc. Under this view, “non-verbal” metaphors have also received considerable attention, while previous studies mainly focused on the mental processing and one-on-one correspondence between language and other mode, especially the pictorial mode. However, this study believes the relations between multimodalities can be viewed as a complex continuum of various kinds of combinations. Therefore, this study proposes an alternative interactive interpretation model, the Multimodal Fusion Model, to account for the one-on-many relationship between language and multimodality (visual and aural modes) and to explore how certain topic/issue has been represented in various forms in distinct genre. Based on the evidence from the multimodal genre of political cartoons and poetry music, we examined (1) how the hotly debated topic, the U.S. beef import issue, has been represented in various political cartoons in different newspapers and (2) how a classic poem (i.e. Serendipity) has been conceptualized in four versions of songs. We further incorporate the audience response analysis from the interview surveys to examine how the audience actually interprets the multimodal genre. Our findings also show how prominent multimodal features and multimodal cues (visual/aural/verbal cues) help facilitate the interpretation process of multimodal metaphor in specific genre.

This study constructs a multimodal corpus involving 56 political cartoons with regard to U.S. beef import issues as reported in two dominant Taiwanese newspapers, the *Liberty Times* and *United Daily News*. Our corpus shows that multimodal fusion combining the visual mode, verbal mode, and the conceptual level evolves from two metonymic-metaphoric networks, i.e., related metonymic network and diversified metaphoric network. The result demonstrates that multimodal fusion is a significant and recurrent representation technique in the genre of political cartoon and has the cognitive function of encapsulating the abstract complex political debates efficiently with irony and humorous effect. In addition, this study emphasizes the important role of metonymy and demonstrates how metonymies and metaphors are interwoven in the process of multimodal fusion, which underlies the metaphorical mappings of conceptual scenarios related to “POLITICS IS GAME” and “POLITICS IS WAR”. Although the critical messages and distinct political stances in two newspapers both emerged through multimodal fusion, they are highlighted and contrasted through prominent visual features and verbal context in political cartoons.

Based on the evidence from four versions of poetry music composed for the poem “Serendipity” by Hsu Chih-Mo, this study provides a metaphor-based musical analysis under the scope of multimodal fusion model. We discovered the “chain of metaphors” as an interactive and echoing relationship between metaphors that creates a holistic representation of multimodal metaphor combining the verbal and aural mode in the genre of poetry music. Furthermore, the results show the metonymic representations of metaphor in poetry music have been realized through the “pictorial effect” of partial visualization of cloud image on the score representing CLOUD metaphor and aural simulation of wave rhythm representing SEA metaphor, etc. Through cross-comparison analysis, we propose two major aspects of metaphor, the “concrete image” and “abstract metaphorical extension”, have been highlighted differently through musical techniques and features such as dynamics, pitch, and timbre, etc. in four versions of songs. Moreover, this study provides a concrete-abstract continuum as a categorization of composing patterns and musical effects. Viewing from the embodied perspective, our findings also show that the concrete relationship of space and abstract attributes of emotion can be mapped onto the representations of prominent musical features in poetry music.

This study further incorporates the quantitative and qualitative analysis from audience response survey. The results show that higher level of audience interpretation in political cartoons relies on the effective visual recognition of caricatures, comprehension of verbal and contextual information, and interactive connection of the verbal mode, visual mode, and conceptual level. On the other hand, the higher level of audience interpretation in poetry music is related to the distinction between the concrete image and abstract metaphorical extension. The results also suggest that the audience’s preference is relevant to the correspondence between the verbal and aural mode, which implicitly reflects how the audience “feels” about the musical effect and aesthetic functions created from multimodal fusion in different versions of songs.

From the perspectives of researcher and audience, this study investigates how verbal mode, visual mode, aural mode, and conceptual level are interacted, integrated, and interpreted through multimodal fusion. We hope to shed light on the interdisciplinary studies on multimodality and the interface between the cognitive mechanism and pragmatic multimodal use of language, picture, music, and the aesthetic experience of the multimodality appreciation process.

Keywords: Multimodal Fusion Model, multimodal metaphor/metonymy, audience response, multimodal cues, multimodal genre

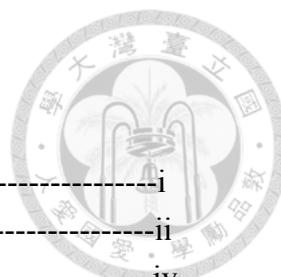


Table of Contents

Signature Page	i
Acknowledgements	ii
Abstract in Chinese	iv
Abstract in English	vi
Table of Contents	1
List of Figures	4
List of Tables	6
Chapter 1 Introduction	7
1.1 Background Information of U.S. Beef Imports	11
1.2 Background Information of Poetry Music: “Serendipity”	14
1.3 Complexity in Multimodality: Multimodal Fusion Model	16
1.4 Multimodal Fusion in Political Cartoons	26
1.4.1 Metonymic-Metaphoric Network and Conceptual Scenarios in Political Cartoon	26
1.4.2 The Conceptual Metaphor and Prominent Visual Features Contrast Functions as Visual Cues	29
1.4.3 Genre: Cognitive-Pragmatic Functions and Critical Effects in Political Cartoons	30
1.4.4 Audience Response to Multimodal Fusion in Political Cartoon..	30
1.5 Multimodal Fusion in Analyzing Poetry Music	32
1.5.1 Metonymic-Metaphoric Network: Metonymic representation of Metaphor and Chain of Metaphor in Poetry Music	33
1.5.2 The Conceptual: Embodied Basis of Conceptual Metaphor..	34
1.5.3 Multimodal Cues in Poetry Music	36
1.5.4 Genre: Musical Effects and Aesthetic Functions	36
1.5.5 Audience Response to Multimodal Fusion in Poetry Music	38
1.6 Organization of the Dissertation	39
Chapter 2 Literature Review and Theoretical Background	42
2.1 Studies on Political Cartoon	42
2.2 Studies on Music and Language	48
2.3 Theoretical Background	53
2.3.1 Conceptual Metaphor Theory and Multimodal Metaphor	54
2.3.2 Conceptual Integration Theory	58
2.3.3 Special Focus on Metonymy and Metaphor	75

2.4 Our Proposal: Multimodal Fusion Model.....	78
Chapter 3 Multimodal Fusion in Political Cartoon: Debates on U.S. Beef Imports into Taiwan	82
3.1 Data Collection: Multimodal Corpus of Political Cartoon.....	82
3.2 Data Analysis	84
3.2.1 Related Metonymic Network: Productive Visual Images.....	87
3.2.2 Diversified Metaphoric Network: Prominent Verbal Cues.....	100
3.3 Conceptual Scenario and Conceptual Metaphor.....	110
3.4 Multimodal Fusion and Political Stance.....	115
3.4.1 Different Targets of Bull Metonymy in the Same Scenario.....	115
3.4.2 The Personified Bull in Different Scenarios.....	119
3.4.3 Different Representations of Uncle Sam in Different Scenarios	123
Chapter 4 Audience Response to Multimodal Fusion in Political Cartoon	128
4.1 Questionnaire Survey: Participants' Response.....	128
4.2 Survey Procedures.....	129
4.3 Results and Discussion.....	132
4.3.1 Significant Differences in the Interpretation Level.....	133
4.3.2 Political Stance Identification	137
4.3.3 The Involvement of Conceptual Scenario	138
4.3.4 The Distribution of the Conceptual Metaphor.....	142
4.3.5 Effective Multimodal Fusion and Interpretation.....	144
Chapter 5 Multimodal Fusion in Poetry Music: The Encounter between Language and Music in “Serendipity”	147
5.1 Data Description of Poetry Music: “Serendipity”.....	147
5.2 Data Analysis	148
5.2.1 Version 1	154
5.2.2 Version 2	159
5.2.3 Version 3	165
5.2.4 Version 4.....	169
5.3 Discussion	171
Chapter 6 Audience Response to Multimodal Fusion in Poetry Music	180
6.1 Questionnaire Survey: Participants' Response.....	180
6.2 Survey Procedures.....	181
6.3 Results and Discussion.....	184
6.3.1 Matching Level of the Concrete Image.....	185
6.3.2 Preference among Four Versions of Songs	190
6.3.3 Multimodal Fusion and Interpretation: Different Focuses on the	

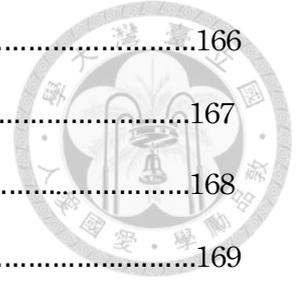
Multimodal and Conceptual level.....	192
Chapter 7 Conclusion	194
7.1 Multimodal Fusion in Political Cartoon.....	196
7.2 Multimodal Fusion in Poetry Music.....	199
7.3 Audience Interpretation and Multimodal Fusion	201
7.4 Concluding Remarks	204
Appendix 1: Word-to-word Translation of the Poem ‘Serendipity’:.....	206
Appendix 2: Political Cartoon Survey Questionnaire	208
Appendix 3: Poetry Music Survey Questionnaire	219
Appendix 4: Statistical Data of Questionnaire Surveys	228
References.....	233



List of Figures

Figure 1. The Conceptual Level	17
Figure 2. The Multimodal Level	18
Figure 3. Other Aspects	19
Figure 4. Multimodal Fusion Model	21
Figure 5. The Conceptual Level	65
Figure 6. The Multimodal Level	66
Figure 7. The Conceptual Level	68
Figure 8. The Multimodal Level	69
Figure 9. The Conceptual Level	73
Figure 10. The Multimodal Level	74
Figure 11. Multimodal Fusion in Political Cartoon.....	85
Figure 12. The Parallel Comparison between Example 3, 4, 7, 8 and 5, 6, 9	135
Figure 13. Multimodal Fusion in Political Cartoon.....	149
Figure 14. Version 1: Paragraph 1.....	156
Figure 15. Version 1: Paragraph 2-Part	157
Figure 16. Version 1: Paragraph 2-Part	158
Figure 17. Version 2: Introduction.....	159
Figure 18. Photo Cloud Image and Pictorial Cloud Image.....	160
Figure 19. Version 2: Paragraph 1.....	161
Figure 20. Version 2: Paragraph 2-Part 1.....	162
Figure 21. Version 2: Paragraph 2-Part	164

Figure 22. Version 3: Paragraph 1.....	166
Figure 23. Version 3: Paragraph 2-Part	167
Figure 24. Version 3: Paragraph 2-Part	168
Figure 25. Version 4: Paragraph 1.....	169
Figure 26. Version 4: Paragraph 2.....	170
Figure 27. The Continuum between Concrete Images and Abstract Metaphorical Extensions	172





List of Tables

Table 1. The Distribution of Political Cartoons in the Multimodal Corpus	83
Table 2. Bull/Animal-related Scenarios of the GAME Domain in LT.....	111
Table 3. Bull/Animal-related Scenarios of the GAME Domain in UDN.....	111
Table 4. War-related Scenarios in LT.....	113
Table 5. War-related Scenarios in UDN.....	113
Table 6. Pairwise Comparison Statistical Data Summary.....	134
Table 7. Positive Response to the Involvement of Conceptual Scenarios.....	139
Table 8. Negative Response to the Involvement of Conceptual Scenarios.....	141
Table 9. Audience Response to the Distribution of Conceptual Metaphor	143
Table 10. Metaphors in “Serendipity”	181
Table 11. The Statistical Data Summary of the Significant Differences in “CLOUD” metaphor	186
Table 12. Participants’ Descriptions of the Concrete Image of CLOUD.....	187
Table 13. The Statistical Data Summary of Preference Score among Four Versions of Songs (Part 2)	190

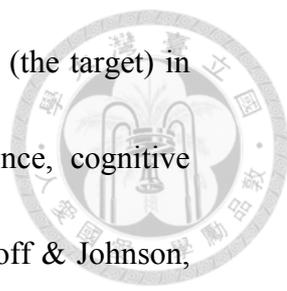
Chapter 1 Introduction



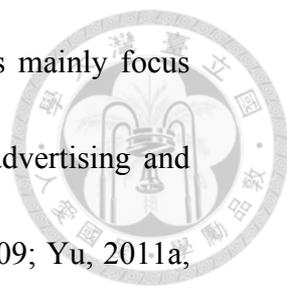
In this study, we focus on the relations and interactions between multimodality, metaphor, and interpretation through the multimodal genre of political cartoon and poetry music with the central concern on the following fundamental questions:

- (1) What are the complex relations and interactions between multimodalities of verbal, visual, aural mode and the conceptual level in multimodal data?
- (2) How are multimodal metonymy and multimodal metaphor represented and interpreted in specific multimodal genre?
- (3) How do the “multimodal cues” and prominent multimodal features help facilitate the interpretation in creating critical effects in specific multimodal genre?

Over the past few decades, there has been mounting evidence of how “verbal” metaphors shape our thoughts and actions, inspired by the tenets of the Conceptual Metaphor Theory (Lakoff and Johnson 1980/2003, 1999; Gibbs, 2008). Previous studies have extended the central claim to explore the



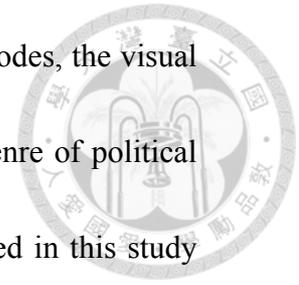
cognitive process of understanding and experiencing one thing (the target) in terms of another (the source) in the field of cognitive science, cognitive linguistics, psychology, mathematics, art, and philosophy (Lakoff & Johnson, 1980; Gibbs, 1994, 1995; Johnson, 1987, 1993; Lakoff, 1993; Lakoff & Johnson, 1999; Lakoff & Nunez, 2000; Lakoff & Turner, 1989; Lin & Chen, 2014; Sweetser, 1990; Teng, 1999; Turner, 1991). Under this view, “non-verbal” metaphors, especially the pictorial/visual metaphors (Forceville, 1996), in the pictorial trope have also received considerable attention. Previous studies have provided theoretical accounts for the pictorial metaphor and pictorial simile in advertising and film (Carroll, 1996; Forceville, 1994, 1996, 2000; Kaplan, 1990; Messaris, 1997; Whittock, 1990). In addition, Cognitive researchers have focused on the process of meaning construction of the visual mode “outside” the language domain in cartoons, political cartoons, visual arts, comics, advertisements, digital design, and film (Cohn, 2007, 2010; Coulson, 2002; Gleason, 2009; Schilperoord & Maes, 2009; Schilperoord, 2013). With a design-based approach, Teng and Sun (2002) further proposed the “image grouping hypothesis” to examine the pictorial grouping, pictorial simile, and pictorial oxymoron in advertisements.



Recent studies viewing from the multimodal perspectives mainly focus on the relations between language and pictorial elements in advertising and films (Forceville, 2007, 2008, 2012a, 2012b; Urios-Aparisi, 2009; Yu, 2011a, 2011b) or between language and gestures (Muller & Cienki, 2009; Mittelbeg & Waugh, 2009; Parrill & Sweetser, 2004). However, scant studies have viewed political cartoons as multimodal entity (instead of “visual” genre alone) and further investigated the readers’ response to political cartoons (El Refaie, 2009a, 2009b). Very few studies (Zbikowski, 1999, 2002, 2009; Hsu & Su, 2014) have examined congruous and incongruous relations between language and music (see Chapter 2). To fill these research gaps, this study aims to extend previous studies by incorporating the perspectives of researcher and audience to investigate the complex relations between multimodalities of verbal, visual, aural mode and the conceptual level of metaphor based on the multimodal genre of political cartoon and poetry music.

In addition, previous studies concerning multimodality mainly focused on the one-on-one relationship between language and other modes. However, this study believes the relations between multimodalities can be viewed as a continuum of various kinds of combinations or competitions. To investigate the fascinating yet complex continuum, this study specifically focuses on the

“one-on-many” relationship between language and the other modes, the visual or aural mode. Based on the evidence from the multimodal genre of political cartoons and poetry music, the central research questions raised in this study are:

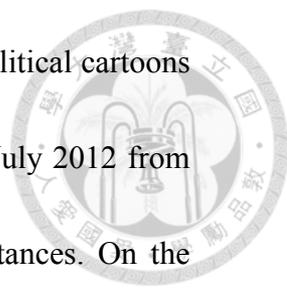


How can a certain topic or concept be represented in various creations of work?

In other words, how can a hotly-debated topic (i.e. U.S. beef imports issue) be represented in various editorial cartoons in different newspapers?

How can a classic poem (i.e. Serendipity) be conceptualized in different versions of songs?

In view of the one-on-many relationship between language and other modes, this study focuses on the data concerning certain topic represented in various forms in the genre of political cartoon and poetry music. With regard to data selection of political cartoon, this study systematically constructs a “multimodal corpus” based on a collection of 56 political cartoons composed of visual and verbal modes concerning the investigated topic of U.S. beef imports as published in two dominant newspapers in Taiwan, namely the *Liberty Times* (LT) and *United Daily News* (UDN), to provide a dynamic picture of political cartoon analysis in terms of a specific issue. Following the



historical timeline of the expansion of U.S. beef imports, the political cartoons in our multimodal corpus are collected from October 2009 to July 2012 from LT and UDN, which are newspapers with distinct political stances. On the other hand, to explore the complex interplay between the conceptual, the verbal mode, and the aural/musical mode, this study focuses on the classic Mandarin Chinese poem “Serendipity” (偶然) written by the famous poet Hsu Chih-Mo (徐志摩), which has been represented and musicalized in four versions of songs and viewed as one of the most popular poetry lyrics in Taiwan music history.

In the following section, the background information of U.S. beef imports and poetry music “Serendipity” will be presented.

1.1 Background Information of U.S. Beef Imports

Looking at the period of 2009 to 2012, we summarized the historical timeline of U.S. beef import issues in Taiwan¹ in the following three stages.

(1) 2009 Protocol: Relaxing of previous ban on U.S. beef products

A huge debate has been raging in Taiwan over the safety of U.S. beef since President Ma Ying-jeou (President Ma) and Washington signed a

¹ In this section, we summarize the main issues from online news reports in *Taiwan News* and *Taipei Times*, the representative English news media in Taiwan, in order to provide an accurate English translation of the issues concerning U.S. beef imports.



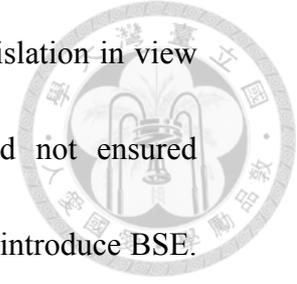
protocol on October 22, 2009 under which Taiwan agreed to lift its ban on U.S. bone-in beef and certain other beef products. The decision sparked controversy in Taiwan, with opposition parties, including the Democratic Progressive Party (DPP)² and civic groups led by the Consumers Foundation and politicians, contending that imports of high-risk beef products could put consumers at risk of Mad Cow Disease (Bovine Spongiform Encephalopathy, BSE). However, the Department of Health (DOH) assured the public that strict measures (i.e., the “three controls” and “five certifications”) had been effective at keeping potentially hazardous U.S. beef out of Taiwan. The Economic Affairs Minister claimed that “the measures have been effective and that public misgivings about the high-risk parts of U.S. beef coming to Taiwan have proven unfounded.”

(2) 2009 Amendment: Banning of U.S. beef imports in contravention of the 2009 protocol

On December 29, 2009, the Legislative Yuan passed an amendment to the Act Governing Food Sanitation that effectively barred U.S. ground beef, bovine offal, and other beef parts such as skulls, eyes, and intestines from access to the Taiwan market, contravening the bilateral beef trade protocol

² Democratic Progressive Party (DPP) is the opposition and the dominant party of the Pan-Green Coalition. DPP is widely classified as the liberal party in Taiwan due to its strong advocacy of human rights, Taiwanese identity, and the independent status of Taiwan.

signed in October, 2009. Taiwanese lawmakers passed the legislation in view of widespread concern in Taiwan that health officials had not ensured sufficient safeguards to prevent the entry of imports that could introduce BSE.



Following this amendment, President Ma stressed that the new beef agreement signed in October remained effective and that the government would seek to reopen talks with Washington on issues related to beef imports, while the controversy remains due to the consideration of public health and international regulations.

(3) 2012 Amendment: New regulations that relax restrictions on U.S. beef imports

Earlier in 2012, the Taiwanese government intensified its inspections of imported beef, resulting in the destruction or return of large quantities of beef containing ractopamine residue and subsequent dramatic reductions in U.S. beef imports. However, in March 2012, under pressure from the United States, the government announced that it would relax these restrictions. The legislature passed an amendment to the Act Governing Food Sanitation in late July 2012, opening the way for the import of U.S. beef containing ractopamine to Taiwan by setting the maximum allowable residue level of the drug at 10 parts per billion for beef while maintaining a complete ban on the drug for

imported pork. On September 11, 2012, a new regulation allowing imported beef containing set levels of ractopamine took effect.



1.2 Background Information of Poetry Music: “Serendipity”

We believe “poetry music” should be considered a unique genre due to its development under the background of Taiwan music history. Since the 1970s’ movement on “marriage between contemporary poetry and music” that promotes “musical” edition of poetry, which is advocated by poets in Taiwan, the combination of poetry and music has become a trend, forming the fad of “poetry music” in Taiwan. The composition of poetry music is unique in its combination of western and Taiwan music elements. The genre of poetry music emphasizes the intimate relationship between literature and music and also the fusion among traditional folk music, nostalgia emotion, and the modern style of western music grammar.

In particular, Hsu Chih-Mo’s poetry “Serendipity” has been viewed as one of the most popular poetry lyrics composed in four versions of songs, keeping a record throughout the poetry music history in Taiwan. The following is the original Chinese version of “Serendipity” and our English translation.³

³ Please see Appendix for word-to-word English translation of the poem.



我是天空裡的一片雲，

I am a cloud in the sky,

偶然投影在你的波心，

by chance reflected on the wave of your heart.

你不必訝異，

Don't be surprised,

無需歡喜，

Or too elated;

在轉瞬間消滅了蹤影，

In an instant I shall vanish without trace.

你我相逢在黑夜的海上，

We meet on the sea of dark night,

你有你的，我有我的方向，

You are on your way, I am on mine.

你記得也好，

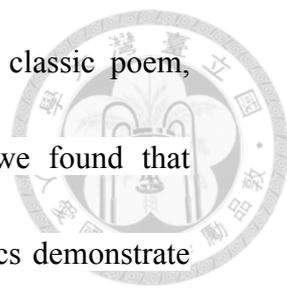
Remember if you will,

最好你忘掉，

Or, better still, forget

在這交會時互放的光亮

The glowing light exchanged in this encounter.



In total, four musical versions have been composed for the classic poem, “Serendipity”. Among a wide a variety of poetry music, we found that different versions of songs composed for the same poetry lyrics demonstrate how the conceptual level (the concept of space, emotion, etc.) and the verbal mode (poetry lyrics) can be alternatively combined with the aural/musical mode in the compositions, creating different musical effects and aesthetic feelings.

Based on the unique characteristics of the data in this study, we will provide detailed multimodal analysis of our multimodal corpus of political cartoon and examine all of the four versions of poetry music “Serendipity” in the following sections.

1.3 Complexity in Multimodality: Multimodal Fusion Model

To examine the relationship between language and other modalities, previous studies demonstrate that Conceptual Blending Theory (Fauconnier & Turner, 1996, 2002) is applicable in analyzing the corresponding relationship. According to blending theory, the conceptual integration incorporates three types of spaces, the *blended* spaces integrates selected information and structure from two or several discrete domains, which are called the input spaces, while the *generic* spaces contains the structural aspects common to the

input spaces. The blended spaces and the input spaces are connected by projection relations, whereas the generic spaces and the input spaces are connected by inheritance relations, and the counterpart connections between entities and relations captured by the inputs are called *vital relations* (cf. Fauconnier & Turner, 2002, pp. 47, 89). Previous studies have viewed Conceptual Blending Theory as the useful approach to account for the relation between (1) different concepts (the conceptual level) or (2) different modalities, such as verbal and visual modes (the multimodal level) or (3) other aspects as summarized in the following Figure 1, 2, 3 (For further details, please see literature review in Chapter 2).

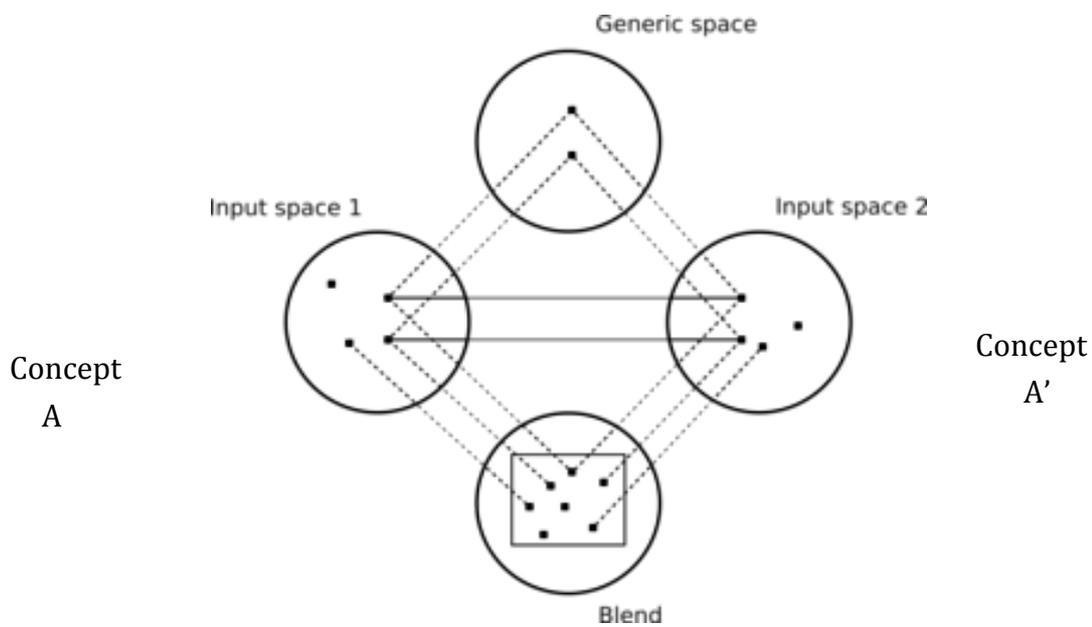
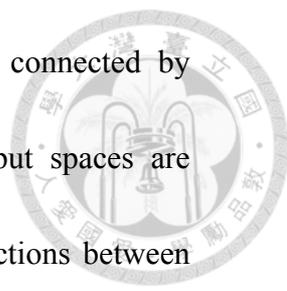
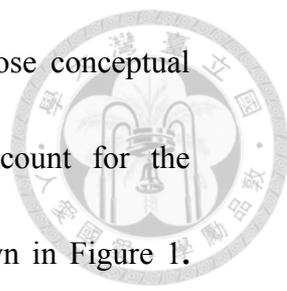


Figure 1. The Conceptual Level



Most previous studies focused on the linguistic data to propose conceptual integration networks that combine different concept to account for the online-processing mechanism on the conceptual level as shown in Figure 1.

Previous studies also adopted the Conceptual Integration Networks (CIN) to account for the corresponding relation between language and other modalities on the multimodal level, such as language (verbal) and picture (visual mode), or language (verbal) and music (aural) mode, as summarized in Figure 2.

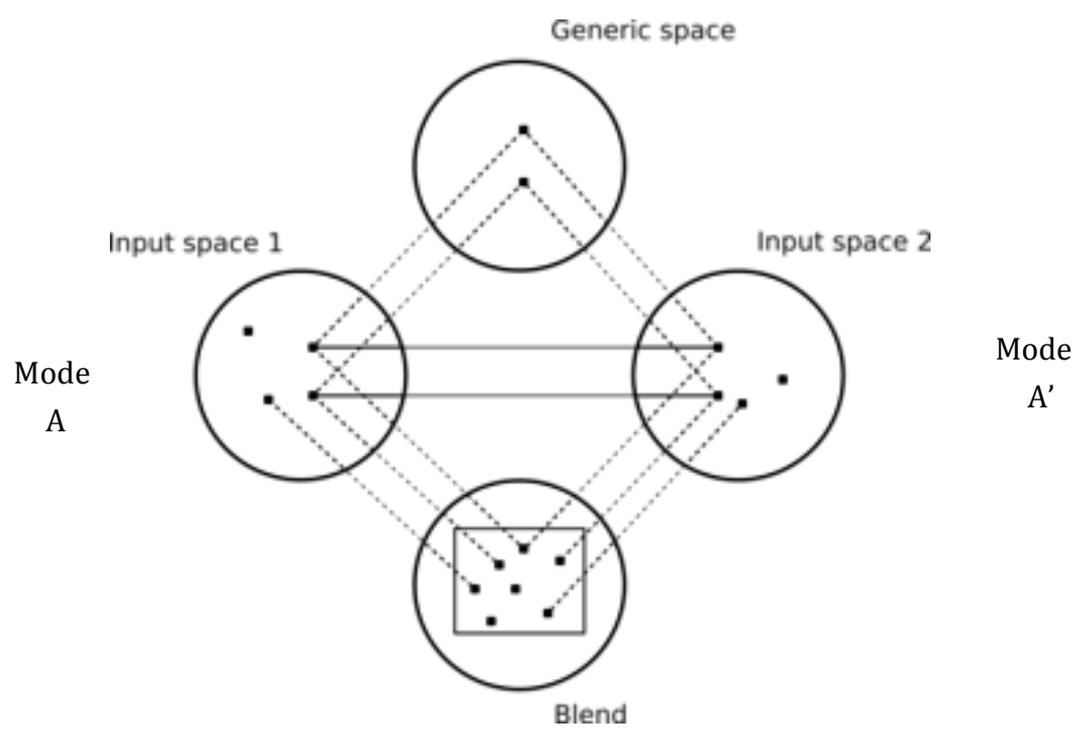


Figure 2. The Multimodal Level

Previous studies also indicate that conceptual blending network is also powerful in explaining how different aspects can be combined and blended to create emergent meanings in many other fields as summarized in Figure 3.

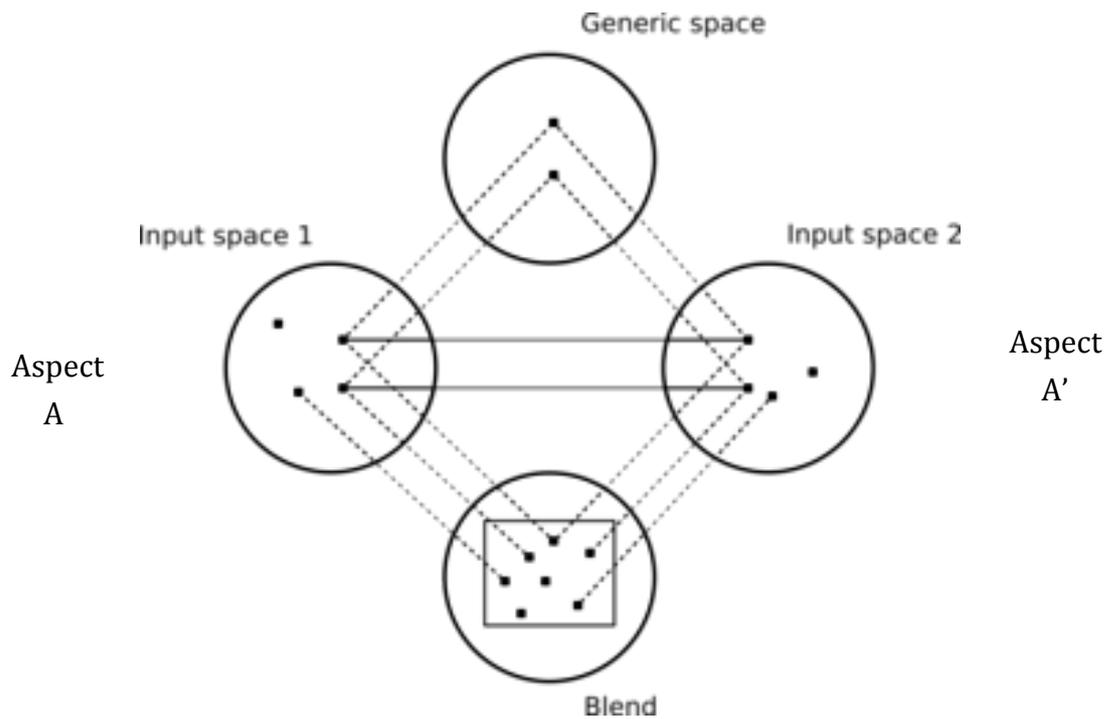
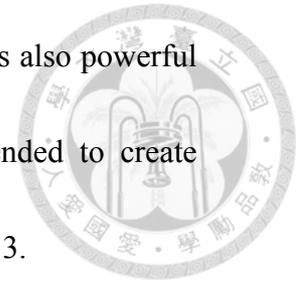


Figure 3. Other Aspects

Thus various Conceptual Integration Networks (CIN) can be constructed for different focuses or purposes (For more details, see Chapter 2). However, we wonder if there is an analytical approach to incorporate the conceptual level and modalities in the multimodal analysis since the conceptual level, the

verbal, visual, and aural modes are all encapsulated and integrated simultaneously in the multimodal discourse, such as in the genre of political cartoons and poetry music.



Through blending theory, we can see the general patterns of how input spaces can be blended to create the emergent structure and emergent meanings in each piece of data, for example in a political cartoon or in an art song. However, as shown in the above Figures (1, 2, 3), we believe at least four aspects of the multimodal genre are hard to be reflected clearly or examined specifically in a single form of CIN:

- (1) The cross-comparison of the data and the overall pattern of a multimodal corpus
- (2) The interaction between the multimodal and the conceptual level
- (3) The involvement of the multimodal cues and the prominent contrasting features within a single mode, i.e. the visual or aural mode
- (4) The audience interpretation of the conceptual and multimodal level

To bridge these gaps, this study proposes an alternative approach of “Multimodal Fusion Model” as shown in Figure 4 that incorporates the conceptual level and multimodalities, including the verbal, visual, and aural modes. This is an interactive model induced from the cross-comparison of the

multimodal corpus of our political cartoon data and poetry music data. We view the multimodal fusion as a modest model with more flexibility for the unique circumstance when one topic or concept has been represented in various multimodal art works.

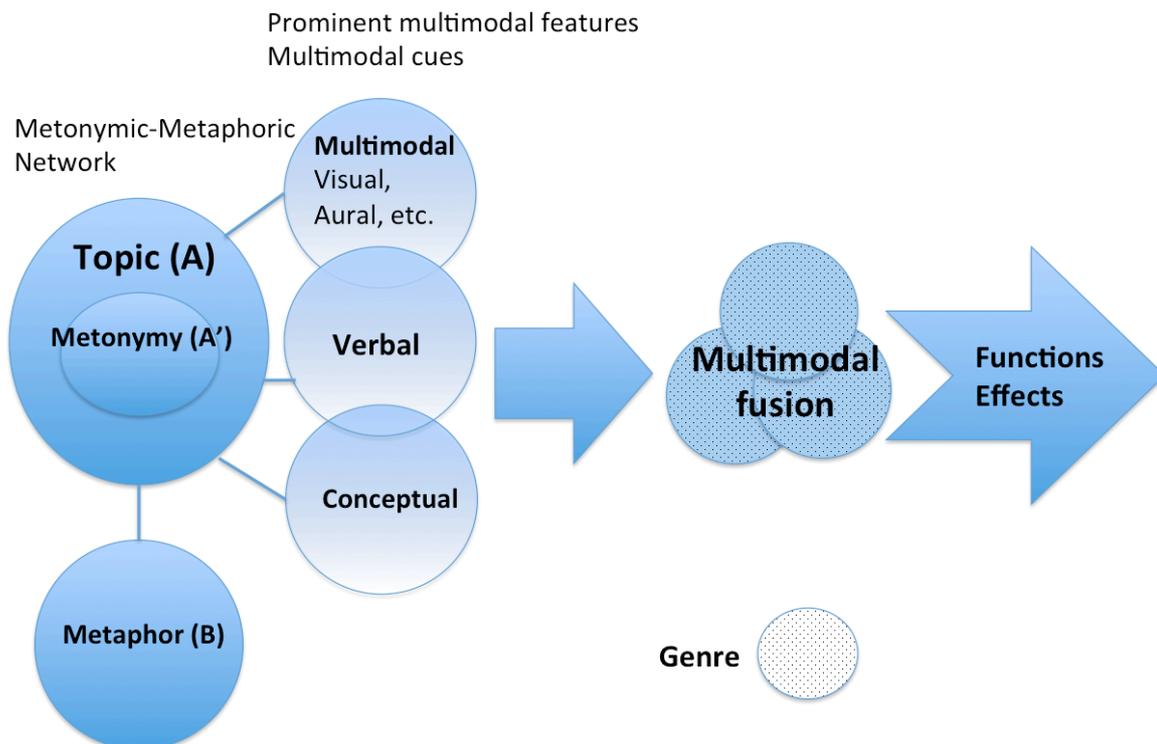
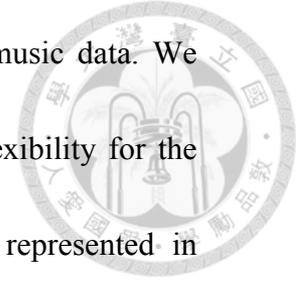


Figure 4. Multimodal Fusion Model

To clarify, our multimodal fusion model emphasizes that the conceptual level, the verbal, visual, and aural modes are all encapsulated and integrated simultaneously in the multimodal discourse thus we consider the conceptual level and the verbal mode, visual mode, or aural mode, naturally connected and overlapped with each other, whereas Blending Theory views different

concepts or modes as in distinct input spaces connected with vital relations that are drawn with the blended lines.



The term “fusion⁴” for our multimodal fusion model can be attributed to the *fusion* metaphor (Carroll, 1996) or *hybrid* metaphor (Forceville, 1996). The film scholar Carroll (1996) strictly viewed visual metaphor as a visual fusion of elements from two separate areas into one spatially bounded entity, while Forceville (1996: 163) specifically defined hybrid metaphor as ‘target and source visually amalgamated into one spatially bounded object’. Previous studies show that hybrid or fusion metaphor is common in the pictorial variety of advertisements and political cartoons (Forceville, 1996; Bell, 2004). Our data indicates that “fusion” not only largely exists in the visual mode but also in the combination of multimodality and conceptual level in multimodal genre. To distinguish from visual fusion or hybrid (Forceville, 1996, 2006) widely discussed in pictorial trope, we propose that multimodal fusion should be defined as “when the target and the source are verbally, visually, aurally (multimodally) and conceptually amalgamated” to emphasize the equivalently significant roles of multimodal and conceptual level in multimodal genre.

⁴ In fact, the concept of “fusion” has been largely applied in other fields of studies, including nuclear physics, biochemistry, medicine, immunology, computing, and electrical engineer, etc. In particular, “multimodal information fusion” has been viewed as an important strategy to accomplish multiple multimedia analysis tasks in the field of applied computer science (Atrey, Hossain, El Saddik, & Kankanhalli, 2010). Viewing from the interface between language and multimodality, we believe “fusion” also naturally exists in multimodal genre, thus this study proposes the “multimodal fusion model” to account for the combination of multimodal, verbal, and conceptual level.

Through our analysis, the multimodal data indicates that the verbal, visual, and aural cues play significant roles in facilitating the interaction between different modes and the conceptual level, thus we believe it is necessary to examine the multimodal cues specifically in each mode in our model while interpreting the multimodal data.

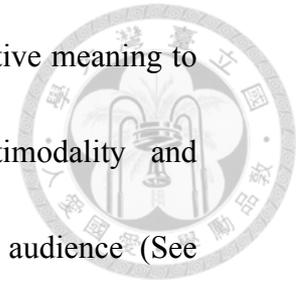
In addition, the contrasting visual/aural features within the visual/aural mode reflected from the data, for example from the cartoon image or the musical score, are often the key elements that influence how certain topic or concept is conceptualized in different versions of work. Thus in our model, we suggest viewing the contrasting features in each mode with the consideration of the unique characteristics of the “genre” in question. In our model, we extend Forceville’s (2014) claim that ‘genre is the most important interface between text and context’ by emphasizing the significant role of “genre” as the “key base” in multimodal fusion (Figure 4), leading to its cognitive-pragmatic functions and effects. We believe the data speaks for itself and provides the best evidence from its genre.

Thus in the following sections, we will elaborate our analysis in the detailed examination of the data (i.e. political cartoons and musical scores) under the scope of our model. Moreover, we will provide the

cross-comparison analysis among the data that helps develop and strengthen the model. In this study, we propose an interactive interpretation model of multimodal fusion that provides a “multimodal discourse analysis” to account for different aspects of the unique one-on-many relationships between language and the visual/aural modes in specific genre.

As Figure 4 shows, the “metonymic-metaphoric network” is considered the fundamental basis of the model since we view multimodal fusion model as the cognitive mechanism that works at the interactive level of multimodal metaphor and metonymy, hoping to “recognize the diversity of forms and semantic content that make up different tropes” (Gibbs and Colston, 2012, p.342). Regarding the theoretical background of multimodal fusion model, we do not agree that a single theory or model can be generalized and account for how all aspects of figurative language is understood. As Gibbs and Colston (2012, p. 3-4) advocates that ‘people may not process all figurative meaning in the same way, precisely because the kinds and forms of different tropes are sufficiently varied as to resist classification within a single theoretical framework’, therefore we extend the central tenets of Conceptual Metaphor Theory, Multimodal Metaphor, and Conceptual Blending Theory, to develop our model (See Chapter 2). We further incorporate “audience responses” with

our analysis to extend the discourse-internal analysis of figurative meaning to an interactional level of interpretation analysis on multimodality and metaphor/metonymy from the perspectives of analyzer and audience (See Chapter 4 and 6).



With special concern on the unique on-on-many relationship, this study focuses on the data collected from political cartoon and poetry music, which combine verbal, visual, and aural modes that reflect Mandarin language and culture, to investigate how certain topic or concept has been represented in various representations of specific multimodal genre. With the primary focus on multimodal metaphor and metonymy, this study will demonstrate how the verbal mode, visual/aural mode, and the conceptual level are interactively fused and interrelated in the genre of political cartoons and poetry music. We will illustrate how certain topic and concept (i.e. politics, emotion, space, etc.) are conceptualized and construed differently in the multimodal genre accompanied with the use of prominent multimodal cues (visual/aural/verbal cues) that facilitate multimodal realization and interpretation.

In the following, we will follow the key elements of Multimodal Fusion Model to introduce the significant findings based on the evidence from political cartoons and poetry music.

1.4 Multimodal Fusion in Political Cartoons⁵



In this study, we propose the multimodal fusion model to account for the cognitive mechanisms involving political cartoons with regard to U.S. beef import issues as reported in Taiwanese newspapers. This study focuses on the topic of “U.S. beef imports” since it has become a worldwide issue causing global and local concerns regarding environmental and public health problems over the past few years.

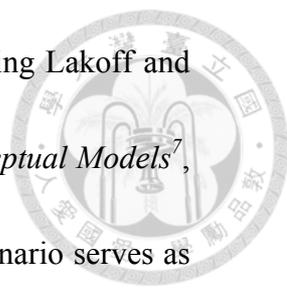
1.4.1 Metonymic-Metaphoric Network and Conceptual Scenarios in Political

Cartoon

In this study, we demonstrate how metonymy and metaphor are interwoven in the process of multimodal fusion and how metonymy-metaphor network has been involved with the conceptual level and realized in the multimodal genre of political cartoons. In view of the conceptual level, we suggest that the “conceptual scenario”⁶ initiated by visual and verbal modes in

⁵ The political cartoon analysis presented in this study has been extended from our paper published in *Metaphor and Symbol* (Lin and Chiang, 2015). The previous version of this article was presented at the 8th International Conference on *Researching and Applying Metaphor* (RaAM) in Amsterdam, the Netherlands in 2010.

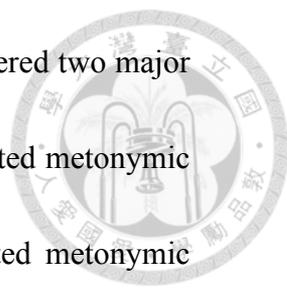
⁶ To clarify, we adopted the term “conceptual scenario” extended from Lakoff and Johnson’s (1980) *scenario metaphor* and Schilerpoord and Maes’s (2009) viewpoint of *scenario* in political cartoon studies. We understand the concept of the terms



political cartoon is essential on the conceptual level by extending Lakoff and John's (1980) view of "scenario metaphors" as *Idealized Conceptual Models*⁷, and Schilperoord and Maes's (2009: p. 224-225) claim that scenario serves as 'a supplier of various metaphorical relations' and 'imposes an additional conceptual structure onto the image by virtue of the conceptual relations that hold between the scenario elements' in political cartoon studies. In the multimodal fusion model, we propose that visual and verbal details in each political cartoon contain the conceptual scenario(s) of the source domain, which involves relevant metaphorical mappings through scenario elements of people, objects, roles, relations, and status (for further discussion on conceptual scenarios, please see section 3.3).

"scenario" and "frame" are often overlapped in their definitions upon the conceptual representations or knowledge of human experience. Viewing from frame semantics terminology, "frame" refers to the semantic frame of encyclopedic meaning activated or evoked by a word that relates to the specific concept it refers to. Originally being applied to lexemes, frame semantics (Fillmore, 1985, 1992, 2006) has been expanded to grammatical constructions as the main semantic principle in Construction Grammar and other larger and more complex linguistic units as in Conceptual Blending theory. In this study, we did not use the term "*frame*" to avoid confusion with the expanded use of political "frame" in *Don't think of an elephant: Know your values and Frame your debate* (Lakoff, 2004).

⁷ "The cognitive networks with causal, temporal and other sorts of relationships between persons, roles, locations, and attributes which all are more or less fixed, and conventionally known by all members of the cultural community in which they appear." (Lakoff and Johnson, 1980; Schilperoord and Maes's, 2009: p. 225)



Through our analysis of the multimodal corpus, we discovered two major metonymic-metaphoric networks of multimodal fusion: (1) related metonymic network, and (2) diversified metaphoric network. In the related metonymic network, the visual image is mainly comprised of metonymies closely related to the key concept or keywords related to U.S. beef events, such as “bull” for “U.S. beef”, thus the relevant visual images, such as bullfighting, cowboys, etc., are productively used to conceptualize U.S. beef events as a (i.e. powerful/playful) bull or a personified bull. Furthermore, these relevant metonymies are frequently associated with the metaphorical mappings of conceptual scenarios related to the metaphors POLITICS IS GAME and POLITICS IS WAR, revealing the close connection between metonymy and metaphor. On the other hand, while in the diversified metaphoric network, the visual image is mainly comprised of novel metaphors that are implicitly related to the key concept or keywords connected to U.S. beef events, thus U.S. beef events are conceptualized as a rock, bomb, or net, etc., in diversified scenarios. In diversified metaphoric network, we found that the verbal mode functions as the crucial link to connect the visual image with the conceptual scenario in conveying the complex political ideologies and critical messages in political cartoons.

1.4.2 The Conceptual Metaphor and Prominent Visual Features Contrast



Functions as Visual Cues

This study demonstrates that the abstract and complex concept of politics has been productively conceptualized through GAME and WAR domains in multimodal discourse of political cartoons. As a result, our multimodal corpus shows that the conceptual metaphors, including POLITICS IS WAR and POLITICS IS GAME, are most frequently associated with the metonymies used to conceptualize the general aspects of the political issue of U.S. beef imports in both newspapers.

However, the distinct political stances of LT and UDN resulted in subtly different representations underlying the multimodal fusion in the political cartoons. Our analysis demonstrates that visual features contrast, for instance, between size, facial expressions, or positions, function as “prominent visual cues” to highlight the critical messages of the political cartoons. We will further elaborate how the distinct political stances of the two newspapers subtly influenced the use of multimodal fusion in representing controversial debates in the political cartoons (For further details, please see Chapter 3).

1.4.3 Genre: Cognitive-Pragmatic Functions and Critical Effects in Political

Cartoons

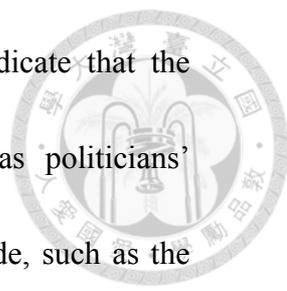


Our findings indicate that the cognitive mechanism of multimodal fusion with the fundamental basis of related metonymic network and diversified metaphoric network is a significant and recurring representation technique in the genre of political cartoons.

From our multimodal corpus, we discovered that the representation technique of “multimodal fusion” has the cognitive function that encapsulates abstractly complex critical messages and political ideologies efficiently to create ironic and humorous effects in the genre of political cartoons. Based on the unique characteristics of the political cartoon genre, our findings clearly demonstrate how knowledge can be encapsulated and condensed through multimodal fusion of the conceptual, the visual mode, and the verbal mode, creating the critical effects of irony and humor while conveying serious debates on political issues (For further details, please see Chapter 3).

1.4.4 Audience Response to Multimodal Fusion in Political Cartoon

Besides the detailed analysis of multimodal fusion, we also conducted surveys to investigate how the audience interprets the conceptualization of political issues concerning U.S. beef imports through multimodal fusion of



visual and verbal modes in political cartoon. The results indicate that the audience needs to first recognize the visual mode, such as politicians' caricatures, and understand the connotations of the verbal mode, such as the meaning of "CODEX", in order to build the connection with the conceptual level. Moreover, prominent multimodal cues of the visual and verbal mode are interactively combined with the conceptual level and influence audience interpretation on the conceptual scenario. However, the results reveal that different conceptual scenarios of the conceptual level may be activated and highlighted for different audience. We propose that various conceptual scenarios are involved on the conceptual level of multimodal fusion, while audience may view specific conceptual scenario as more prominent and significant based on their interpretation. In view of the conceptualization of the political issues among the representative examples, 89% of the participants consider the frequency of GAME domain is high, whereas 67% of them remain neutral stance toward the frequency of WAR domain. Based on the quantitative and qualitative analysis of the survey, we suggest that the higher level of audience interpretation relies on the effective recognition, comprehension, and connection of the verbal mode, visual mode, and conceptual level of multimodal fusion in political cartoons (For further details,

please see Chapter 4).

In the next section, we will present significant findings on multimodal fusion in the genre of “poetry music”.



1.5 Multimodal Fusion in Analyzing Poetry Music⁸

In this study, we adopt the multimodal fusion model to examine the unique genre of “poetry music”, which refers to the music compositions purposefully composed for the poetry as its lyrics (previously introduced in section 1.2). This study provides a detailed account of how the verbal mode, the aural/musical mode, and the conceptual level are interactively combined and contrasted to represent specific metaphors, metonymies, and metaphorical expressions in four versions of poetry music to examine three insightful aspects: (1) the interactive relationship between metaphor and metonymy, the metonymic representation of metaphor and holistic representation of the “chain of metaphor”; (2) the multimodal representation extended from the embodied basis of conceptualization of abstract concepts like space or emotion; and (3) the prominent multimodal features and multimodal cues that facilitate metaphor realization and interpretation. We will further elaborate how these

⁸ The poetry music analysis presented in this study has been extended from our paper presented at the 5th Conference on Language, Discourse, and Cognition (CLDC) in Taipei, Taiwan in 2011.

aspects are analyzed through multimodal fusion model.

1.5.1 Metonymic-Metaphoric Network: Metonymic representation of

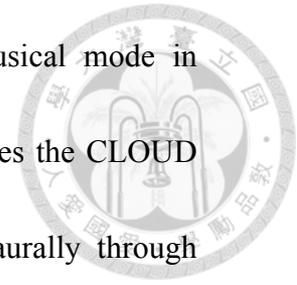
Metaphor and Chain of Metaphor in Poetry Music



In this study, we discovered the metonymic-metaphoric network evolves through the metonymic representations of metaphor in poetry music. We found the “pictorial effect” on the score assembles or simulates the concrete image of certain part of the source domains of the metaphor, for example, the visual shape of cloud appears on the score to represent CLOUD metaphor. In addition, we also found the beating rhythm of the ocean waves occurs to represent SEA metaphor. Such part-for-whole metonymic representations of metaphor have been realized through the multimodal fusion of the visual, aural, and verbal mode. The examples demonstrate the inseparable relationship between metonymy and metaphor in the multimodal genre of poetry music.

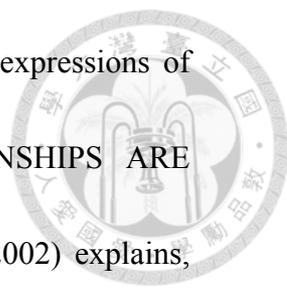
Extending from the metonymic representations of metaphors, we discovered the “chain of metaphors”, an interactive and echoing relationship between metaphors in poetry music, connects and highlights the concrete image of metaphors, such as CLOUD, REFLECTION, and LIGHT metaphor. The “chain of metaphor” has been connected through the echoing musical features represented by each metaphor, creating a holistic representation of

multimodal metaphor combining the visual, verbal, and musical mode in poetry music, for example, the REFLECTION metaphor echoes the CLOUD metaphor visually through similar notes on the score and aurally through music patterns and techniques. We propose the echoing status of the chain of metaphors may be due to the coherent characteristic of melody in the musical mode.

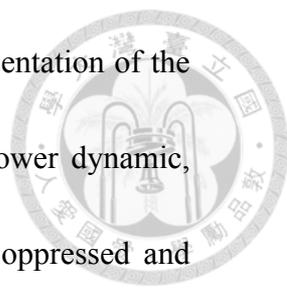


1.5.2 The Conceptual: Embodied Basis of Conceptual Metaphor

In view of the conceptual level, we examine how the conceptualization and construal of abstract concept including emotion, space, etc. and the embodied basis of conceptual metaphor are naturally and closely connected with the visual, aural, and verbal modes of poetry music through multimodal fusion. Our findings reveal the embodied basis of how the concrete image of metaphor and the abstract concept of emotion that underlies the metaphor have been conceptualized through prominent musical features, such as the change on rising and falling figures, major and minor key, speed, volume, pitch, tone range, timbre, etc., which we consider the “aural cues” that activate listeners’ imagination and interpretation of metaphors in poetry music. Previous studies on embodiment and music have shown ample evidence that the characterization of musical pitches in terms of high and low is basically



metaphorical. The conceptual metaphor behind the linguistic expressions of playing the higher or lower notes is PITCH RELATIONSHIPS ARE RELATIONSHIPS IN VERTICAL SPACE. As Zbikowski (2002) explains, mapping up-down onto pitch allows us to import the concrete relationships through which we understand space into the domain of music, and thereby provide a coherent account of relationships between musical pitches. Extending this point, our findings demonstrate that besides pitches, the concrete relationship of “space” can also be mapped onto the prominent musical features, such as increasing and descending figures, lower and higher key/notes, and tone range in the musical domain. We propose that the high-low representations of the prominent musical features not only realize the conceptualization of space in poetry music but also function as the aural cues that activate listeners’ simulation of the image of metaphors in poetry lyrics and thus enable listeners to appreciate the emotions conveyed through poetry music. Extended from the embodied basis of the orientational metaphor HAPPY IS UP and music metaphor HIGH IS BRIGHT, LOW IS DARK, we also found the scale of emotion can be mapped onto the representations of prominent musical features in poetry music. For example, increasing dynamic, higher key, and bright timbre are often associated with positive emotions like



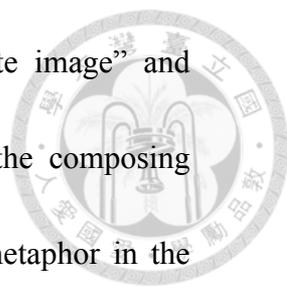
expectation, excitement, and enthusiasm as shown in the representation of the LIGHT metaphor in versions 1 and 2. On the contrary, the lower dynamic, lower pitch, and dark timbre are often associated with more oppressed and sentimental emotions as shown in the representation of the SEA metaphor in versions 1 and 2 (For further details, please see Chapter 5).

1.5.3 Multimodal Cues in Poetry Music

Based on multimodal fusion model, our analysis suggests that prominent musical features, such as increasing and descending figures, wide or narrow tone range, lower and higher keys/notes, bright or dark timbre, are considered as aural cues, while the pictorial effect shown on the score is considered as visual cues. The visual cues and aural cues that underlie the composing patterns and musical techniques not only facilitate metaphor realization in representing the abstract concept, such as love, emotion, and space, in poetry lyrics but also link to the imagination process and aesthetic experience of poetry music.

1.5.4 Genre: Musical Effects and Aesthetic Functions

Our findings provide a categorization of the concrete-abstract dimensions of composing patterns and musical techniques, indicating that the poem as poetry lyrics has been transformed into the music mode visually and aurally



via each composer's highlights of two aspects, the "concrete image" and "abstract metaphorical extension" of metaphors in view of the composing patterns and musical techniques. For instance, the CLOUD metaphor in the poetry lyrics has been visualized and musicalized through repetitive rising-falling figures on the score, which resemble the concrete image of cloud and create the "pictorial effect"; on the other hand, the abstract metaphorical extension refers to the underlying feelings and emotions expressed through metaphors and metaphorical expressions, for example, the abstract metaphorical extension of the sentimental emotion that underlies the CLOUD and REFLECTION metaphors has been expressed through "sighing figures".

Our findings show that the four musical versions are distributed on the continuum between the concrete and abstract dimensions of metaphorical representations in the musical mode. Overall, this study demonstrates the significant difference between the classical art music of versions 1, 2, and 3, and the popular music of version 4; while art songs "reinterpret" and "transform" each metaphor from the language mode to the musical mode in visual and aural aspects of the concrete image of metaphor, the popular song expresses the gestalt emotion and feelings of the poetry instead of translating the representations of each metaphor in poetry lyrics in detail (For further

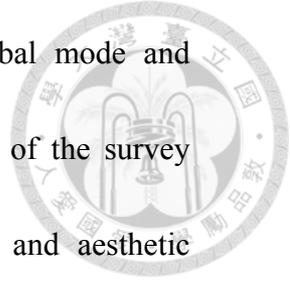
details, please see Chapter 5).



1.5.5 Audience Response to Multimodal Fusion in Poetry Music

In this study, we conducted surveys to investigate how the audience interprets the transformation of metaphor and musical effects through multimodal fusion in poetry music. The results show that the audience distinguishes different musical effects of the concrete image of “CLOUD” metaphor in four versions of songs. Interestingly, the audience responses to CLOUD metaphor correlate with our analysis of the “concrete-abstract continuum”. We found that the audience unconsciously describes the concrete aspect of shape, texture, weight, and status of CLOUD metaphor when they interpret the metaphor in a more concrete version (version 1 and 2), while the audience focuses more on the abstract aspect of the atmosphere, feelings, and emotion conveyed through CLOUD metaphor when they interpret the metaphor in a more abstract version (version 3 and 4). The audience response echoes with our claim that each version of song is in fact displaying the unique combination of multimodal level and conceptual level in multimodal fusion, demonstrating how the abstract concept in the poem can be conceptualized and construed in various ways with distinct focuses and musical effects. In addition, the results suggest that the audience’s preference for certain version

of song is relevant to the correspondence between the verbal mode and aural/musical mode. The quantitative and qualitative analysis of the survey reflects how the audience “feels” about the musical effect and aesthetic functions created from multimodal fusion in different versions of songs composed for the specific poem (For more details, please see Chapter 6).

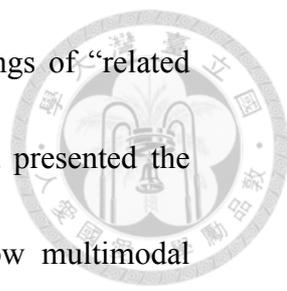


1.6 Organization of the Dissertation

The remainder of this dissertation is divided into eight chapters.

Chapter 2 presents the literature review of studies on political cartoon and studies on music and language. We also extended previous viewpoints concerning metonymy and metaphor to stress the significant role of metonymy and its inseparably connection with metaphor. In addition, we reviewed the theoretical background and central tenets of Conceptual Metaphor Theory, Multimodal Metaphor, and Conceptual Integration Theory in order to develop the fundamental basis for Multimodal Fusion Model. Furthermore, we will demonstrate how Multimodal Fusion may fill the gaps in Conceptual Blending Networks with the consideration on the unique characteristics of the specific genre in the multimodal analysis to connect with the following analysis in the next chapter.

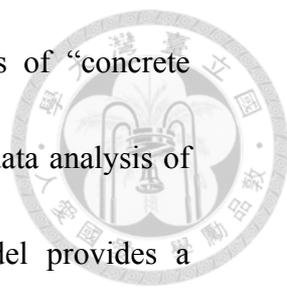
In Chapter 3, we first introduced the data description regarding the research



topic of U.S. beef imports. Then we illustrated the major findings of “related metonymic network” and “diversified metaphoric network” and presented the data analysis of representative examples in detail to show how multimodal fusion model provides a fundamental basis in analyzing political cartoon. Moreover, we identified the significant role of multimodal cues, visual and verbal cues. We also examined the distribution of the conceptual scenarios in the multimodal corpus and found that POLITICS IS GAME and POLITICS IS WAR are the most frequently identified scenarios on the conceptual level.

Chapter 4 provides the significant findings of audience responses in interpreting political cartoon based on the quantitative analysis and qualitative analysis. As a result, there are significant differences in the interpretation level and in the distribution of the conceptual metaphors, POLITICS IS GAME and POLITICS IS WAR, among ten cartoons; there are significant differences in the distribution of the involvement of the conceptual scenario in the interpretation process in each cartoon; and there are significant frequencies. We have examined the general tendency of audience response from the statistical analysis and discussed it with the qualitative analysis of the surveys.

In Chapter 5, we introduced the data description regarding the topic of poetry music and the poem “Serendipity”. Then we illustrated the



metaphor-based musical analysis to present the major findings of “concrete image” and “abstract metaphorical extension” with the detailed data analysis of four versions of songs to show how multimodal fusion model provides a fundamental basis in analyzing poetry music. Moreover, we discovered the unique representation of “chain of metaphor” and proposed how embodied basis of prominent musical features conceptualizes the abstract concept (i.e. space, emotion) in poetry music.

Chapter 6 provides the significant findings of audience responses in interpreting poetry music based on the quantitative analysis and qualitative analysis. We have found the general tendency of audience response from the statistical analysis and further examine it with the qualitative analysis of the surveys. The results indicate that there are significant differences in the matching level of the concrete image of “CLOUD” metaphor and in the preference scores among four versions of songs. We have examined the general tendency of audience response from the statistical analysis and discussed it with the qualitative analysis of the surveys.

Chapter 7 summarizes the significant findings and concluding remarks of this study. We hope to shed light on the interdisciplinary studies on the interface between multimodality, metaphor, and interpretation.

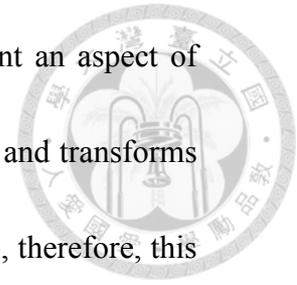
Chapter 2 Literature Review and Theoretical Background

In this chapter, we will first review literature and previous studies concerning our topic genre, political cartoon and poetry music, to clarify the central claims of multimodal fusion model. And we will review the central tenets of Conceptual Metaphor Theory, Multimodal Metaphor, and Conceptual Integration Theory to develop the fundamental basis of the theoretical background of multimodal fusion model.

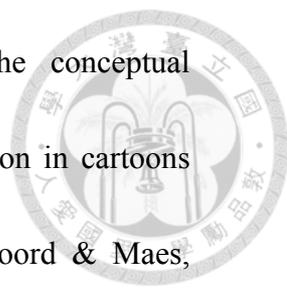
2.1 Studies on Political Cartoon

Political cartoons, unlike advertisements, aim to expose something negative or shameful instead of highlighting the positive, to affect states of minds, beliefs, points of view, and perspectives on socio-political affairs, and to express a particular critical stance on the topic (El Refaie, 2009a), which simultaneously appeal to several integrated but distinguishable domains of knowledge to define and evaluate the topic (Schilperoord, 2013). Typically published on the editorial or comments pages of a newspaper, the frame of a political cartoon has been viewed as an implicit metacomment that functions as the signal to the readers as part of the dramatic cartoon world contrasting to the real world of news reports, commentaries, and newspaper photographs (Baldry & Thibault, 2006). The genre of political cartoons frequently uses

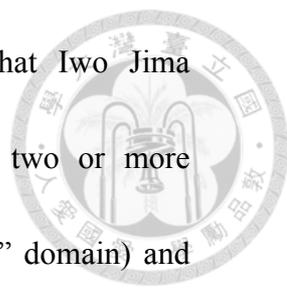
multimodal metaphors of the verbo-visual variety to “represent an aspect of social, cultural, or political life in a way that condenses reality and transforms it in a striking, original, and humorous way” (El Refaie, 2009a), therefore, this study believes the genre of political cartoon provides us interesting examples as evidence to explore the complexity lying in the interface between multimodality, metaphor, and interpretation.



Previous studies show that most scholars are interested in exploring how political cartoons reflect public opinions, race, gender, and cultural identities (Benoit, Klyukovski, Mchale, & Airne, 2001; Connors, 1998; 2010, Diamond, 2002; Edwards, 1997, 2001; Green, 2009; Kelley-Romano & Westgate, 2007; Lumsden, 2010; Morrison 1992; Seymour-Ure, 1998; Vultee, 2007). Several studies have examined the rhetorical forms of political cartoons by examining how the metaphors represent certain issue (Bounegru, Liliana, & Forceville, 2011) or by focusing on cartoonists’ intentional means to comment or criticize socio-political actors and events (Bostdorff, 1987; Medhurts & Desousa, 1981). Some scholars investigated the functions of political cartoon as drawing style and aesthetics (Davies, 1995), as visual argumentation (Groake, 2002), or as vehicle for political purposes (Kluver, 2000).

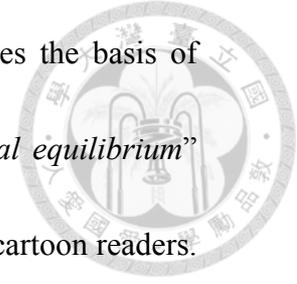


In the past decade, cognitive scientists focused on the conceptual representations by studying the process of meaning construction in cartoons (Cohn, 2007, 2010; Coulson, 2002; Gleason, 2009; Schilperoord & Maes, 2009; Schilperoord, 2013) and other genres like comics, advertising, digital design, film, etc. as well. Most previous studies advocate the use of conceptual blending to investigate the online processing of political cartoon. Coulson (2002) started to examine how the processes of conceptual blending work in humorous examples of political cartoons, claiming that conceptual integration is flexible and significant in perpetuating and modifying culturally relevant concepts. We agree that humorous examples in political cartoons ‘allow us to test the flexibility of our conceptual system and navigate the space of possible construals’ (Coulson, 2002), however, it remains problematic to verify if and how the conceptual structure of humor actually “arise” in “blended spaces”. With a mental-space approach, Schilperoord (2013) explored the conceptual representations of readers by examining the structure and processing of Iwo-Jima-Inspired political cartoons, while the author views cartoons as multi-domain cartoons that are capable of expressing complex conceptual structures (i.e. analogy, causation, time passing, etc.), thus meaning construction results from representing and interconnecting these



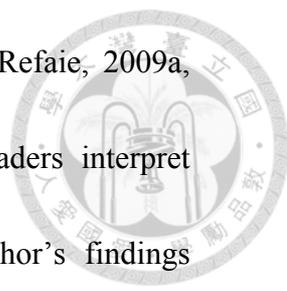
multi-domains. Under such view, Schilperoord proposed that Iwo Jima Inspired cartoons are visually integrated representations of two or more conceptual domains (the reference domain and the “Iwo Jima” domain) and are processed and interpreted by *deconstructing* the blended space (of which the cartoon represents) and *compare* the aligned domains. In other words, Schilperoord (2013) proposed that “de-blending” the blend represented by the cartoon helps interpret the cartoons concerning Iwo Jima images and suggested establishing *conceptual equilibrium* based on the *vital relations* between the contents of the input spaces to account for the negative and positive evaluations of the topics in the cartoons. Schilperoord’s (2013) analysis aims at showing how contrastive vital relations between input spaces requires readers capability of double-scope blending, demonstrating that the interpretation process is constructed on a reader’s conceptual equilibrium to reach positive or negative evaluation. The author concluded that ‘the theory of conceptual blending allows *isolating* the graphic and/or verbal features that provide access to conceptual domains and analyzing various possibilities of establishing equilibrium and thus, various ways of interpreting and comprehending cartoons’. However, we do not agree that visual or verbal features should be “isolated” in the first hand, we believe the interaction

between the visual, verbal, and conceptual altogether underlies the basis of interpretation. In addition, we are not sure if such “*conceptual equilibrium*” does exist in representing the real interpretation process of the cartoon readers.



As a matter of fact, it remains difficult to verify the mental-space approach with empirical data. Therefore, we are still uncertain if this may be the proper approach from a cognitive perspective.

Moreover, most previous studies have focused more on the internal aspects of analysis than on the interaction and communication between political cartoon and its audience. Very few studies have pinpointed the unique representation and effect of multimodal metaphor and metonymy in the combination of distinct modes (i.e. verbal/visual modes), which have their own ‘affordances and constraints’ as advocated by Forceville (2014). And most studies paid more attention to the cognitive online processing than how audience actually interprets political cartoons and how multimodal cues function in the interpretation process. To fill these gaps, this study proposes the multimodal fusion model and incorporates “audience response” with our analysis, hoping to view the issue of interpretation and multimodal cues from the perspectives of analyzers and readers.



Regarding the audience responses, previous studies (El Refaie, 2009a, 2009b) have investigated readers' responses to see how readers interpret multimodal metaphors of the verbo-visual variety. The author's findings suggest that most readers understand some metaphorical mappings in cartoons, such as those between size and power/status, between movement through space, and the passing of time more generally at a more intuitive level, while they tend to interpret more elaborate structural metaphors in different ways (El Refaie, 2009a). The author include readers' responses and interviews with cartoonists to suggest how cartoon interpretation can be a complex process that 'requires people's broad knowledge of past and current events, a familiarity with the cartoon genre, a vast repertoire of cultural symbols, and experience of thinking analytically about real-world events and circumstances' (El Refaie, 2009b). Although previous studies conducted a relatively small-scope study of interviews with small number of participants and data, the author pointed out significant aspects and provided insightful findings for the exploration of audience response in the genre. In this study, we will extend the scope of response surveys in the number of data and participants, and we will further incorporate qualitative analysis with quantitative analysis to present a general picture and tendency of the audience with detailed

information that sincerely reflects the unpredictably specific features of audience response.



2.2 Studies on Music and Language

With regard to the relations between language and music in song, they have long been debated in musicological studies throughout the history of Western music although rarely been discussed in the field of linguistics (Hsu and Su, 2014). The relations between language and music in song can be identified as four models (Agawu, 1992; Zbikowski, 2002; Hsu & Su, 2014), which are clearly reviewed and summarized from Hsu and Su (2014) as follows.

(1) Langer's (1953) assimilation model:

Langer considers song as not a compromise between language and music; words are transformed into musical elements, completely assimilated by music when words and music come together in a song. However, Agawu (1992) pinpointed that Langer's account neglects the process in which the music assimilates the words and what is left of the words after the assimilation.

(2) Kramer's (1984) model

Kramer views the text of a song as simply incorporated with the music instead of assimilating the music. Although this model is effective in viewing the interaction between the two modes, Kramer specifies the “ingredients” of song as an “alloy” (Schoenberg, 1983: 76; Agawu, 1992), which did not account for its identity.

(3) Agawu's (1992) modified model

Agawu first viewed song as a compound structure where word carry the primary semantic content and music serves to enhance the significant meanings of the words, however, this model fails to offer an independent account of musical meaning. Agawu proposes a modified model of viewing song as a confluence of three independent but overlapping systems: music, words, and song. However, Agawu (1992) also mentioned that this model is difficult to account for the dynamic process when text and music are combined as a song since this is a more static account of song.

(4) Cook's (1998) model:

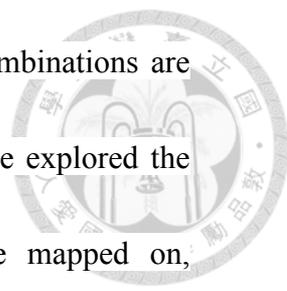
Cook holds the multimedia view of song as an instance of multimedia where words and music form two separate media, while music and words are two independent and separate media that interact and contest with each other, which is effective to account for many songs but not all of the songs. However, this model neglects the fact that song should be viewed as an integrated whole.

Regarding the models within traditional analysis of text and music, Zbikowski (2002) argued that few models or attempts viewed song as an integrated whole to examine how meaning is constructed by language and music. According to Zbikowski (2008), over the past 50 years, studies focused on the connections between music and metaphor encompass a range of disciplines, including philosophy, semiotics, cognitive science, and the critical and analytical study of music; scholars in these fields have investigated the topics of music and metaphor from distinct perspectives with the main focus on how "descriptive language" was used to characterize music. Recent studies have started to view music as a manifestation of human cognitive capacities inspired by the contemporary theory of metaphor (Lakoff, 1993) and the notion of image schema (Johnson, 1987), which provide a theoretical basis for metaphorical descriptions of music grounded in embodied experience (Feld,

1981, 1982; Saslaw, 1997-1998; Brower, 1997-1998; Larson, 1997-1998; Aksnes 2002; Borgo, 2004; Chuck 2004; Cox 2001; Johnson, 1997-1998; Walker, 2000) and a broader perspective on the cognitive capacities that shape humans' understanding of music (Spitzer, 2004; Zbikowski, 1991, 1998, 2002).

Under such view, Zbikowski (1999, 2002, 2009) proposed an integrated account of conceptual blending theory to explain language and music is combined to create artistic works with rich meanings and emergent structure. Zbikowski's works demonstrated how conceptual blending can be applied to song to explain the rich interpretative possibilities arising from the interaction between language and music. However, Hsu and Su (2014) pinpointed the ignored aspect of incongruity between text and music of song in Zbikowski's works. Hsu and Su (2014) thus propose a modified blending framework and provide insightful findings to highlight the significant aspect of incongruous text and music, claiming that such incongruity in song brings rich emergent meaning and lead to the iconic effect that reflects composers' intention, which also helps bring out the aesthetics of the works.

Though previous studies have shown exquisite findings on language and music from a range of topics, we believe the relations between language and



music can be distributed on ‘a continuum where countless combinations are possible’ (Hsu and Su, 2014: 145). However, few studies have explored the complex relations of how certain verbal metaphor can be mapped on, transformed to, and combined with musical modes in different compositions. Nearly none of the above studies provided a detailed metaphor-based musical analysis incorporating with audience interpretation responses to account for the one-on-many complex relations between language and music in compositions based on the same lyrics. Therefore, this study aims to fill the gap by providing a thorough examination of four musical versions composed for a classic poem in order to explain how certain metaphors and metaphorical expressions in poetry can be transformed to and represented in different musical forms with various musical approaches and effects, becoming multimodal metaphors in the verbal-musical mode of songs. Through our analysis, we aim to show how musical techniques combined with the verbal mode are used to create distinct musical effects and aesthetic feelings. We will also explore how aural cues and verbal cues facilitate metaphor interpretation and realization on the one-to-many correspondent relationship between the poetry and its four compositions.

2.3 Theoretical Background



In this study, we hope to propose a broad theoretical framework to understand the dynamic status of how people interpret various forms of multimodal metaphor and metonymy in different contexts and in specific multimodal genre. As Gibbs and Colston (2012, p. 3-4) advocates that ‘people may not process all figurative meaning in the same way, precisely because the kinds and forms of different tropes are sufficiently varied as to resist classification within a single theoretical framework’, we believe that there may not be a generalized single theory or model that accounts for how all aspects of figurative language is comprehend. Therefore, in developing the theoretical background of multimodal fusion model, we believe it is necessary to review the relevant viewpoints from the significant theories of figurative meaning interpretation in developing multimodal fusion model. In the following, we will review the central tenets of Conceptual Metaphor Theory, Multimodal Metaphor, and Conceptual Integration Theory and pinpoint the advantages and limitations of these theories to explain how our “multimodal fusion model” may extend the solid findings of these theories yet fill the research gaps in order to provide more flexibility in describing how the verbal mode, visual mode, aural mode, and the conceptual are interactively combined

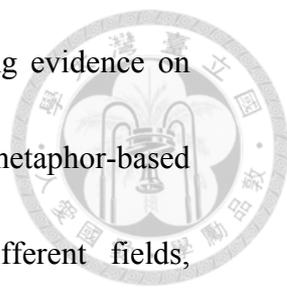
to create various meanings and effects in specific multimodal genre.



2.3.1 Conceptual Metaphor Theory and Multimodal Metaphor

Over the past few decades, Conceptual Metaphor Theory (CMT) (Lakoff & Johnson, 1980/2003, 1999; Gibbs, 2008) holds that metaphor is central to thought and language since it organizes human experiences through cognitive processing and structures our thinking and knowledge. Previous studies extended the central claim of CMT to explore the cognitive process of understanding and experiencing one thing (the target) in terms of another (the source) in the field of cognitive science, linguistics, psychology, mathematics, art, and philosophy (Lakoff & Johnson, 1980/2003; Gibbs, 1994; Johnson, 1987, 1993; Lakoff, 1993; Lakoff & Johnson, 1999; Lakoff & Nunez, 2000; Lakoff & Turner, 1989; Lin & Chen, 2014; Sweetser, 1990; Teng, 1999; Turner, 1991).

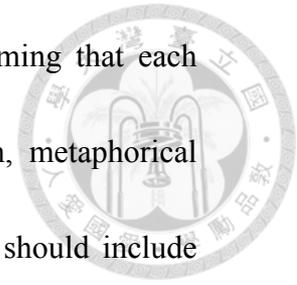
In addition, CMT claims that conceptual mechanisms arising from our sensorimotor experience and neural structures, including image schemas, metaphors, and other embodied imaginative structures, help us conceptualize and experience our world. The mechanisms of image schemas operating with conceptual metaphors enable human beings to employ the logic of our sensory-motor experience to perform high-level cognitive operations for



abstract entities and domains (Johnson 2005: 26); as mounting evidence on cognitive science shows, the process of image-schematic and metaphor-based understanding has been demonstrated for concepts in different fields, including mathematics (Lakoff & Nunez 2000), law (Winter 2001), morality (Johnson 1993), analogical problem solving (Craig, Nersessian, & Catrambone 2002), scientific causality (Lakoff & Johnson 1999), psychology (Gibbs & Colston 1995, Fernandez-Duque & Johnson 1999).

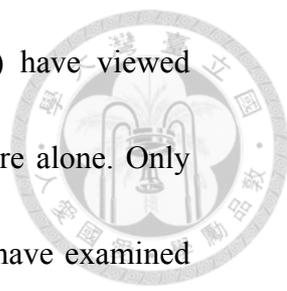
However, few researches focused on abstract reasoning and theorizing in the interface between language and multimodality. As the pioneer in multimodal metaphor research, Forceville & Urios-Aparisi (2009) pinpointed the fact that CMT will be restricted if verbal metaphor is the main evidence to validate its claim, which entails the risk of “circular reasoning” in cognitive linguistics research that “it starts with an analysis of language to infer something about the mind and body which in turn motivates different aspects of linguistic structure and behavior” (Gibbs and Colston, 1995: 354; Cienki, 1998). Therefore, Forceville & Urios-Aparisi (2009) emphasized that it is necessary to demonstrate how the central claims of CMT can occur non-verbally and multimodally as well as purely verbally. Furthermore, Forceville held the belief that “the medium is the message” (McLuhan 1964;

24 *et passim*) to stress the significant role of medium⁹, claiming that each medium or mode can cue, independently or in combination, metaphorical targets as well as sources, thus a full blown theory of CMT should include verbal metaphor as well as non-verbal and multimodal metaphor.



Under this view, “non-verbal” metaphors, especially the pictorial/visual metaphors (Forceville, 1996), in the pictorial trope have also received considerable attention in advertising, film, and art (Carroll, 1996; Forceville, 1994, 1996, 2000; Kaplan, 1990; Messaris, 1997; Teng & Sun, 2002; Whittock, 1990). Focusing on the visual mode “outside” the language domain, cognitive researchers investigated the meaning construction process in cartoons, political cartoons, visual arts, comics, advertisements, digital design, and film (Cohn, 2007, 2010; Coulson, 2002; Gleason, 2009; Schilperoord & Maes, 2009; Schilperoord, 2013). Recent studies viewing from the multimodal perspectives mainly focus on the relations between language, pictorial elements, and multimodal metaphor in advertising and films (Forceville, 2007, 2008, 2012a, 2012b; Urios-Aparisi, 2009; Yu, 2011a, 2011b) or between language and gestures (Muller and Cienki, 2009; Mittelbeg and Waugh, 2009;

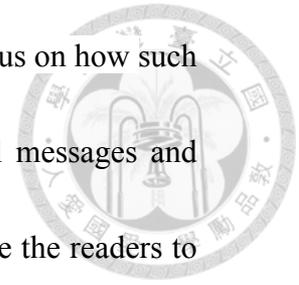
⁹ Forceville & Urios-Aparisi (2009:21) defined each medium “as a material carrier and transmitter of formation communicates via one or more signaling systems”.



Parrill and Sweetser, 2004). Scant studies (El Refaie, 2003) have viewed political cartoons as multimodal entity instead of “visual” genre alone. Only few studies (Zbikowski, 1999, 2002, 2009; Hsu & Su, 2014) have examined congruous and incongruous relations between language and music. Previous studies on multimodality have rarely focused on one-on-many relations between multimodality, metaphor, and conceptual level or the “embodied basis” of multimodal conceptualization in specific genre. Moreover, very few studies have incorporated researcher’s perspectives with audience responses to explore how audience actually interprets multimodal metaphor.

Another central tenets of CMT is that metaphor is ideological, since the interpretation of situations and events presented by any metaphor through “a coherent network of entailments that highlight some features of reality and hide others” is only partial and therefore flawed (Lakoff & Johnson, 1980/2003, p. 157; Deignan, 2005). Moreover, researchers have found that metaphors, based on their feature of highlighting, hiding, and simplifying, can be used to deliberately present a biased view of a situation or an event. Recent studies on multimodal metaphors concerning certain topic (Bain, Chaban, & Kelly, 2012; Bounegru & Forceville, 2011; El Refaie, 2003; Yu, 2011a, 2011b) demonstrate how metaphors are used to conceptualize specific topic or issues

in a particular genre, however, very few multimodal studies focus on how such use of multimodal metaphor is essential to delivering critical messages and how the contrasting multimodal features serve as cues to enable the readers to comprehend the underlying ideologies and power structure in different news media.

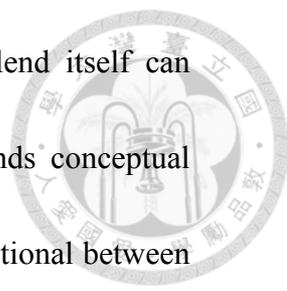


Incorporating the perspectives of researcher and audience, this study aims to fill the above research gaps to investigate the complex one-on-many relations and interactions between multimodalities and the conceptual level. We will further elaborate such research aims with data analysis in the following sections.

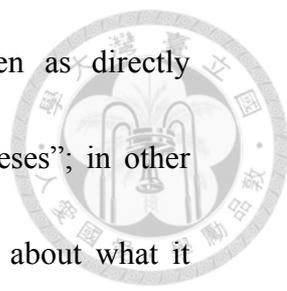
2.3.2 Conceptual Integration Theory

Conceptual Blending Theory¹⁰ (Coulson, 2001; Fauconnier & Turner, 2002/2008) claims that mental spaces can be invoked as partially specified constructs of frames or mental models during thought and communication, while multiple mental spaces can participate in the mapping processes involved in higher-order cognition such as counterfactual thinking and several aspects of language use, such as figurative speaking and understanding. Under this view, metaphor meaning can be captured by a blending space that inherits

¹⁰ In this section, we mainly summarized the review on Conceptual Blending Theory from Gibbs and Colston (2012: 109-113) to provide a unified theoretical overview based on recent studies in the field.



partial structure from each of the input spaces, while the blend itself can develop emergent content of its own. Blending Theory extends conceptual metaphor theory by allowing for mappings that are not unidirectional between multiple domains. Blending theorists providing linguistic analyses of humorous discourse, irony, and sarcasm (Coulson, 2005; Oakley & Hougaard, 2008; Veale, T., Feyaerts, K., & Brône, G, 2006) argued that the basic cognitive operation of “frame-shifting” underlies the process of incongruity resolution, which is critical to people’s interpretation and appreciation of verbal irony. Blending theory provides an explanation of how verbal irony is understood through the incongruity, claiming that “frame-shifting” involves a mental reorganization of the message level aspect of a speaker’s utterance. Experimental studies have shown that frame-shifting processes may require additional cognitive effort when interpreting language, however, as Gibbs and Colston (2012) suggests, measures of reading time and specific brain activity may only reflect overall mental effort and do not necessarily indicate that very specific cognitive processes, such as frame-shifting, must be operating when irony or humor is understood. Although the Conceptual Blending theorists see it as a highly generalizable tool that can explain a broad range of linguistic and cognitive phenomena, the challenge for blending theory is “to produce unique

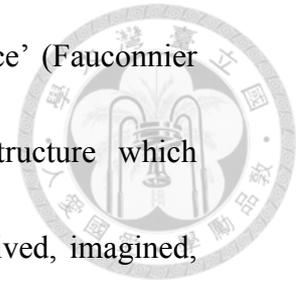


empirical predictions that can be tested explicitly and seen as directly supportive of blending theory compared to alternative hypotheses”; in other words, Blending Theory suffers from “a lack of both details about what it precisely and uniquely claims and interpreting figurative meaning that can be readily tested in experimental contexts” (Gibbs and Colston, 2012: 113; Tendahl, 2010).

In the field of Cognitive Linguistics, the framework of Conceptual Metaphor Theory (CMT) (Lakoff and Johnson, 1980/2003; 1993) and Blending Theory (BT) (Fauconnier and Turner, 2002) seek to investigate and unify a wide variety of linguistic and conceptual phenomena. Although both approaches view metaphor as a conceptual phenomenon rather than a purely linguistic phenomenon, involving systematic projection of language with constraints, there are crucial differences between CMT and BT (Grady, Oakley and Coulson, 1999), which is briefly summarized in the following:

(1) In CMT, language and conceptual structure from the ‘source domain’ are used to depict the situation in the ‘target domain’; particular elements of the source and target domains are chosen through the principled ‘mapping’, which is stored in the long-term memory that tells us how elements in the two domains correspond with each other. In comparison, the basic unit of

cognitive organization in Blending Theory is the ‘mental space’ (Fauconnier 1985; 1994), ‘a partial and temporary representational structure which speakers construct when thinking of or talking about a perceived, imagined, past, present, or future situation’.



(2) Unlike CMT that involves mappings between two conceptual domains, Blending Theory has a four-space model of mappings, which include two ‘input’ spaces, a ‘generic’ space, and the ‘blend’ space. The two ‘input’ spaces can be associated with the source and target domain of CMT in metaphorical cases, the ‘generic’ space represents the conceptual structure shared by both inputs, and the ‘blend’ space shows how elements from the input spaces blend and interact.

(3) While there is a “unidirectional” mapping from source to target posited by CMT, the four-space model material in Blending Theory is projected partially from both input spaces to the blend. The blend space not only inherits partial structure from each input space but also develops its ‘emergent structure’ resulting from the juxtaposition of elements from the inputs. Therefore, the four-space model of Blending Theory can account for phenomena that are not explicitly addressed by the two-domain model in CMT.

(4) Blending Theory has often been used to deal with novel, innovative, and unique examples which do not arise from entrenched or conventional mappings on which CMT mainly focus.



In this study, the central tenets of two approaches are viewed as the theoretical basis to unify different aspects of metaphoric conceptualization through our data. As previous research (Grady, Oakley and Coulson, 1999) proposes, these two approaches are complementary since ‘the conventional conceptual pairings and one-way mappings studied within CMT are inputs to and constraints on the kinds of dynamic conceptual networks posited within BT’, while the metaphor theorists aims to capture generalizations across a broad range of metaphoric expressions resulting from entrenched associations in long-term memory, the blending researchers typically focus on the particulars of individual examples through online processing, and they believe the ‘principles which speakers use to interpret metaphor may operate similarly across a wide range of non-metaphorical phenomena’ (Grady, Oakley and Coulson, 1999). Kovecses (2005) claims ‘metaphor and blending are viewed as creative mental processes that account for a great deal of variation in the use of figurative conceptualization’. Therefore, this study proposes that it is useful to extend the viewpoints of Conceptual Metaphor Theory, Multimodal

Metaphor, and Conceptual Blending Theory as the theoretical basis to develop the “Multimodal Fusion Model”.



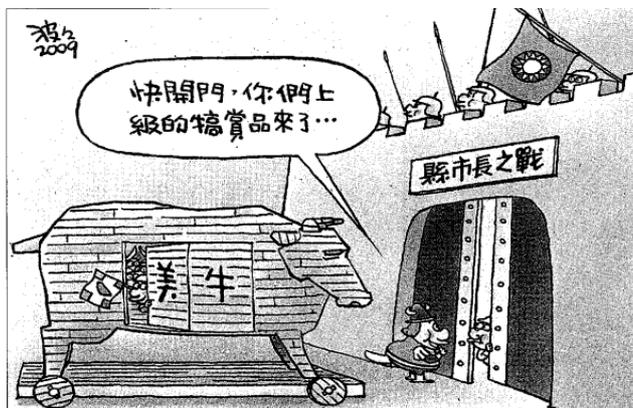
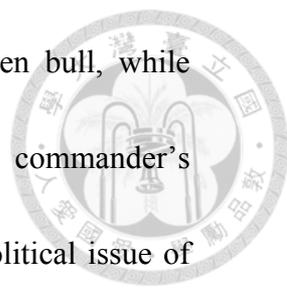
However, we will further explain why we need to develop multimodal fusion model. Although conceptual blending theory provides us to analyze multimodal data on different levels of correspondent mapping relationships, it has its limitations in the construction of “input spaces”. In the following, we will demonstrate this viewpoint with representative examples¹¹ from political cartoon and poetry music. The basic form of the Conceptual Integration Network (CIN) will be presented with specific focus on the conceptual and multimodal level as most previous studies did¹².

In view of example 1, the mapping relationships on the conceptual level constructed through the blends are clearly shown as in Figure 5. In example 1, the verbal tag “U.S. beef” on the body of the bull indicates the bull is the metonymy for U.S. beef import event, while the visual image of the gate setting resembles the Trojan war setting. The flag of the DPP (Democratic

¹¹ The representative examples are listed here with the Conceptual Blending Networks for clarification. These examples will be listed and discussed again in the following section in Chapter 3 and 5 for consistency and coherence of our data analysis in political cartoon and poetry music.

¹² To pinpoint the main problems concerning input space construction, we mainly focused on the input spaces constructions and graphically presented the main mapping elements in the basic form of input spaces of CIN in this study. In addition, it remains difficulties in deciding what elements should be included in the generic space as mentioned in previous studies (Teng and Sun, 2002; Coulson and Oakley, 2003; Hsu and Su, 2014), thus we followed the previous studies and did not graphically present the generic space in our discussion. Due to the complexity emerged from the emergent structure in multimodal data, we will include this part in our description instead of the graphics as well.

Progressive Party) waved by the person hiding in the wooden bull, while another person leading the bull said “Open the door, your commander’s rewards are coming”. The concept of “Trojan war” and the political issue of mayor election is integrated through vital relations of input spaces, thus the critical message can be conveyed through the emergent structure of the CIN that U.S. beef import event has been used as DPP’s political winning weapon/strategy to defeat the KMT in the upcoming mayor election as the Trojan horse in Trojan War.



Example 1 UDN (2009-11-16)

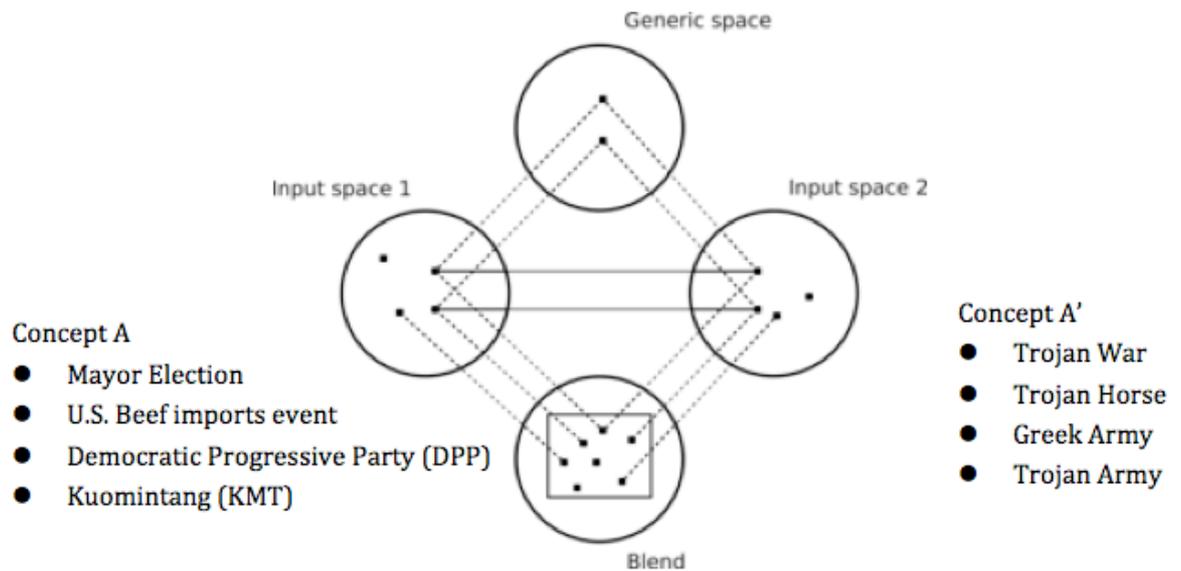


Figure 5. The Conceptual Level

On the conceptual level of CIN analysis, the Trojan horse in the original story is mapped onto the “U.S. beef imports event” in mayoral election. However, based on CIN, it remains unclear to identify how many input spaces should be included and to understand the principles governing the “selective projection” into the blend or the exact sequence of the vital relations constructions through online-processing of the conceptual integration of the “wooden bull”. On the other hand, the mapping relationships on the multimodal level can also be constructed through the blends in Figure 6. While examining the input spaces on the conceptual level or multimodal level, the vital relations between the input spaces are the main focus.

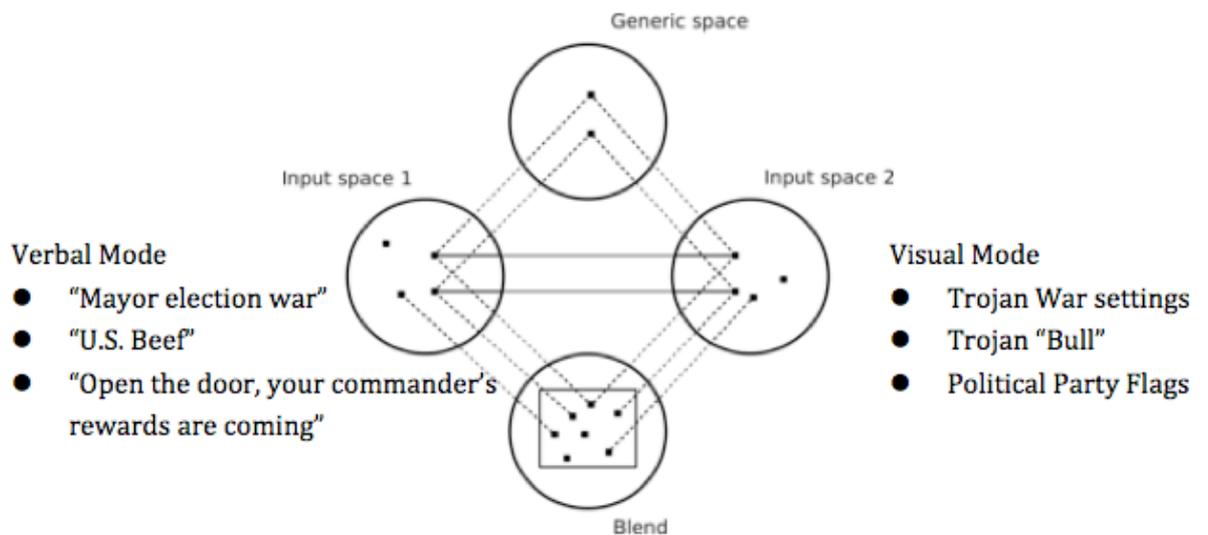


Figure 6. The Multimodal Level

However, the central question arises from the conceptual networks is whether it is possible to incorporate the conceptual level and multimodal modes simultaneously in the analysis since the multimodal modes and the conceptual level are in fact all integrated simultaneously in the multimodal discourse, as shown in example 1.

We believe the input spaces provide us a good starting point to mark clear mapping relationships in blending networks, however, we believe the interactions and relations on the multimodal and conceptual level are complex and beyond the identified mappings in the input spaces. Although different aspects can be specified and blended to create rich emergent meanings, the



blending network does have its limitations in accounting for important aspects from different levels in the data through the designated “blended” platform of mental spaces.

Furthermore, if we use conceptual blending network to analyze example 2, as Figure 7 and 8 show, we can at least obtain the basic forms of blended networks on the conceptual level and multimodal level as well.



Example 2 UDN (2012-07-27)

In example 2, the visual image of the rock marked with “U.S. beef” is the metaphor for U.S. beef events. However, only a part of the mountain road is shown in the foreground, whereas the other mountain with the flag “TIFA”¹³ appears far in the background. In this cartoon, the man resembling President

¹³ “TIFA” is the abbreviation for the “Trade and Investment Framework Agreement.” The Taiwan-U.S. TIFA serves as a primary platform for bilateral high-level economic and trade consultations between Taiwan and the United States.

Ma is pushing a huge rock with the verbal tag “U.S. beef” off the cliff, saying, “I can finally get rid of this huge rock that was blocking my way...”. Through CIN in Figure 7, the concept of mountain hiking has been integrated with the concept of President Ma’s political policy. The emergent structure thus implicates that President Ma is reaching his goal of signing TIFA with the U.S. government by agreeing to sign the new protocol on U.S. beef imports with more relaxations, however, it seems unrealistic and in vain.

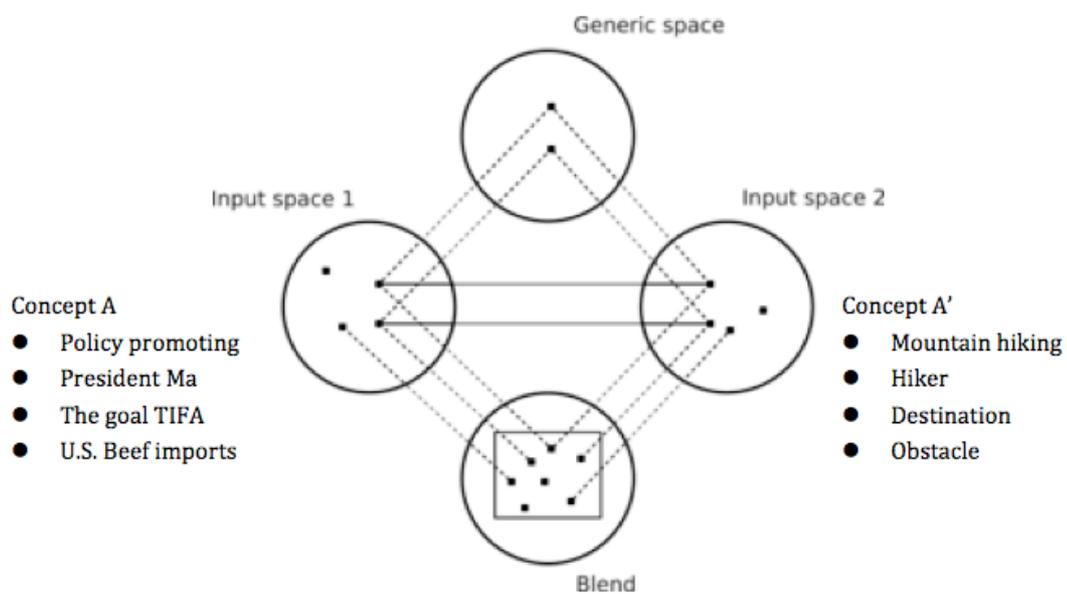


Figure 7. The Conceptual Level

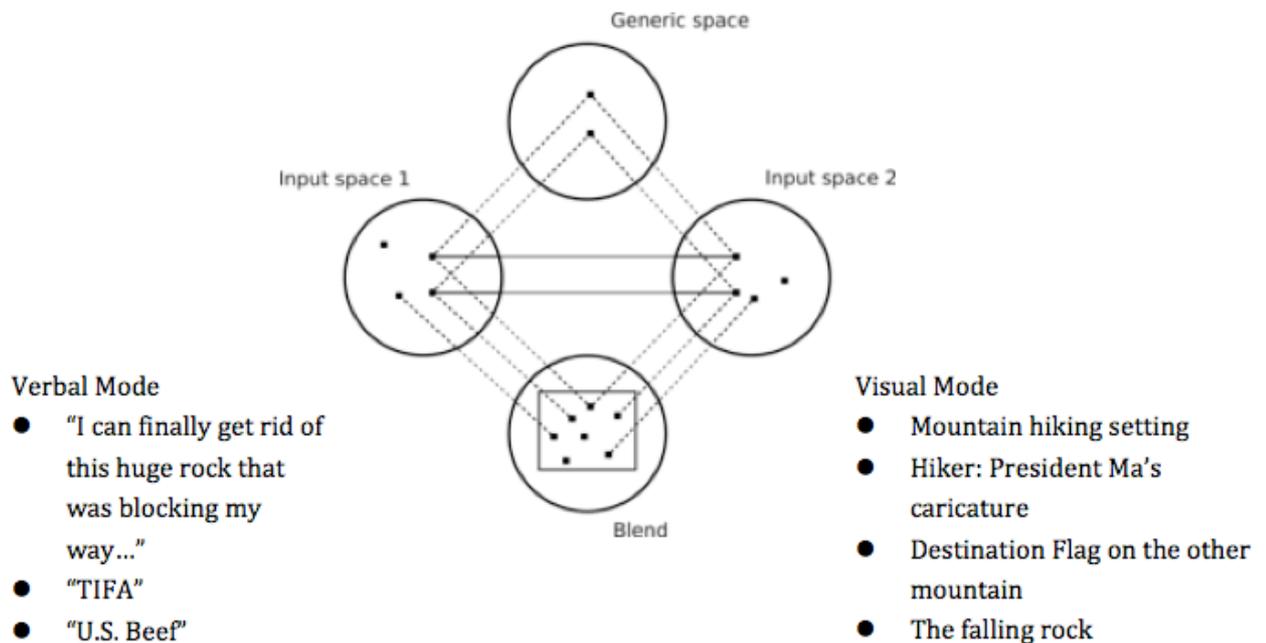


Figure 8. The Multimodal Level

As Figure 8 shows, viewing from the multimodal level of CIN, the verbal tag “I can finally get rid of this huge rock that was blocking my way” is in fact the crucial link that connects the conceptual, verbal, and visual mode based on our findings of audience interpretation. However, it remains difficult to emphasize the crucial role of verbal mode in such form of CIN.

In addition, our analysis indicates that distinct focus on metonymy and metaphor may lead to different multimodal representations, while example 1 represents the category of “related metonymic network” based on relevant metonymies, example 2 represents the other category of “diversified

metaphoric network” based on diversified metaphors.

However, such distinction is difficult to be found through the blending networks. Moreover, it is also hard to make cross-comparison through the data or present the general pattern and distribution of metaphor or metonymy in a multimodal corpus through the blended spaces as the above Figures show.

If we examine another example “version 1” of the songs (for detailed analysis of version 1, please see Chapter 5.2.1) from our data in poetry music conceptual integration networks, we can definitely construct one of the possible integration networks that connected input spaces of text/verbal mode and music/aural mode on the multimodal level as most previous studies (Zbikowski, 1999, 2002, 2009; Hsu & Su, 2014.) did.



Andante

mf

我 是 天 空 裡 的 一 片
I am sky in a

mf

雲， 偶 然 投 影 在 你 的 波 心，
cloud by chance reflect on your wave heart

mf

你 不 必 訝 異， 更 無 須 歡 喜； 在

f

Fa Re Do

Higher notes
(Similar melody and tempo)

Fa

Lower notes
(Similar melody and tempo)

Re Do

Version 1 (Paragraph 1)

Arpeggio(Ascending figures)

mp

你我相逢在黑夜的海
You I meet on dark night sea

上,
You on your (way) I on my way 向,

你記得也好, 最好是一忘
You remember better forget

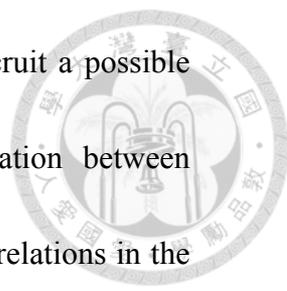
Version 1: Paragraph 2-Part 1

p *mf*

你記得也好, 最好是一忘
You remember better forget

掉, 在這交會時, 互放的光亮。
In this encounter exchanged glowing light 亮。

Version 1: Paragraph 2-Part 2



On the conceptual level of poetry music, we can also recruit a possible basic form of CIN that constructed the corresponding relation between “Narrator” and “Cloud” through input spaces to view the vital relations in the poetry lyrics as Figure 9 shows.

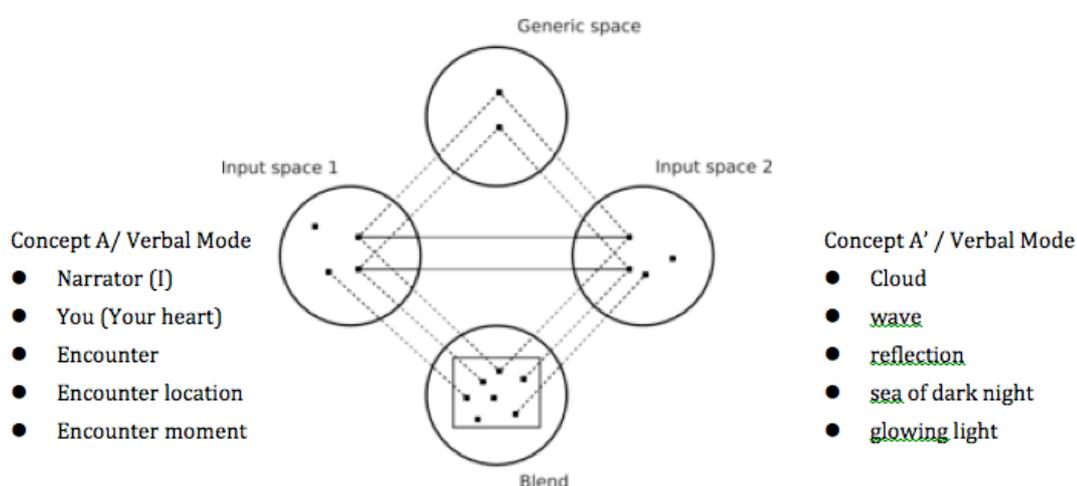


Figure 9. The Conceptual Level

In view of the relations between the pictorial effects and musical effects of poetry music, we can also recruit another possible integration network that binned input spaces of visual mode and aural mode on the multimodal level as Figure 10 shows.

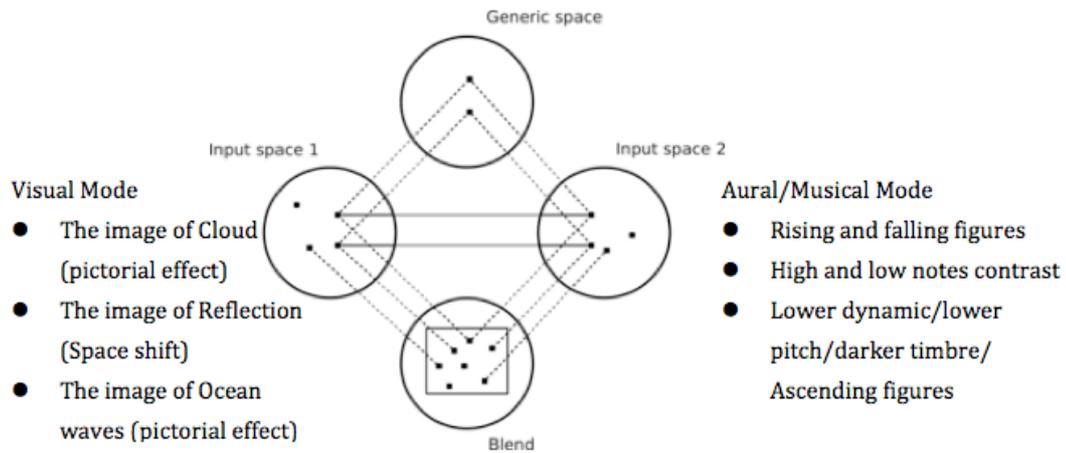


Figure 10. The Multimodal Level

However, if our research aim is to examine and compare four versions of songs based on the same poetry lyrics, it is difficult to apply blended integration networks to incorporate the visual, aural/musical, verbal, and conceptual level to make cross-comparisons among all. Blending networks may provide us a general pattern of each song with the consideration of specific input spaces, However, as previously mentioned, the input spaces can be recruited and constructed for different purposes. It remains problematic to determine how many input spaces should be included, what the principles are to govern the “selective projection” into the blend, and what the exact sequence of the vital relations construction process is. Thus we believe Blending Theory has its limitations in incorporating and analyzing visual, verbal, and musical mode in detail with its static drawing of blended spaces.

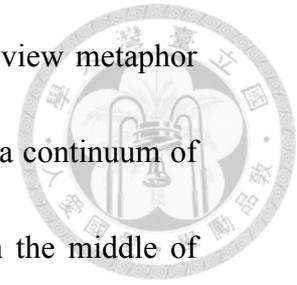
We will further clarify our proposal and elaborate how multimodal fusion model may provide more flexibility and possibilities in analyzing multimodal genre in section 2.4.



2.3.3 Special Focus on Metonymy and Metaphor

Previous studies show that metonymy and metaphor have long been viewed differently, while metonymy has often received less attention than metaphor. As Dirven (2002) pinpoints the fact that ‘Jakobson (1956) was the first to pay equal attention to both metonymy and metaphor.’ In the past few years, it remains problematic that most scholars have focused on the distinction between metaphor and metonymy. The domain approach (Barcelona 2002:211-15; Croft 2002) views ‘metonymy as a cognitive mechanism whereby one experiential domain is partially understood in terms of another experiential domain included in the same common experiential domain.’ However, demarcating clear boundaries between domains (or domain matrices) remain difficult and subjective. In addition, the traditional approach marks the distinction between metaphor and metonymy on the difference between similarity and congruity, which has also been found to be ambiguous and limiting (Piersman & Geeraerts, 2006:27-72). Recently, scholars have agreed that the traditional distinction between metaphor and

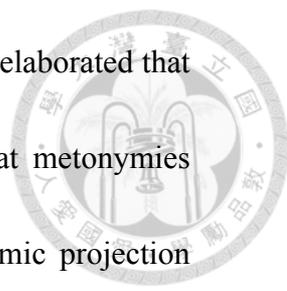
metonymy cannot be maintained, and that instead we should view metaphor and metonymy as ‘prototypical categories at the end points of a continuum of mapping processes,’ while metonymy-based metaphors are in the middle of the metonymy-metaphor continuum, providing an experiential motivation for a metaphor (Radden 2000: 105).



Under such viewpoint, we aim to pinpoint the significant role of metonymy and the interactive level of metaphor and metonymy by proposing the cognitive mechanism of multimodal fusion with the fundamental basis of a metonymic-metaphoric network, which will be elaborated in section 2.4.

2.3.3.1 Metonymy Projection in Metaphoric Blend

Conceptual Integration Theory views the “metonymy projection constraint” as one of the competing optimality principles/constraints on integration networks. In the integration networks, the “metonymy projection constraint” is defined as ‘when an element is projected from an input to the blend and a second element from that input is projected because of its metonymic link to the first, shorten the metonymic distance between them in the blend’ (Fauconnier and Turner 2000: p. 139), for example, in the integration networks of the canonical representation of death as “the Grim Reaper”, the skeleton becomes the bodily form of The Grim Reaper from



metonymy projection constraint. Fauconnier and Turner (2000) elaborated that ‘the metonymy projection constraint additionally specifies that metonymies get tighter under projection’. However, to satisfy the metonymic projection constraint is not a matter of blindly projecting metonymic links, since the internal integration of the blend provides opportunities for some acceptable metonymies but not for others. Tightening metonymies under projection typically optimizes Integration in the blend, since it helps build a tighter and more easily manipulated unit.

Blending theory demonstrates how metonymy is centrally important to conceptualization and human thought through integration networks (Coulson, 2003). However, the examples from Blending theory mainly focused on the conceptual level of analysis. In the multimodal level of analysis, metonymy and other tropes such as hyperbole, oxymoron, etc. has received relatively scant attention (Forceville, 2014). Therefore, this study aims to fill this gap by specifically examining the relationship between metonymy and metaphor in the multimodal genre to explore how metonymies and metaphors are interacted and represented through the combination of multimodalities and conceptual level.

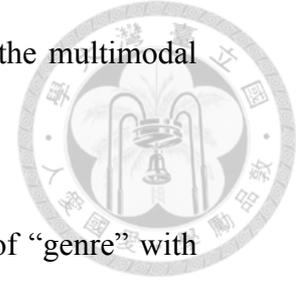
2.4 Our Proposal: Multimodal Fusion Model



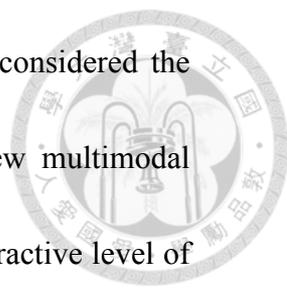
This study proposes the “multimodal fusion model” that includes the multimodal level (i.e. visual, aural, and verbal mode) and conceptual level to analyze the complex interaction between the multimodalities, the conceptual level, and interpretation. In other words, multimodal fusion model embraces the fact that the conceptual level, the verbal, visual, and aural modes are all encapsulated and integrated simultaneously in the real interpretation process of multimodal discourse thus we consider the conceptual level and the verbal mode, visual mode, or aural mode, naturally connected and overlapped with each other as shown in Figure 4, while Blending Theorists view different concepts or modes as in distinct input spaces connected with vital relations that are drawn with the blended lines. We acknowledge the significant integration of conceptual level as blending theorists suggested yet further assign a specifically important role of the conceptual level connecting with the verbal and visual/aural mode in multimodal fusion model.

Our analysis of the multimodal data indicates that multimodal cues including the verbal, visual, and aural cues each play significant parts in facilitating the interaction between different modes and the conceptual level, thus we stress that it is necessary to examine the multimodal cues specifically

in each mode in multimodal fusion model while interpreting the multimodal data.



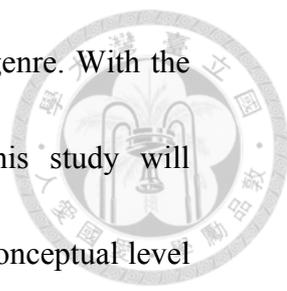
In multimodal fusion model, we stress the central role of “genre” with the consideration on distinct affordances and constraints of each medium as Forceville (2012, 2014) advocates that genre is the most significant interface between text and context by emphasizing the significant role of “genre” as the “key base” in our model (Figure 4). Through our analysis, we found the “contrasting visual/aural features” within the visual/aural mode and in comparison with the “verbal mode” from the data are often the key elements that influence how certain topic or concept is conceptualized in different versions of work. Thus in our model, viewing the contrasting features in each mode and in comparison with the other mode is the main step to interpret multimodal data. We believe “genre” plays an important part in multimodal fusion model to predetermine and dominate how the conceptual level, the verbal mode, and the visual/aural mode will be fused and intertwined. In multimodal fusion model, we emphasize how different mediums and modalities, including the verbal mode, visual, and aural mode, are fused with the conceptual level, leading to various combinations and representations in specific genre, such as in political cartoon and poetry music.



In this model, the “metonymic-metaphoric network” is considered the fundamental basis of multimodal fusion model since we view multimodal fusion model as the cognitive mechanism that works at the interactive level of multimodal metaphor and metonymy, hoping to “recognize the diversity of forms and semantic content that make up different tropes” (Gibbs and Colston, 2012, p.342).

In view of interpretation, we believe that ‘people may not process all figurative meaning in the same way, precisely because the kinds and forms of different tropes are sufficiently varied as to resist classification within a single theoretical framework’ (Gibbs and Colston, 2012, p. 3-4), thus we do not agree that a single theory or model can be generalized and account for how all aspects of figurative language is understood. Under such view, this study incorporates “audience responses” with our analysis to extend the discourse-internal analysis of figurative meaning to an interactional level of interpretation analysis on multimodality and metaphor/metonymy from the perspectives of analyzer and audience (See Chapter 4 and 6).

This study focuses on the data collected from political cartoon and poetry music, which combine verbal, visual, and aural modes that reflect Mandarin language and culture, to investigate how certain topic or concept has been



represented in various representations of specific multimodal genre. With the primary focus on multimodal metaphor and metonymy, this study will demonstrate how the verbal mode, visual/aural mode, and the conceptual level are interactively fused and interrelated in the genre of political cartoons and poetry music. We will illustrate how certain topic and concept (i.e. politics, emotion, space, etc.) are conceptualized and construed differently in the multimodal genre accompanied with the use of prominent multimodal cues (visual/aural/verbal cues) that facilitate multimodal realization and interpretation.

In this study, we also examine how specific effects, such as humor and irony in political cartoons or musical effects in poetry music, result from multimodal fusion in political cartoon. Our findings indicate that the audience interpret the data and perceive/sense the effects through the comparison and contrast between multimodal cues and prominent multimodal features.

In the following chapters, we will provide detailed examination of the data (i.e. political cartoons and musical scores) and cross-comparison analysis under the scope of multimodal fusion model.

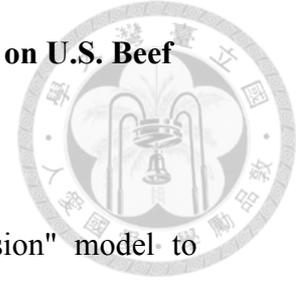
Chapter 3 Multimodal Fusion in Political Cartoon: Debates on U.S. Beef

Imports into Taiwan

In this chapter, this study proposes a "multimodal fusion" model to account for the cognitive mechanisms involving political cartoons with regard to U.S. beef import issues as reported in Taiwanese newspapers. In the following sections, data collection is provided in Section 3.1, data analysis with representative examples illustrating the process of multimodal fusion are presented in detail in Section 3.2, the analysis of conceptual scenario and conceptual metaphor is in Section 3.3, which is followed by further discussion in multimodal fusion and political stance in Section 3.4.

3.1 Data Collection: Multimodal Corpus of Political Cartoon

Following the historical timeline of U.S. beef imports in Taiwan as previously presented in Section 1.1, this study constructs a multimodal corpus based on a collection of 56 political cartoons comprised of visual and verbal modes with reference to the topic in question that were published in LT and UDN from October 2009 to July 2012. As summarized by Lin (2009), LT and UDN are distinct in political stances and national identifications. LT has the pro-independent political stance that reflects Taiwanese identity and localization (Tsai, 2006), aligning it with Taiwan's Pan-Green Coalition led by



the DPP and standing for a stronger position on Taiwanese sovereignty, whereas the political stance of the UDN is closer to the pan-blue coalition led by Kuomintang (KMT)¹⁴ (Lin, 1987; Kuo, 2002) of which President Ma is the chairman. Thus, UDN shows stronger support for President Ma and favors a softer policy and positive attitudes toward China (Chiang & Duann, 2007).

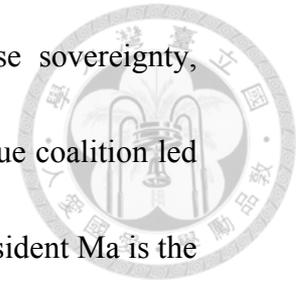


Table 1 presents the distribution of the political cartoons in the multimodal corpus. In total, our corpus included 36 political cartoons from LT and 20 political cartoons from UDN.

Table 1. The Distribution of Political Cartoons in the Multimodal Corpus

Year/Month	LT	UDN	Total
2009/10	3	3	6
2009/11	10	3	13
2009/12	5	4	9
2010/01	11	1	12
2010/02	1	1	2
2012/06	6	5	11
2012/07	0	3	3
Total	36	20	56

¹⁴ Kuomintang (KMT) is the ruling party and leader of the Pan-Blue Coalition. KMT is generally considered as the conservative party that advocates a supportive stance for Taiwan-China relationship.

In the following section, we will present data analysis from the representative examples of our multimodal corpus.



3.2 Data Analysis¹⁵

In this section, we will illustrate our analysis based on the theoretical basis of multimodal fusion model. We propose that the cognitive mechanism of multimodal fusion evolves from the metonymic-metaphoric network, which combines the visual, verbal, and conceptual level in the genre of political cartoon as shown in Figure 11. Through the multimodal fusion process, the metonymies and metaphors suggested by visual and verbal modes provide metonymic associations and metaphorical mappings with the main concept to present the conceptual scenario underlying the conceptual metaphor.

¹⁵ To remain the intact image of the cartoons without interruption, all the translation of the Mandarin Chinese verbal tag will be included in our analysis.

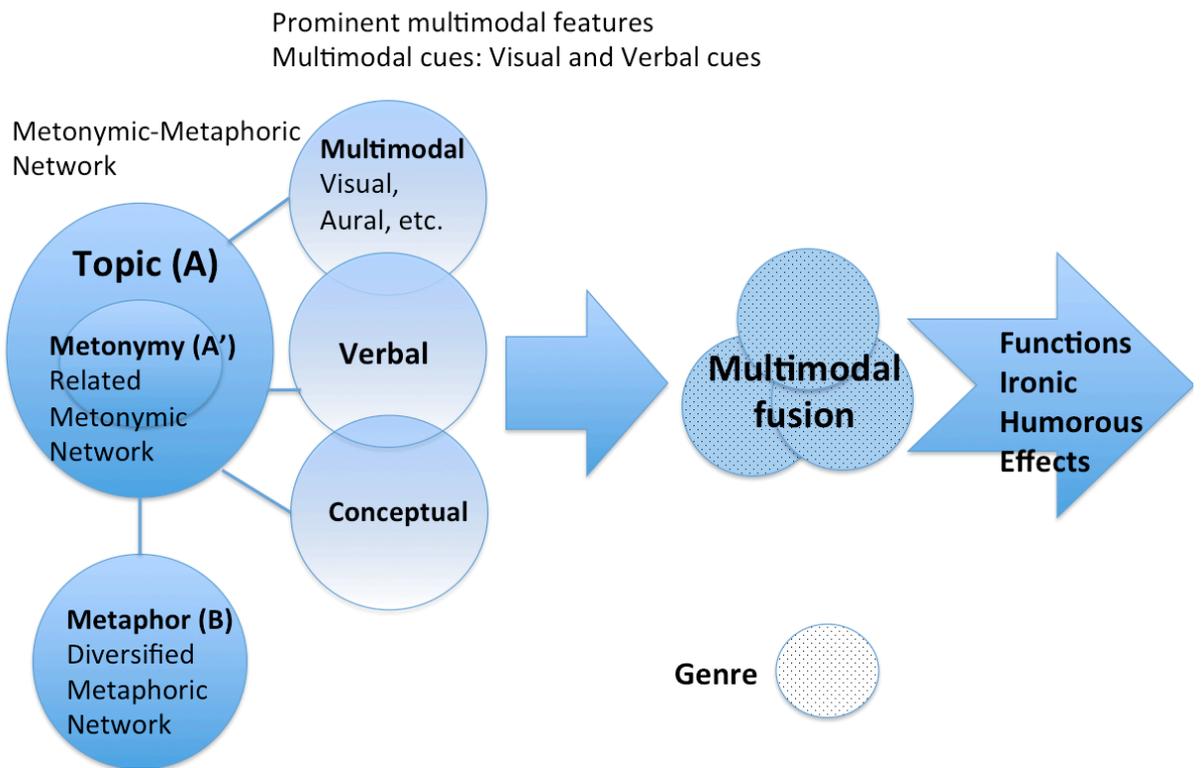
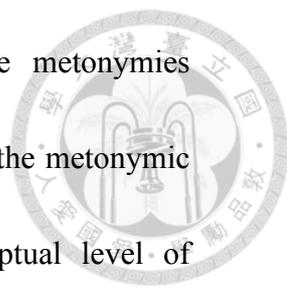


Figure 11. Multimodal Fusion in Political Cartoon

In this model, we suggest that the visual mode displays prominent visual features as visual cues to highlight the main focus, while the verbal mode serves as the link to connect visual details with conceptual scenarios. The integration of conceptual scenarios, visual modes, and verbal modes thus leads to multimodal fusion, creating the ironic and humorous effect and resulting in a critical message of political cartoons. Interestingly, we found the conceptual is often implicitly intertwined with the visual image combining the verbal tags. Through our analysis, we found a large number of metonymies and metaphors



representing the key elements of U.S. beef events, while metonymies demonstrate the part-for-whole features in efficiently building the metonymic association between visual mode, verbal mode, and conceptual level of multimodal fusion, and metaphors display the metaphorical mapping features in creatively linking visual and verbal context with conceptual scenarios. Thus, we propose a metonymic-metaphoric network of multimodal fusion starting with metonymies and metaphors representing the main topic that activate the conceptual scenario and then connecting the visual, verbal, and conceptual level based on the metonymic associations and metaphorical mapping. Furthermore, we discovered two major types of metonymic-metaphoric network of multimodal fusion on the basis of our data: (1) related metonymic network and (2) diversified metaphoric network.

For more details, we will provide a comprehensive analysis on how the visual mode, verbal mode, and the conceptual level interact and combine to convey the critical message with representative examples of “related metonymic network” and “diversified metaphoric network” in the next section. In each example, the main focus of the multimodal fusion process of each aspect, including the visual, verbal, and conceptual level, will be pinpointed, and the multimodal fusion concerning the main topic of U.S. beef event will

be summarized and highlighted in a box figure, which will be further elaborated with detailed analysis¹⁶.



3.2.1 Related Metonymic Network: Productive Visual Images

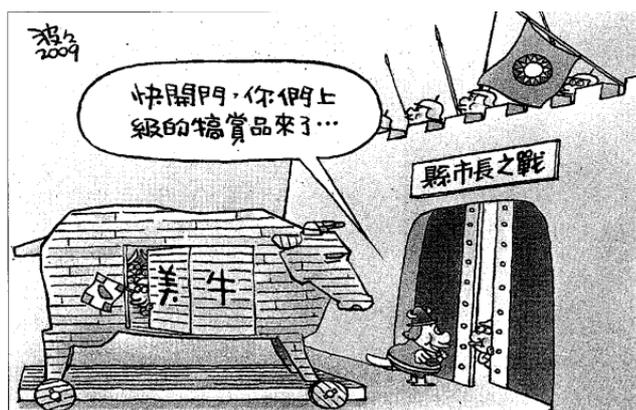
Through the detailed analysis of the multimodal corpus of 56 political cartoons, this study finds that the bull, productively used as metonymy for “U.S. beef event,” forms a basis for the related metonymic network and is extended to various kinds of bull-related activities and scenarios shown in the cartoons, including bullfighting, cowboys, bullchasing, bullriding, horse-related, and beef-related images, etc. In other words, the visual image is mainly composed by metonymies closely related to the key concept or keywords related to U.S. beef events, such as a “bull” for “U.S. beef”, thus the relevant bull-related or animal-related visual images are productively used to conceptualize U.S. beef events as a (i.e. powerful/playful) bull or a personified-bull. Consequently, the relevant metonymies in visual images contribute to our understanding of how “U.S. beef events” have been conceptualized. In addition, the visual images of a bull can represent “a

¹⁶ We understand there may be multiple explanations from different angles to analyze the following examples due to the complexity of the genre of political cartoon, however, this section mainly provides the interpretative analysis under the scope of multimodal fusion model to examine the one-on-many relationship of how the topic “U.S. beef imports” has been conceptualized as various metonymies/metaphors and represented through different combinations of verbal, visual, and conceptual level in the cartoons. We believe other possible interpretations flexibly exist under the scope of multimodal fusion model as well.

huge/powerful/poor/playful bull,” through prominent visual features, such as size or facial expression contrasts. Some bulls are even visually personified as emotional human beings that can express their own thoughts through verbal tags. The highly relevant metonymies and productive visual images marked with the verbal tags thus efficiently integrate the conceptual scenario based on readers’ related background knowledge.

In our data, this type of multimodal fusion accounts for 52 percent (29/56) of the corpus, which reveals its productivity and high frequency. In the following part, we will provide the representative examples of related metonymic network.

As shown in Example 1, the verbal tag “U.S. beef” on the body of the bull indicates the bull is the metonymy for the U.S. beef import event, while the visual image of the gate setting resembles the famous setting in Troy, activating the conceptual level of the “Trojan War” scenario. Specifically, the “wooden horse” in the original story of Trojan War is transformed into the “wooden bull” due to the relevance between the bull and beef. Therefore, the innovative metonymy “Trojan Bull” has become the main focus of the multimodal fusion of the U.S. beef event.



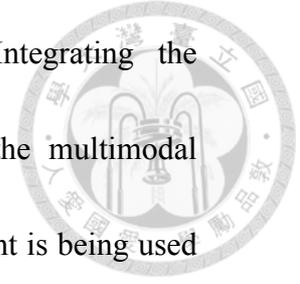
Example 1¹⁷ UDN (2009-11-16)

Visual (“wooden bull” metonymy + Trojan War settings+ Party Flags) +
 Verbal (“Open the door, your commander’s rewards are coming”, “Mayor election war”, “U.S. Beef”) +
 Conceptual (Trojan war scenario, POLITIC IS WAR)
 => U.S. beef event is Trojan Bull

As the huge wooden bull shows in the cartoon, the conceptual scenario of the Trojan horse is implicitly integrated with visual images and verbal details. To be more specific, the Trojan horse scenario is conceptually integrated with the visual shape of the Trojan bull and the verbal tag U.S. beef. The flag of the DPP (Democratic Progressive Party) waved by the person hiding in the wooden bull is the visual cue, while the person saying “Open the door, your

¹⁷ The quotation of the cartoons in this study is either permitted by the authors or permitted under the doctrine of fair use (Article 52 of Copyright Act: Within a reasonable scope, works that have been publicly released may be quoted where necessary for reports, comment, teaching, research, or other legitimate purpose). Each cartoon is marked with the source of newspaper and published date.

commander's rewards are coming" is the verbal cue. Integrating the conceptual scenario, visual cue, and verbal cue together, the multimodal fusion thus efficiently conceptualizes the U.S. beef import event is being used as DPP's political winning weapon/strategy to defeat the KMT in the upcoming mayor election.



In example 3, the bull marked with verbal tag of "U.S. beef" indicates the bull metonymy for U.S. beef imports. The conceptual scenario of a circus is reflected through the visual image of circus settings and the verbal tag of "2010 circus performance" on the top. However, the bull with the verbal tag of "U.S. beef" is attacking the man in suit with the verbal tag of "Ma", showing the bull as the metonymy for U.S. beef events. However, the bull with the verbal tag of "U.S. beef" is attacking the man in the suit with the verbal tag of "Ma." "Ma" stands for the animal "horse" in Mandarin Chinese, the surname of President "Ma," and also the first character "Ma" of circus (the literal meaning: horse-trick-show in Mandarin Chinese).



Example 3. LT (2010-01-01)

Visual (“bull” metonymy, “President Ma” caricature) +
 Verbal (“2010 circus performance”, “Ma”, “U.S. Beef”, “Happy new year to you all!”) +
 Conceptual (circus scenario, POLITICS IS GAME)
 => U.S. beef event is circus bull.

To interpret the circus scenario, we understand that the performing animals are normally tame and obedient, so the attacking-animal scene should be rarely seen. In addition, the bull generally does not perform in the circus. Thus, the ironic effect arises from the combination of the bull-attacking image and circus scenario, while the bull metonymy appears in this cartoon due to the relevance between the bull image and U.S. beef.

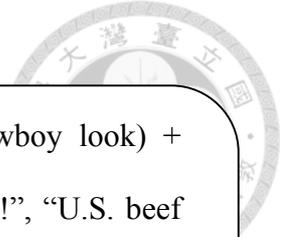
On the other hand, the audience seems happy with the clapping hands while the host says “Happy new year to you all” when seeing the

bull-attacking scene as the circus performance. The ironic effect results from the multimodal fusion of the circus scenario, which is integrated with the bull-attacking and audience-clapping image and the verbal cue “Happy new year to you all.” This strengthens the critical message that while the U.S. beef imports issue is not going as well as President Ma would like, most people are happy about it since they resist hazardous U.S. beef being imported into Taiwan. This example shows how critical messages can be elicited and emphasized through the ironic conflict created from the multimodal fusion process of visual, verbal, and conceptual level in political cartoon.

In example 4, the visual image of the bull’s tail on fire marked with the verbal tag “U.S. beef event” indicates the fire-tail metonymy for U.S. beef event. The cowboy outfits of the man resembling President Ma and the cowboy hat with KMT symbol depicts the conceptual scenario of cowboy.



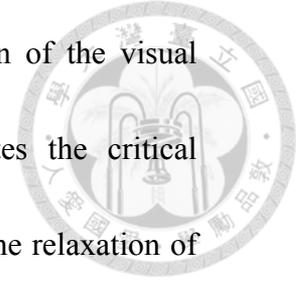
Example 4. LT (2012-06-04)



Visual (“fire-tail” metonymy, angry bull, President Ma’s cowboy look) +
Verbal (“U.S. beef has its *butt on fire* and I’m going crazy too!”, “U.S. beef event”)+
Conceptual (cowboy scenario, POLITICS IS GAME)
=> U.S. beef event is fire.

Although the conceptual scenario of the cowboy is often understood as an energetic cool cow-rider, the facial expression of Ma looks tensed and his leg is tied with bandages, which are not commonly associated with the cowboy frame. In addition, as the verbal tag says “the KMT opposing group is not getting away with this! U.S. beef has its *butt on fire* and I’m going crazy too!” . In fact, the colloquial term “butt on fire” in Mandarin Chinese means having urgent problem unsolved while running out of time, thus President Ma expresses his frustration and anxiety while riding a bull with “butts on fire”. Under such circumstances, the fire marked with the verbal tag “U.S. beef import” will not only make the bull angry but also accelerate its speed in danger, thus U.S. beef issue is conceptualized as the main reason causing President Ma’s current dangerous and chaotic situation.

The ironic conflict reflected from the multimodal fusion of the visual cues, verbal cues, and conceptual scenario clearly indicates the critical message that President Ma is facing a serious situation since the relaxation of U.S. beef import regulations that he promoted has not been passed and is still being opposed and boycotted by the legislators in the Legislative Yuan during this period. This cartoon shows how the critical messages can be elicited and interpreted through the ironic effect of the multimodal fusion.



The following Example 5 is one of the productive examples where U.S. beef issue is represented by a personified bull. As the visual image shows, the personified bull with the verbal tag of “U.S. beef” on his body is the metonymy for the U.S. beef event. The personified bull is standing next to the pan-blue party (pro-government stance) and pan-green party (anti-government stance) on a small island with bones on the ground as they all seem struggled to survive without enough food.



Example 5. UDN (2012-07-07)

Visual (“bull” metonymy, DPP and KMT personified metonymies, rescue ship metonymy) +

Verbal (“U.S. Beef”, “CODEX”, “10 ppb”, “thanks god”, “political parties acronyms” 朝/野) +

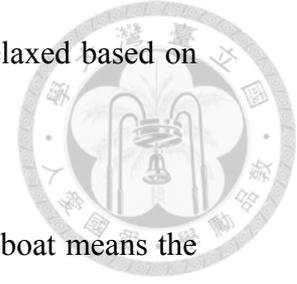
Conceptual (isolated island scenario, POLITICS IS SURVIVAL STRUGGLE)

=> U.S. beef event is personified bull.

The visual details reveal the conceptual scenario of an isolated island, which stresses both parties are struggling against the harsh dilemma of opposing U.S. beef import regulations. However, the bull waves a sheet toward a rescue ship marked with the verbal cues, “CODEX”¹⁸ and “10 ppb”, and expresses his gratefulness by the verbal tag saying “thank the gods....” with a tear falling

¹⁸ CODEX is the abbreviated form of the Codex Alimentarius Commission. *Codex Alimentarius* (Latin for "Book of Food"), according to Wikipedia, is a “collection of internationally recognized standards, codes of practice, guidelines and other recommendations relating to foods, food production and food safety.”

from his eye, indicating that U.S beef imports law has been relaxed based on the level of 10 ppb CODEX requirements.



In an isolated island scenario, the appearance of a rescue boat means the trapped victims will be saved. The multimodal fusion of the visual cues, verbal cues, and conceptual scenario conceptualizes the U.S. beef issue as a desperate bull being rescued, revealing the critical message that the rescue boat represents a new beef examination standard that allows the relaxation of U.S. beef import regulations, which is a relief to the pan-blue and pan-green legislators after the long-term boycott. Interestingly, the main character in conceptual scenario of isolated islands, such as the famous Robinson Crusoe is usually a human being. However, this bull is personified as an emotional character being rescued and stands in the center of the cartoon. Examining the isolated island scenario within the visual context of the bull's facial expression, central position, posture, and verbal expression, the bull creates an ironic and humorous effect that easily attracts most readers' attention. We believe such use of multimodal fusion provides the reader an innovative, humorous, and efficient way of understanding the event.

In example 6, as the visual image shows, the head of a bull and a leg are hanging on the office wall as hunting trophies, representing the metonymy for

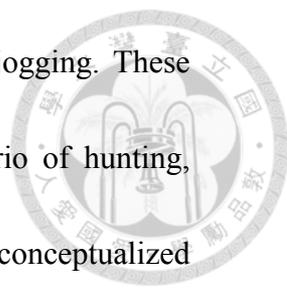
U.S. beef and metonymy for President Ma. The man with the verbal tags of “Ko” (on the shirt) and DPP “party leader” (on the desk) is sitting with content smile in his office desk with a DPP flag standing beside, whereas at the other side of the window, another man resembling President Ma, who appears to have lost one of his legs, looks crippled and unhappy. The contrast between the status and facial expressions of these two men and the alignment of the two trophies implies a paralleled relationship between them and the trophies.



Example 6 UDN (2012-06-16)

Visual (“bull” metonymy, “Ma’s leg” metonymy, Ko’s smiling face, Ma’s injured look, DPP’s flag, office setting) +
 Verbal (“party leader”, “Ko”) +
 Conceptual (hunting scenario, POLITICS IS GAME)
 => U.S. beef event is hunting trophy.

One of the trophies hang being portrayed as a “leg” can be attributed to President Ma’s nickname “Brave-leg Ma”, which is a Taiwanese colloquial



phrase praising how active and strong Ma is since he loves jogging. These visual and verbal cues, integrated with the conceptual scenario of hunting, revealing the critical message that the U.S. beef issue has been conceptualized as the “victory trophy” that the DPP leader won by defeating President Ma. In other words, DPP used the U.S. beef import event to successfully oppose President Ma’s policy on U.S. beef imports and also bring down Ma’s support ratings. This example vividly demonstrates how contrasting visual features through position and facial expression help to highlight the critical message in multimodal fusion. As our analysis shows, we believe it is important to have background knowledge to interpret the accurate meanings of the verbal mode, such as when colloquial phrases or idioms are used, to fully grasp how the multimodal fusion works in political cartoons.

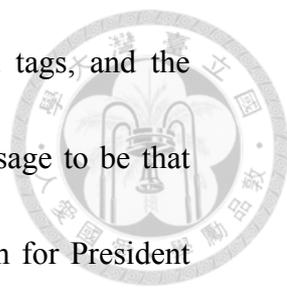
In the following, example 7 demonstrates the productive use of multimodal fusion. In this cartoon, the arrow-attacking war scenario and the head of the D.O.H. official, Yang (who supports U.S. beef imports) visually fused with the bull’s horns and a golden armored bull indicate the “war bull” as the metonymy for Yang promoting U.S. beef imports. The verbal tag, “Don’t worry! My skin is tough enough to protect you all,” is combined and integrated with the conceptual scenario of “war animals,” which are often

horses that carry warriors in battle. However, the horse transforms into a bull, which is similar to the “Trojan Bull” in example 1, due to the relevance between the bull and U.S. beef. The protective feature of the war animal scenario is highlighted with the arrows being shot at the golden bull with one of the arrows on the right being marked “U.S. beef imports,” thus indicating that the U.S. beef import issue has been conceptualized as attacking arrows being shot at the government.



Example 7. LT (2009-11-01)

Visual (“war bull” metonymy, visual fusion of the bull with Yang’s head, shooting arrows, President Ma’s and Vice President Wu’s caricatures) +
 Verbal (“Don’t worry! My skin is hard enough to protect you all”, “U.S. beef imports”) + Conceptual (war horse scenario, POLITICS IS WAR)
 => U.S. beef event is shooting arrow.



Through the multimodal fusion of visual images, verbal tags, and the conceptual frame of this cartoon, we perceive the critical message to be that the D.O.H. official, Yang, is taking all the blame and criticism for President Ma and Vice President Wu on health issues, including that on U.S. beef imports. The contrasting facial expressions of Yang's smile and the frightening look of others ironically expose Ma and Wu's negative attitudes in terms of avoiding responsibility and the ignorant obedience of the D.O.H. official. The contrast in Yang's larger head and Ma and Wu's smaller heads also indicates that Yang has become the main target for criticism, while Ma and Wu timidly hide behind him to avoid taking responsibility. This example demonstrates the use of multimodal fusion with emphasis on contrasting visual features, such as in depicted size and facial expressions, as prominent visual cues that help to highlight the critical message.

In the next section, we will provide detailed analysis of the representative examples in the type of diversified metaphoric network, showing how prominent verbal cues play an important part in the multimodal fusion process.

3.2.2 Diversified Metaphoric Network: Prominent Verbal Cues

In this network type, the visual image is mainly composed of novel metaphors that are implicitly related to the key concept or keywords

concerning U.S. beef events, thus U.S. beef events are conceptualized as different objects, such as a rock, bomb, or net, etc., in diverse scenarios.

Unlike previous types, the visual images are implicitly associated with the keywords of U.S. beef event, which are more diversified, thus a wide variety of scenarios, such as bombing, fishing, rolling stones, being in offices or on highways, etc. comprise the visual images in this type of metaphoric network.

In these cartoons, the U.S. beef event is conceptualized as an object (i.e. a rock, a bomb, a net, etc.), creating the novel metaphor. In this type, the verbal mode is the crucial link that establishes the relationship between the conceptual scenario and visual details, highlights the main focus, and combines the critical message underlying the verbal, conceptual, and visual modes. The verbal mode plays an important role in this type of multimodal fusion since the verbal tags are the prominent cues that enable readers to efficiently build the connection between the metaphors and the U.S. beef event and associate the verbal and visual contexts with the conceptual scenarios. In the following section, we will illustrate several examples of diversified metaphoric networks.

In example 2, as the visual image shows, the rock marked with “U.S. beef” is the metaphor for U.S. beef events, while the conceptual scenario is

that of a mountain road. However, only a part of the mountain road is shown in the foreground, whereas the other mountain with the flag “TIFA”¹⁹ appears far in the background. Such positioning of visual features should be further examined in the context of verbal tags and conceptual level. In this cartoon, the man resembling President Ma is pushing a huge rock with the verbal tag “U.S. beef” off the cliff, saying, “I can finally get rid of this huge rock that was blocking my way...”. We understand that in the conceptual scenario of the mountain road, most mountain-climbers/hikers aim to reach the top of the mountain, thus the man in the cartoon is pushing the obstacle away to reach his destination, which is conceptualized as the “TIFA” flag in the background. However, it is impossible and unrealistic to reach the top of the other mountain under such circumstances.



Example 2. UDN (2012-07-27)

¹⁹ “TIFA” is the abbreviation for the “Trade and Investment Framework Agreement.” The Taiwan-U.S. TIFA serves as a primary platform for bilateral high-level economic and trade consultations between Taiwan and the United States.

Visual (“rock” metaphor, ²⁰president Ma’s caricature, mountain road setting, destination flag)+

Verbal (“U.S. Beef”, “TIFA”, “I can finally get rid of this huge rock that was blocking my way...”)+

Conceptual (mountain road scenario, POLITICS IS GAME/OBJECT)

⇒ U.S. beef event is falling rock

Interpreting the multimodal fusion of the visual details, verbal clues, and conceptual scenario, we understand that President Ma is aiming to sign the TIFA with the U.S. government by agreeing to sign the new protocol on U.S. beef imports (with more relaxation on the previous ban). With the cues of verbal mode, the critical message implicates that Ma’s strategy is unrealistic and in vain. As shown in example 7, we believe verbal mode is crucial in connecting the relations between the conceptual scenario and visual details of the multimodal fusion, which enables the readers to comprehend the whole picture of the critical message and ironic effect in the cartoon.

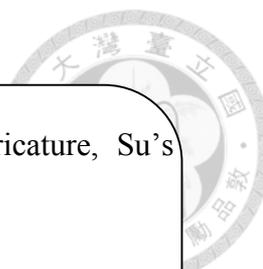
In example 8, the bomb marked with the verbal tag, “U.S. beef chaos,” is the bomb metaphor for the U.S. beef event. The cartoon illustrates the conceptual scenario of disarming a bomb, which is a highly dangerous and

²⁰ Regarding the “rock” metaphor, we understand that the rock as obstacle is conceptualized through JOURNEY metaphor in a broad sense. However, this does not affect the results in our analysis since our main concern is to examine how U.S. beef event has been represented and conceptualized in each cartoon, while it has been conceptualized as the “rock” verbally and visually in this cartoon.

life-threatening process. However, the visual office setting of the “National Security Council” suggested by the verbal tag creates an ironic effect at first sight. Moreover, President Ma looks nervous and clumsy with the huge tool in his hand as he is disarming the bomb named “U.S. beef chaos,” while a government official hiding behind the desk with the verbal tag “National Security Council” says, “President, have you disarmed the bomb yet?” when the huge bomb is already on fire and may explode very soon. This cartoon implicates the U.S. beef issue has been conceptualized as an exploding bomb that is uncontrollable and dangerous, which will threaten national security on the basis of public health and safety.



Example 8 LT (2010-01-07)



Visual (“bomb” metaphor, office setting, President Ma’s caricature, Su’s hiding figure, the nervous dog) +

Verbal (“U.S. beef chaos”, “President, have you disarmed the bomb yet?”, “National Security Council”) +

Conceptual (bomb disarming scenario, POLITICS IS WAR/OBJECT)

=> U.S. beef event is an exploding bomb.

Incorporating the visual, verbal, and conceptual scenario of multimodal fusion in this cartoon, the critical message is that in spite of knowing the uncontrollable danger caused by U.S. beef imports, president Ma still insists on signing the protocol and believes he can solve the problem by using ineffective protective strategies. The ironic effect highlighted through the verbal tags critically strengthens the ignorance and weakness of how the government handles the U.S. beef import issue. Similarly, the ironic effect of the critical message is also illustrated in the following example.

In example 9, the three monkeys sitting on the bench named “the Ma administration” with the verbal tags of “Ma”, “Wu”, and “Chin” representing from left to right and the caricatures of President Ma, Vice President Wu, and Secretary Chin conveying the monkey metaphor for the Ma administration. The visual image of the three monkeys covering their ears, nose, and eyes

relate to the famous “three wise monkeys” originating from ancient Chinese Confucius culture. Viewing from the conceptual level, the three-monkey scenario reminds people to face the truth, stay rational, and not to listen, smell, or see anything without a clear mind. The verbal information in this cartoon is the key element to interpreting the critical message, the ironic, and humorous effect.



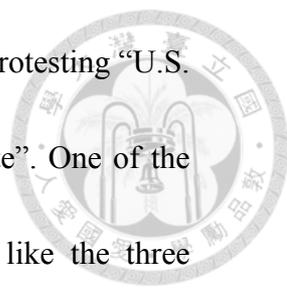
Example 9 LT(2010-01-26)

Visual (“Three-monkey” metaphor, Ma’s fusion figure covering his ear, Wu’s fusion figure covering his mouth, Chin’s fusion figure covering his eyes, Protestors raising placards) +

Verbal (“government Ma”, “Ma”, “Wu”, “Chin”, “U.S. beef chaos”, “vaccine problems”, “high unemployment rate”) +

Conceptual (Three-monkey scenario, POLITICS IS (ANIMAL) BEHAVIOR)

=> Government Ma is three monkeys.



On the right of the cartoon, there are two men raising placards protesting “U.S. beef chaos”, “vaccine problems”, and “high unemployment rate”. One of the men pointing at the three monkeys says, “why do you act like the three monkeys when facing public criticism?” These verbal tags provide the clues to understand the ironic and critical message in this cartoon. Integrating the multimodal fusion of visual images, verbal details, and the conceptual scenario, the ironic conflict pinpoints the fact that these three monkeys who refuse to face the truth are the government authority that has the power and resources to solve the problems. Moreover, the critical message underlying the multimodal fusion is that although the government is encountering serious issues resulting from their problematic policy, they still choose to ignore the facts and are reluctant to solve the problems.

In the following example 10, the broken net marked with the verbal tag “U.S. beef chaos” at the bottom is a metaphor for the U.S. beef event. Meanwhile, President Ma stands at the top saying “Little Su, hurry! Get the broken net fixed/mended for me!” to the caricature of government official Su, who looks frustrated and poor with his bare feet as he strives to fix the broken net. The verbal tag on top of his head says “if you don’t say what is bothering you...” enclosed by musical symbols are from the lyrics of a sorrowful

Taiwanese song. The conceptual scenario of net-fixing activity finds its origin in the Taiwanese idiom “fix the broken net”, which refers to fixing serious problems that are hard to solve.



Example 10 LT (2010-01-04)

Visual (“net” metaphor, Su’s caricature, president Ma’s caricature, broken net image, music symbols) +

Verbal (“U.S. chaos”, “Little Su, hurry! Get the broken net fixed/mended for me!”) +

Conceptual (broken net-fixing scenario, POLITICS IS OBJECT)

=> U.S. beef event is a broken net.

Therefore, in the net-fixing scenario, the U.S. beef import issue is conceptualized as a broken net that is difficult to fix. In other words, the problem is difficult to resolve using the government’s ineffective strategies. In

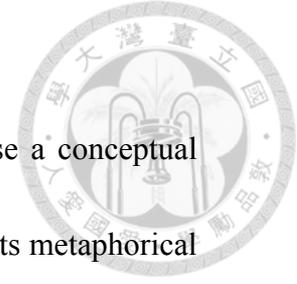
this cartoon, the ironic effect is created by the contrast between Ma's ignorant order and Su's miserable status of fixing the broken net.



Like the previous examples, this cartoon also demonstrates how multimodal fusion is processed in conceptual, visual, and verbal modes in political cartoons. Without the cultural background knowledge of the conceptual scenario and verbal cues, we would hardly be able to comprehend the meaning emerging from the visual and verbal details. As the above examples show, it is necessary to first identify the metonymy or metaphor for the main topic of U.S. beef event and then recognize the conceptual scenario. While interpreting the multimodal fusion, having access to background knowledge and cultural reference of the multimodal context is also crucial to achieving full understanding of the critical message, the irony, and humor conveyed.

In view of the conceptual level in political cartoons, we believe it is necessary to identify the conceptual scenarios and conceptual metaphors of all the cartoons in our multimodal corpus to have a better understanding of the general pattern and distribution of the conceptual scenario and metaphorical content in the data from our multimodal corpus, which will be discussed in the next section.

3.3 Conceptual Scenario and Conceptual Metaphor



In each political cartoon, the multimodal details comprise a conceptual scenario of the source domain. Each scenario is involved with its metaphorical mappings through various elements, such as people, objects, roles, relations, and status. Based on the scenario identification through the prominent multimodal features in each political cartoon, we can identify the major conceptual metaphor by analyzing the conceptual content using the “POLITICS IS Y” template to show which aspects of POLITICS, while the target domain are mostly mapped by specifying Y, the source domain, in each political cartoon. As a result, we found the conceptual metaphors, including POLITICS IS GAME and POLITICS IS WAR, are most frequently used to conceptualize the general aspects of the political issue of U.S. beef imports in both newspapers. Through the analysis of the multimodal corpus, we discovered that the most often identified scenarios, which account for 31% in LT and 35% in UDN, are “bull/animal-related” activities of the GAME domain, such as bullfighting, bull riding, bull-chasing, cowboys, and circus scenarios. Tables 2 and 3 list the most frequently identified scenarios in LT and UDN in the corpus.



Table 2. Bull/Animal-related Scenarios of the GAME Domain in LT

Scenario	Incidence
Bullfighting	2
Bull riding	2
Racing	1
Carriage	1
Cowboys	1
Horseback jumping	1
Bull chasing	1
Circus	1
Strip show	1
(Total/Percentage)	11 (31%)

Table 3. Bull/Animal-related Scenarios of GAME Domain in UDN

Scenario	Incidence
Bullfighting	4
Cowboys	1
Gambling	1
Hunting	1
(Total/Percentage)	7 (35%)

In these scenarios, “bull” image is the most-frequently used metonymy for the U.S. beef imports event, while other characters, including President Ma or government officials, play different roles in each scenario and interact with the bull (e.g., by being attacked by the bull, riding the bull, playing with the bull).



Focusing on the topic of U.S. beef imports, these political cartoons conceptualize the complex political debates as novel representations of human–bull interactions or entertainment to convey critical messages with irony and humor. The high frequency and productive use of bull metonymy and relevant scenarios provide the fundamental basis of “related metonymic network” in multimodal fusion. Overall, the “intensive” interactions between metonymies and metaphors shown in our study indicate the metonymies and metaphors are interwoven in the creation and realization of political cartoons to achieve the humor or ironic effect serving as the function of political cartoons.

The second most often identified conceptual scenarios, which account for 14% in LT and 10% in UDN, are relevant to “war” such as bombing, arrow-attacking, tank-attacking, post-war meeting, and Trojan horse scenarios. These metaphorical mappings of the defending side, attacking side, retreating side, or negotiating side in war-related scenarios clearly show different stances

and the conflict between the government and other political parties. The victory or defeat status in war scenarios also indicates the consequential effect of the issue of U.S. beef imports. Tables 4 and 5 list the identified scenarios relevant to war in LT and UDN.

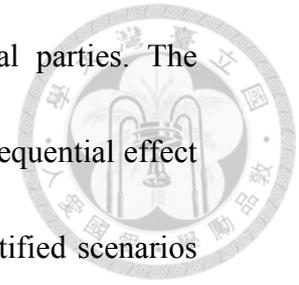
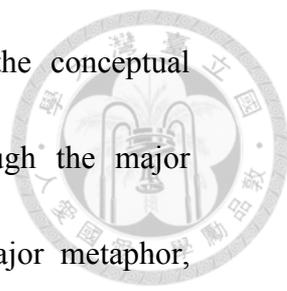


Table 4. War-related Scenarios in LT

Scenario	Number
Bombing	3
Arch-attacking	1
Post-war meeting	1
(Total/Percentage)	5 / 14%

Table 5. War-related Scenarios in UDN

Scenario	Number
Tank-attacking	1
Trojan war	1
(Total/ Percentage)	2 / 10%

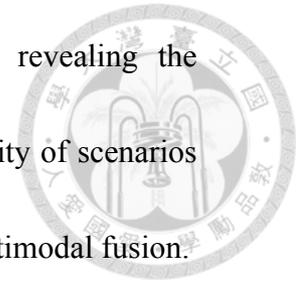


Through the scenario identification process, we analyzed the conceptual content to understand how politics is conceptualized through the major metaphors in political cartoons. We generalized that the major metaphor, POLITICS IS GAME, underlies the most often identified bull/animal-related scenarios, whereas the major metaphor, POLITICS IS WAR, underlies war-related scenarios. The result indicates that the political issue of U.S. beef imports is mainly conceptualized through the GAME and WAR domains in political cartoons in LT and UDN.

In our multimodal corpus, POLITICS IS A GAME accounts for 32% of the database, while POLITICS IS WAR is the second most often used metaphor (13%). In other words, these two major metaphors account for 45% of the database. This result indicates that the political issue of US beef imports is mainly conceptualized through the GAME and WAR domains in political cartoons in LT and UDN.

Besides the war-related scenarios, we also found a wide variety of conceptual scenarios including an isolated island scenario, office scenario, mountain road scenario, three monkeys scenario, broken net-fixing scenario, disaster scenario, highway scenario, and emperor scenario, etc. in the other cartoons. In these cartoons, the conceptual scenarios were mainly comprised

of novel metaphors, which are creative yet unproductive, revealing the innovative feature of political cartoon. We believe such diversity of scenarios provides the basis for “Diversified metaphoric network” in multimodal fusion.



3.4 Multimodal Fusion and Political Stance

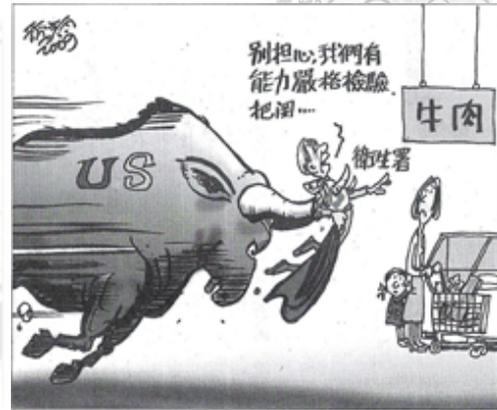
In this section, we will show how distinct political stance leads to subtle differences in multimodal fusions and critical messages within similar conceptual scenarios.

3.4.1 Different Targets of Bull Metonymy in the Same Scenario

In our corpus, the bullfighting scenario is the conceptual scenario that appears most frequently in the GAME domain in LT and UDN. As examples 11 and 12 show, the image of a bullfighter wounded by a bull appears similarly at first sight. However, the bull and bull-fighter metonymy in example 11 and 12 each stands for different things.



Example 11. (LT 2010-01-02)



Example 12. (UDN 2009-10-25)

Example 11:

Visual (“bull” metonymy, “bull-fighter” metonymy, Ma’s caricature, the bullfighter’s falling clothes image) +

Verbal (“Ma”, “Authoritarian leadership”, “KMT legislative representative with regard to the issues of beef imports”) +

Conceptual (bull-fighting scenario, POLITICS IS GAME)

=> KMT legislative representative is the attacking bull.

Example 12:

Visual (“bull” metonymy, “bull-fighter” metonymy, DOH official’s caricature, Mother’s and Daughter’s caricatures, supermarket setting) +

Verbal (“US”, “DOH”, “beef”, “Don’t worry, we will carryout strict inspections...”) +

Conceptual (bull-fighting scenario, POLITICS IS GAME)

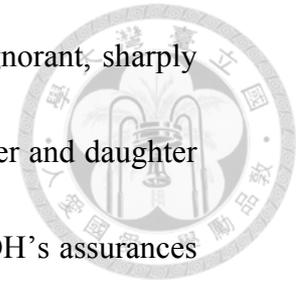
=> U.S. beef event is the attacking bull.



In example 11, the cartoon illustrates that President Ma is under serious attack from the KMT, with the verbal tag “authoritarian leadership” on the falling clothes of the bullfighter and the other verbal tag on the bull “KMT legislative representatives with regard to the issue of beef imports” show that KMT legislators, who are supposed to be supporting President Ma, chose to oppose U.S. beef imports instead. The KMT legislators are depicted as the destructive and powerful bull that stabs President Ma’s back with its horn, warning him that his “authoritarian leadership” may lead to fatal results. However, in example 12, the targets are the DOH and the U.S. The cartoon illustrates that the DOH official (the bullfighter) is being attacked by the bull of the U.S., while a DOH official still claims, “Don’t worry, we will carryout strict inspections (of U.S. beef imports)” in order to put consumers at ease despite being under serious attack from the powerful U.S.

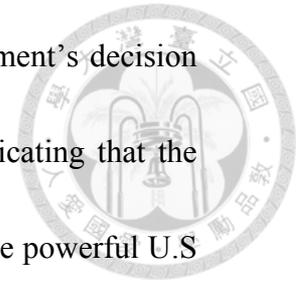
As mentioned earlier, the comparison between facial expressions is a prominent feature in political cartoons. In example 11, President Ma, as the bullfighter, is wounded and looks shocked and frightened, while the attacking bull looks firm and determined with its eyes closed. This contrast shows that the KMT legislative representative is powerful and steadfast in opposing the issue of U.S. beef imports and fighting against President Ma. In example 12,

the facial expression of the DOH official, looks painful yet ignorant, sharply contrasts with the confused and frightened look from the mother and daughter shopping in the beef section. This contrast implies that the DOH's assurances are weak and ineffective at restricting potentially hazardous U.S. beef from entering the country.



The size comparison arises from the enormous size of the bull and the small bullfighter, which is a recurring image in our corpus used to indicate the unequal power relationship and threatening status of the U.S. beef import issue. In example 11, the comparison between the enormous and powerful bull and relatively small bullfighter illustrates that the opposing KMT and Ma have a serious difference of opinion on U.S. beef imports, whereas the injury scene represents the conflict resulting from this divergence. The underlying critical message is that LT not only attacks Ma's incompetent leadership but also emphasizes Ma's policy on U.S. beef imports being so problematic that he cannot even gain support from his own political party. However, in example 12, the relative sizes of the huge bull and small bullfighter represent the unequal power relationship between the U.S. and Taiwanese government. In our corpus, this power relationship is mainly highlighted in the UDN. From the perspective of political stance, since the UDN is closer to the

government's political stance, it sympathizes with the government's decision and its criticism by emphasizing the power of the U.S., implicating that the government is in fact helpless, being forced or threatened by the powerful U.S. on the U.S. beef imports.



The above examples demonstrate how different targets of the same metonymy (i.e. bull and bull-fighter metonymy) in the same scenario may lead to different representations and critical messages in politically distinct newspapers. Our analysis shows such subtle differences can be found in the contrast between visually prominent features, such as facial expressions and size.

3.4.2 The Personified Bull in Different Scenarios

The results show that the personified bull is a recurring element in related metonymic networks of multimodal fusion in our corpus. The personified bull image has been productively represented as a rich, superior, dominant person, extending from the prototypical image of the powerful bull in some scenarios. Although the personified bull metonymies all represent U.S. beef, distinct political stances and critical messages can be elicited through the interactions between the personified bull and the other characters in the cartoon, as examples 13 and 14 show.



Example 13. LT (2012-06-10)



Example 14. UDN (2009-10-27)

Example 13:

Visual (the personified “bull” metonymy, Ma’s caricature, the coachman image) +

Verbal (“U.S. beef”, “KMT legislators”) +

Conceptual (carriage scenario, POLITICS IS GAME)

=> U.S. beef is the manipulator

Example 14:

Visual (the personified “bull” metonymy, DOH official’s caricature, Ma’s caricature, the “consumer” metonymy, casino setting) +

Verbal (“U.S. beef”, “DOH”, “consumer”, “military procurements, diplomacy, and economics”, “To be honest, we hardly have any chips...”) +

Conceptual (casino scenario, POLITICS IS GAME)

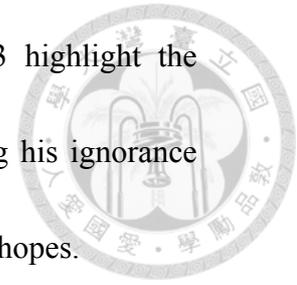
=> U.S. beef is the dealer



In the carriage scenario in example 13, the visual and verbal details show the personified bull “passenger” sitting comfortably on the couch marked “U.S. beef”, smoking a cigar in nice clothes, being served with his legs crossed, representing the typical image of a rich and powerful tycoon. The caricature of the coachman indicates that the target is President Ma, who looks bossy, mean, and emotionless while waving the whip to control the carriage, whereas the exhausted and suffering animals pulling the carriage are personified as KMT legislators. In example 14, the personification of the bull is also the main element of the scenario. However, President Ma and the DOH are represented as passive and helpless participants in this cartoon, whereas the bull is personified as the dominant dealer who has piles of chips marked with the verbal tag of “military procurements, diplomacy, and economics” in front of him.

In example 13, the comparisons between the different facial expressions and contrasting states of the rich passenger, bossy coachman, and tired carriage animals imply the complex power relationship between the U.S., President Ma, and KMT legislators. Under pressure from the powerful U.S., President Ma actively controls the legislative issues by forcing KMT legislators to cooperate with him, hoping to pass the amendment and relax the

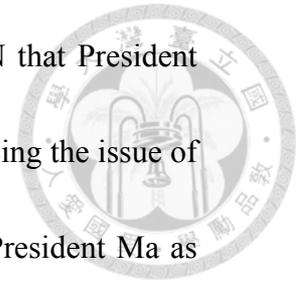
ban on U.S. beef imports. The comparisons in example 13 highlight the critical response to the firm stance of President Ma, revealing his ignorance and insistence on relaxing the U.S. beef import ban as per U.S. hopes.



In example 14, the facial expression of the dealer is serious and determined as he points at the person lying on the gambling table presenting as the chip with the verbal tag “consumer” next to his head. The contrasting size between the huge figure of the dealer and small figure of the consumer emphasizes how consumer health is threatened by U.S. beef imports from the powerful U.S. Sitting behind the consumer, the DOH official as the gambler showing his empty pockets, claims, “To be honest, we hardly have any chips...” with a helpless look on his face.

The comparison between the states of the consumer as the chip and the DOH official as the gambler implies that the DOH is sacrificing consumer health in its negotiations on U.S. beef imports. Interestingly, the facial expression of the onlooker, who resembles President Ma, shows that he is helpless and depressed upon seeing the gambling scene. The critical message can be understood only by comparing the contrasting states of the dominant dealer, small consumer, helpless DOH official, and President Ma, which reflect the passiveness of the government under pressure from the U.S. Such a

comparison also strengthens the political position of the UDN that President Ma is as passive, helpless, and innocent as the consumers in facing the issue of U.S. beef imports. Unlike example 13 in which LT presents President Ma as the “active controller” in the central position of the cartoon, example 14 presents Ma as the “passive bystander” in a relatively peripheral position. Such different representations of President Ma subtly show the contrasting position between LT’s harsh criticism of President Ma’s manipulation and full responsibility for U.S. beef imports and UDN’s depiction of an innocent President Ma.

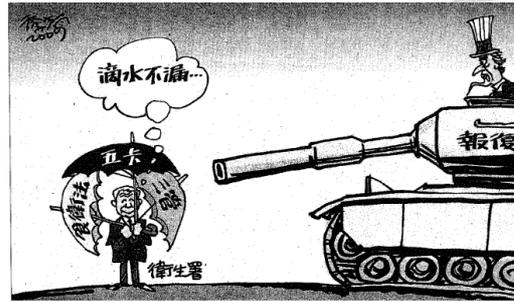


3.4.3 Different Representations of Uncle Sam in Different Scenarios

In our corpus, the “Uncle Sam” metonymy has appeared productively in different scenarios. The following examples 15 and 16 demonstrate how the productive visual figure, namely Uncle Sam representing the US, is combined with the verbal mode and conceptual level in multimodal fusion. Our analysis shows that through the contrasting multimodal features of Uncle Sam and the other characters, the relationship between the U.S. and Taiwan resulting from the issue of US beef imports is depicted and highlighted differently in LT and UDN.



Example 15. (LT 2009-10-27)



Example 16. (UDN 2009-11-04)

Example 15:

Visual (“Uncle Sam” metonymy, Ma’s caricature, flag image, strip show performer image, strip show setting) +

Verbal (“U.S. beef show/strip show (homonym)”, “sovereignty”) +

Conceptual (strip show scenario, POLITICS IS GAME)

=> U.S. beef event is strip show

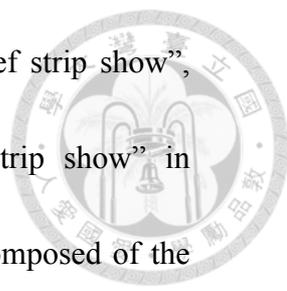
Example 16:

Visual (“Uncle Sam” metonymy, “umbrellas” metonymy/metaphor, DOH official’s caricature, tank image) +

Verbal (“revenge”, “DOH”, “three controls”, “five certifications”, “Food Sanitation Act”) +

Conceptual (war scenario, POLITICS IS WAR)

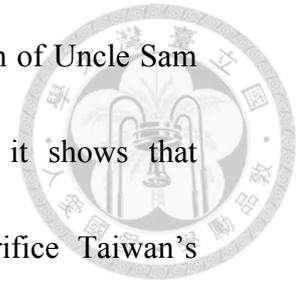
=> U.S. beef is revenger/attacker



In example 15, the verbal tag marks the stage as the “US beef strip show”, which is wordplay of the homonyms “beef show” and “strip show” in Mandarin. This verbal cue initiates an entertaining scenario composed of the sexy dancer and audience. The main target, President Ma, is the strip show performer who is throwing his pants written with “sovereignty” to Uncle Sam. However, in example 16, the DOH official is protecting himself with umbrellas while he is under the tank’s attack by Uncle Sam; ironically, the verbal tag on the tank is “revenge”, whereas the verbal tag above the DOH official states “without a single drop leaking out (well-protected)” and the three umbrellas list the three measures, “three controls”, “five certifications”, “Food Sanitation Management Act” that the DOH advocates while relaxing US beef imports.

The facial expressions contrast vividly to reflect the relationship between Taiwan and the US during negotiations over US beef imports. In example 15, President Ma’s firm and careless look not only contrasts with Uncle Sam’s cunningly smiling face but also with the shameful look from the girl on stage and the shocked, confused expression of the audience. This contrast implies that President Ma actively proposes to relax Taiwan’s ban on US beef to please Uncle Sam regardless of opposition. As a result, the contrast between

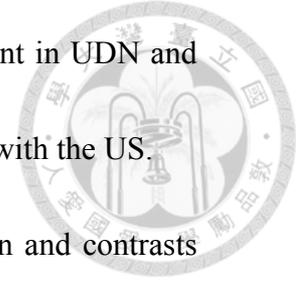
the central position of President Ma and the peripheral position of Uncle Sam reveals the give-and-take power relationship. Specifically, it shows that President Ma is the “active giver” who is willing to sacrifice Taiwan’s sovereignty and yield to the demands of the US, while the US is the “passive taker” who waits to gain the benefits.



However, in example 16, the smiling face of the DOH sharply contrasts with the revengeful fierce-looking face of Uncle Sam, highlighting the tense relationship as well as the ignorant status of the DOH official, who believes he is well protected from Uncle Sam. In addition, the size comparison between the huge tank and small DOH official and the material comparison between the soft texture of his umbrellas and the hard surface of the tank highlight the unequal power relationship between the weak status of Taiwan and the powerful status of the US.

In fact, we found that many cartoons in UDN like example 16 place strong emphasis on the powerful aspect of Uncle Sam, while President Ma is absent and the DOH is presented instead. The representation of Uncle Sam and absence of President Ma in UDN subtly indicate that Taiwan’s government is passively pressured by the powerful US on the issue of US beef imports, while President Ma is not directly responsible for the relevant policy

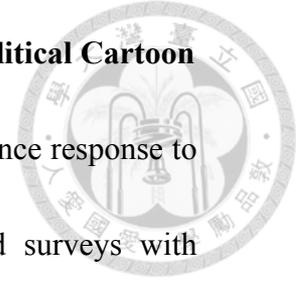
or negotiation. In contrast to LT, President Ma is usually absent in UDN and the DOH is represented as the passive and ignorant cooperater with the US.



It is widely found in our corpus that cross-cartoon comparison and contrasts sharply elicit the critical ideologies on different US-China relationship that reflects distinct political stances of LT and UDN. While LT represents a more cooperative and interactive relationship between President Ma and US, UDN emphasizes on the powerful aspects of US and the tension between Taiwan and US by highlighting the weakness of DOH and excluding President Ma in the cartoons.

These examples indicate that, although political cartoons in both LT and UDN use multimodal fusion to achieve the effect of humor or irony, different perspectives and critiques of LT and UDN are revealed through their comparisons of similar objects, namely the personified bull, President Ma, and Uncle Sam. Our analysis also shows that the complex power relationship can be simplified and highlighted through sharp contrasts and comparisons through prominent visual features and verbal details in the cartoon.

Chapter 4 Audience Response to Multimodal Fusion in Political Cartoon



In this chapter, we will present significant results of audience response to multimodal fusion in political cartoon. We have conducted surveys with participants to investigate how they interpret political cartoons and comprehend the abstractly complex political stance and critical messages. We will incorporate qualitative and quantitative analysis from our surveys to provide a detailed account combining the general pattern and specific variations of the audience response.

4.1 Questionnaire Survey: Participants' Response

To conduct the survey, we designed the questionnaire to interview how the participants interpret political cartoons concerning U.S. beef imports issues through multimodal fusion by reading the cartoon, identifying the topic, and recognizing the political stance and conceptual scenario during the interpretation process. In total, 27 participants were interviewed in this survey²¹. The participants are undergraduate students from National Taiwan University and Yuan Ze University, and they were trained with basic instructions on the background knowledge of U.S. beef imports and political

²¹ Most previous studies on readers' response adopting qualitative analysis have included the amount of participants below 20 to carefully examine the results. Such qualitative analysis usually involves time-consuming interviewing process and detailed investigation on participants' opinions, thus the amount of participants is not necessarily high to present the best result based on qualitative analysis. According to the reliability statistical data, the number of 27 participants has been considered a reliable amount of participants for the qualitative and quantitative results in this study.

stances of two newspapers (*Liberty Times* and *United Daily News*) from which the political cartoons of our multimodal corpus are collected²².



4.2 Survey Procedures

In this section, the survey procedures are introduced as follows.

During the survey, each cartoon is presented to the participant without showing the newspaper title and followed with three main questions. The survey was conducted in Mandarin Chinese (for more details, please see Appendix 3 for the original version of the political cartoon survey questionnaire). We will provide the translated English version of the political cartoon survey questionnaire in the following. The three main questions of the survey are:

1. Based on your interpretation, is it easy to understand the perspectives toward U.S.

beef imports in this cartoon?

Very easy Easy Neutral Difficult Very Difficult

(Please explain how and why)

2. Based on your interpretation, which newspaper was this political cartoon published in?

Liberty Times United Daily News I don't know

(Please explain how and why)

²² Previous studies on political cartoon interpretation have pinpointed that participants with certain level of background knowledge are more likely to read and understand political cartoons due to the fact that political cartoons often ironically represent complex political issues in subtle ways. Therefore, to understand how the participants interpret the political cartoons concerning U.S. beef imports, we believe it is necessary to inform them the background knowledge as the basic training.

3. With the abovementioned, during your interpretation process, did the concept of

“Trojan War” appear in your mind?

Yes. Of course it did. Yes, it did. I don't know. No, it didn't.

No, it didn't at all. (Please explain how and why)

In total, ten cartoons are included in the survey, and the above three questions are listed below each cartoon for participants' evaluation. For the third question, we expect each cartoon to be interpreted on its own concept level based on our previous analysis, thus different concept, such as “Trojan War” or “mountain”, was asked in each cartoon.

To answer the first question, the participants are asked to read the cartoon and evaluate whether it is easy or not to interpret how U.S. beef issue is represented in each cartoon. They are asked to describe how they interpret this cartoon and why they had such interpretation.

To answer the second question, the participants are asked to identify the newspaper in which it is published. They are asked to explain how they made the judgement and why they think it is published in the particular newspaper.

To answer the third question, the participants are asked to recall whether the specific concept, which is the conceptual scenario based on our analysis, has appeared in their mind during their interpretation process. If the answer is “yes”, they are asked to illustrate how the concept involved in the



interpretation; if the answer is “no” or “I don’t know”, they are asked if they had come up with other concepts during their interpretation (instead of the one being asked in the question).



Lastly, in the closing of the questionnaire, the participants are asked to evaluate the frequency of the mappings of conceptual metaphor POLITICS IS GAME and POLITICS IS WAR based on the following two questions:

1. Based on the previous political cartoon interpretations, what is the frequency of

the “game metaphor” representing political issues?

Very high High Neutral Low Very Low

(Please explain how and why)

2. Based on the previous political cartoon interpretations, what is the frequency of

the “war metaphor” representing political issues?

Very high High Neutral Low Very Low

(Please explain how and why)

To answer the above questions, the participants are asked to explain why they consider the frequency of “game” or “war” metaphor representing political issues is high or low.

As a result, we have conducted the survey by interviewing 27 participants

to know how they interpret each political cartoon. The results show several insightful findings, which will be introduced in the next section.



4.3 Results and Discussion

In this section, we will present the quantitative analysis to account for a general pattern from the statistical data, and we will further evaluate and examine the general findings and distribution of the statistical analysis with the qualitative analysis from our survey.

In this study, the statistical data of political cartoon survey is conducted with SPSS analysis. The reliability statistics²³ (Cronbach's Alpha= 0.902) indicates that the reliability of survey is high based on the percentage of 90.2. This suggests that the participants' answers are reliable and the design of the questionnaire is well arranged.

In general, the statistical analysis of political cartoon survey indicates several significant findings:

- (1) Among ten cartoons, there are significant differences in the interpretation level
- (2) Among ten cartoons, there are no significant differences in political stance identification.

²³ The reliability statistics from SPSS analysis indicates that the reliability of the survey is high if the Cronbach's Alpha is higher than 0.75.

(3) In each cartoon, there are significant frequencies in the distribution of the involvement of the conceptual scenario in the interpretation process.

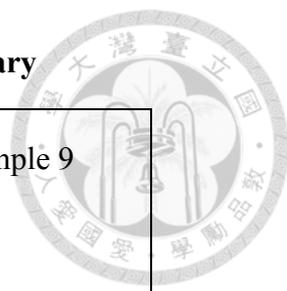
(4) Overall, there are significant frequencies in the distribution of the conceptual metaphors, POLITICS IS GAME and POLITICS IS WAR.

In the following, we will discuss the above general findings of the statistical analysis and incorporate the quantitative analysis with the qualitative analysis from our survey.

4.3.1 Significant Differences in the Interpretation Level

In the first question, we asked the participants to evaluate the level of easiness or difficulties to interpret the political cartoon. Among ten cartoons, we found significant differences exist in the pairwise comparisons (see Appendix 4), which suggests significantly different interpretation levels in the comparison between each cartoon. Specifically, the SPSS analysis shows the parallel significant differences between cartoon “examples 3, 4, 7, 8” and “examples 5, 6, 9”. The mean difference of examples 3, 4, 7, 8 and examples 5, 6, 9 from pairwise comparison statistical analysis is summarized in Table 6. The mean difference is significant at the 0.05 level (P value < 0.05).

Table 6. Pairwise Comparison Statistical Data Summary

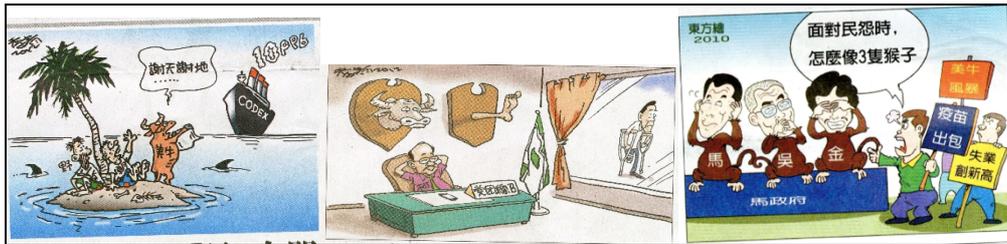


Mean difference	Example 5	Example 6	Example 9
Example 3	.926	.778	.630
(P Value)	(0.002)	(0.009)	(0.034)
Example 4	.815	.667	.741
(P Value)	(0.006)	(0.025)	(0.013)
Example 7	.889	.741	.667
(P Value)	(0.003)	(0.013)	(0.025)
Example 8	.889	.741	.667
(P Value)	(0.003)	(0.013)	(0.025)

In particular, the mean difference is the highest between example 5 and examples 3, 4, 7, 8 as shown in Table 6, which implicates that most participants find example 5 much more difficult to interpret than the other examples. The cartoon images of these two sets of examples are listed in comparison in the following Figure 12.



Example 3, 4, 7, 8 (From left to right)



Example 5, 6, 9 (From left to right)

Figure 12. The Parallel Comparison between Example 3, 4, 7, 8 and 5, 6, 9

During our interviews, over 80 % of the participants find it difficult to interpret example 5 due to the fact that they do not know what the verbal tags “CODEX” and “10ppb” stand for. This finding suggests that if the audience cannot interpret the meaning of the “verbal mode” in the first hand, they can hardly build the link between the multimodal level (visual and verbal mode) and the conceptual level, thus it is difficult for them to interpret the multimodal fusion with humorous or ironic effects.

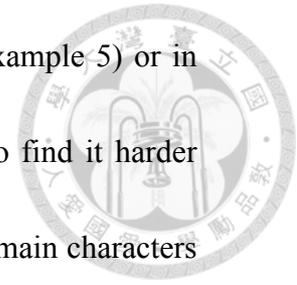
In addition, example 6 has the second-highest mean difference with examples 3, 4, 7, 8 among all. In example 6, more than 50 % of the participants found it hard to interpret this cartoon due to the fact that they

cannot recognize the visual caricature of “Mr. Ko” and they do not know what the verbal tag of “Ko” stands for. Therefore, the results show that the participants find it harder to interpret political cartoons if they encounter difficulties in recognizing and understanding the visual and verbal mode on the multimodal level.

The results from the statistical analysis may reflect an interesting pattern of the participants’ interpretation tendency. In our survey, we found that when the participants see prominent visual figures of “President Ma” as the main visual cue as in example 3, 4, 7, 8, they seem to find it easier to interpret the cartoons by giving a relatively straightforward critical comment concerning President Ma and the topic of U.S. beef imports. When the visual or verbal cues are directly and prominently associated with president Ma and U.S. beef imports, the participants tend to notice and interpret the close connection between president Ma and U.S. beef imports represented metonymically and metaphorically by the bull (example 3, 4), shooting arrow (example 7) or bomb (example 8) through the multimodal fusion of visual image and verbal tags in the cartoons.

However, when the participants read the cartoons in which the visual or verbal cues are not directly and prominently associated with president Ma and

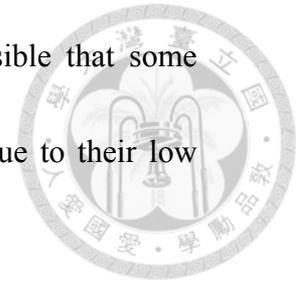
U.S. beef imports, especially when President Ma is absent (example 5) or in the peripheral position (example 6, 9), the participants tend to find it harder and spend more time figuring out the relationship between the main characters and the topic U.S. beef imports. As a result, some of them find it difficult to interpret these cartoons by giving misleading or irrelevant comments concerning the topic.



4.3.2 Political Stance Identification

In the second question, we asked the participants to judge which newspaper, Liberty Times (LT) or United Daily News (UDN), has the cartoon been published in. The statistical data shows no significant differences (P value > 0.05) in political stance identification among ten cartoons. Based on the qualitative analysis of our survey, most participants confess that they are not interested in politics and they don't often read newspapers. Some of the participants express their dislikes in certain politicians and political situations in Taiwan, while some of them show neutral stance in politics. Only few of the participants show their interest in politics and further express their political stance and their opinion on certain political issues, including the sun flower student movement and recent elections of mayors. The results show that some participants may find it too complex or difficult to make the correct judgment

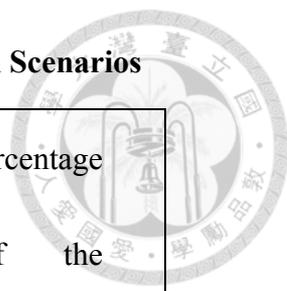
through the interpretation on political cartoons, or it is possible that some participants unconsciously resist in answering this question due to their low interest in politics or negative attitudes toward politics.



4.3.3 The Involvement of Conceptual Scenario

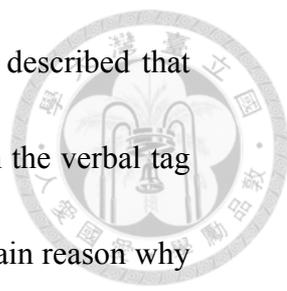
In the third question, we have conducted the statistical analysis of Chi-Square test (See Appendix 4) to evaluate the involvement of each conceptual scenario in participants' interpretation of each political cartoon. The results indicate significant differences among the frequencies in the distribution of the involvement of conceptual scenarios in each cartoon. In all of the tests, P values are all under 0.05, which suggests a strong correlation between the participants' answer and the question. In viewing the test conducted for each cartoon, we found higher frequencies of distribution occur in "recognizing" the involvement of the conceptual scenarios of Troy, Circus, Cowboy, Island, Bomb, Three Monkey, and Net-fixing. In other words, more than 50 % of the participants agree that these conceptual scenarios appear in their mind and involve the interpretation process. To clearly present the positive response to the involvement of these conceptual scenarios, the p value and the percentage of audience response will be summarized in Table 7.

Table 7. Positive Response to the Involvement of Conceptual Scenarios



Example NO.	Conceptual Level Conceptual Scenario	P value (<0.05)	Agree percentage (%) of the participants
1	Trojan War	0.029	70
2	Mountain	0.000	74
3	Circus	0.000	78
4	Cowboy	0.000	67
5	Island	0.000	59
8	Bomb	0.000	89
9	Three Monkey	0.009	70
10	Net fixing	0.037	63

The above table indicates that the agree percentage of the participants is high in view of the involvement of the above conceptual scenarios. Among all, the highest percentage (89%) of the participants agree that the concept of “bomb”



appear in the interpretation process. In fact, these participants described that the prominent cues of the visual image of “bomb” marked with the verbal tag “U.S. beef chaos” in the central position of the cartoon is the main reason why they found the bomb scenario involved in their interpretation of example 8. Similarly, as Table shows, the circus scenario is highly involved in the participants’ response with 78 % (the second highest percentage), while the participants also mentioned that the visual settings of circus, visual image of the animals, host, and audience, and the verbal tag of “2010 circus performance”, altogether clearly remind them of the “circus” scenario while interpreting example 3. Therefore, the results indicate that prominent multimodal cues of the visual and verbal mode are interactively connected with the conceptual level and influenced audience interpretation on the conceptual level.

As a matter of fact, as Table 8 shows, higher frequencies of distribution occur in “denying” the involvement of the conceptual scenarios of “Hunting” and “War Horse”, which shows that over 50 % of the participants disagree that these conceptual scenarios appear in their mind during the interpretation process. In our interviews, 59% of the participants claimed that the “hunting scenario” is not the idea or concept that came up in their mind while

interpreting example 6, instead, some of the participants came up with “office” scenario when interpreting the cartoon.

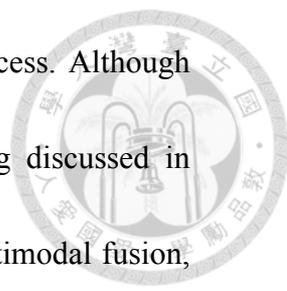


Table 8. Negative Response to the Involvement of Conceptual Scenarios

Example NO.	Conceptual Level Conceptual Scenario	P value (<0.05)	Disagree percentage (%) of the participants
6	Hunting	0.002	52
7	War Horse	0.000	59

In addition, as Table 8 shows, 52 % of the participants disagreed and mentioned that they did not associate the visual image of the “bull” in example 7 (fused with Yang’s head) with the “war-horse” scenario at all. Some of the participants suggested that the “war scenario”, “escape scenario”, or “defense scenario” instead of “war horse scenario” involved in their interpretation.

Interestingly, the results revealed from the quantitative and qualitative analysis of the survey provides us a better understanding of how different conceptual scenarios may be activated and how conceptual level may work



differently for different audience during the interpretation process. Although our analysis proposed that certain conceptual scenario (being discussed in Chapter 3) as the main element on the conceptual level of multimodal fusion, we are NOT suggesting that the conceptual scenario as the only scenario involved in each cartoon. In fact, our analysis proposes that certain conceptual scenario may be the most prominent and significant conceptual scenario associated with the topic and combined with the visual and verbal mode, which enables the readers to comprehend critical messages, ironic, and humorous effect. However, various conceptual scenarios may be activated on the conceptual level and highlighted through multimodal fusion from different perspectives of audience.

4.3.4 The Distribution of the Conceptual Metaphor

In the third question, we asked the participants to evaluate whether the frequency of the distribution of GAME domain and WAR domain in conceptualizing political issues is high, low, or neutral with five degrees of scores. The statistical analysis of Chi-Square test has been conducted to evaluate the distribution of the conceptual metaphors, POLITICS IS GAME and POLITICS IS WAR, as summarized in Table 9.

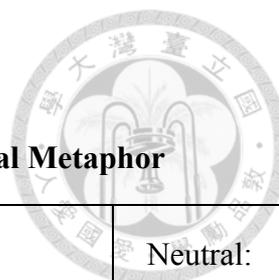


Table 9. Audience Response to the Distribution of Conceptual Metaphor

Conceptual Metaphor	P value	High distribution: percentage (%) of the participants	Low distribution: percentage (%) of the participants	Neutral: percentage (%) of the participants
POLITICS IS GAME	<0.05	89	4	7
POLITICS IS WAR	0.000	30	4	67

Overall, there are significant differences among the frequencies in the distribution of the conceptual metaphors, **POLITICS IS GAME** and **POLITICS IS WAR**, while P values are under 0.05. As Table 9 indicates, 89% of the participants consider **GAME** domain conceptualized **POLITICS** (**POLITICS IS GAME**). However, only 30 % of the participants found **WAR** domain conceptualized **POLITICS** domain (**POLITICS IS WAR**), instead, 67 % of the participants remain a neutral stance toward the occurrence of

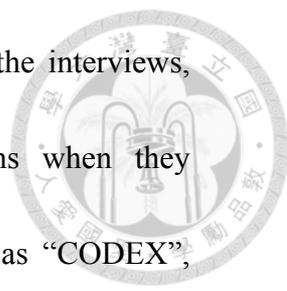


POLITICS IS WAR. In fact, this reveals the fact that the participants may have an overall picture that POLITICS IS GAME occur most frequently through their interpretation on the representative examples. However, this finding also implicitly correlates with our analysis since we found that POLITICS IS GAME is the most identified conceptual metaphor (32%) while POLITICS IS WAR is the second most often used metaphor (13%) in our multimodal corpus.

4.3.5 Effective Multimodal Fusion and Interpretation

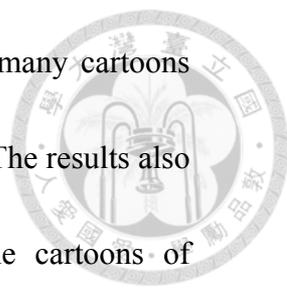
Based on the above findings, we propose that the higher level of interpretation may be relied on an effective multimodal fusion of the verbal mode, visual mode, and conceptual level in political cartoons. We will discuss the findings from the qualitative analysis to explore how each mode works and functions in multimodal fusion to achieve meaningful interpretations.

In view of the verbal mode, the participants' answers show that "verbal mode" plays a crucial part in the interpretation process. Most participants interpret the cartoon by reading the "verbal tags" of the images carefully (read out loud or to draw the lines under the words) to receive verbal cues and build the connections between the verbal and visual mode. While reading the verbal message, they start to make inferences and connect the relationships between



the verbal mode, visual mode, and conceptual level. During the interviews, participants often have questions and ask for explanations when they encounter few verbal tags that are unfamiliar to them, such as “CODEX”, “Little Su”, etc. In fact, these terms are relevant to the detailed background knowledge of a certain event concerning U.S. beef imports. Therefore, this echoes our claim that reading the connotations of the “verbal mode” provides a key to an in-depth comprehension of the critical message in political cartoons. As we proposed in multimodal fusion model, the results of audience response also reflects that the verbal mode can be viewed as the crucial link between the visual mode and conceptual level to build the connection between the visual image and the conceptual scenario, which enables the readers to further interpret the complex political ideologies and critical messages in the cartoons.

In the visual mode, we found readers have difficulties in recognizing some of the politicians’ caricatures. When the readers cannot recognize the correct politician’s caricature, such as the caricature of “Mr. Su” in example 8 or “Mr. Ko” in example 6, they can still make inferences and judgments concerning the matter. However, some of their interpretations are irrelevant and misleading. Interestingly, President Ma’s caricature is the most



recognizable figure among all since he is the main target in many cartoons concerning U.S. beef imports (as shown in example 3, 4, 7, 8). The results also show that most participants found it easier to interpret the cartoons of president Ma with the productive “bull-related metonymies” such as in examples 3, 4, 7, while many participants found it harder to interpret the cartoons comprised of novel metaphors implicitly related to the key concept or keywords connected to U.S. beef events in the diversified scenarios, such as in the isolated island scenario, office scenario, and three monkey scenario in examples 5, 6, 9.

Regarding the conceptual level, we found that understanding the conceptual scenario, such as “Trojan war”, “three monkey”, and “net-fixing” (Taiwanese Idiom and Taiwanese Song), requires readers’ cultural background of language-specific and culture-specific comprehension on the conceptual level. The results show that readers without such cultural background couldn’t interpret the critical meaning and irony at all.

To summarize, our survey suggests that the interpretation level is relevant to the audience capability in recognizing, interpreting, and combining the verbal mode, visual mode, and conceptual level through multimodal fusion in political cartoons.

Chapter 5 Multimodal Fusion in Poetry Music: The Encounter between

Language and Music in “Serendipity”

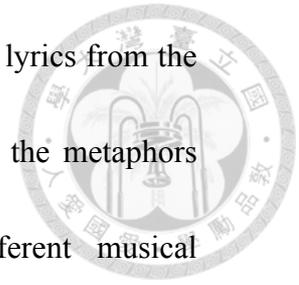


This chapter aims to provide a detailed account of how verbal metaphors can be mapped on, transformed to, and combined with the aural/musical mode in the compositions. In the following sections, we will further identify the important aspects of compositions that realize the metaphors and metaphorical mappings in each version of song and conclude with a classification of the verbal-musical representations and composing tendency in the four versions. Section 5.1 presents data description, while Section 5.2 present data analysis. Results and discussion will be included in Section 5.3.

5.1 Data Description of Poetry Music: “Serendipity”

In total, four musical versions have been composed for the classic poem, “Serendipity”: (1) Version 1 by Lien Pi-Kuang (連璧光); (2) Version 2 by Hsu Te-Chu (許德舉); (3) Version 3 by Li Wei-Ning (李惟寧); and (4) Version 4 by Chen Chiu-Hsia (陳秋霞). The compositions of versions 1, 2, and 3 are based on the classical performing style of art songs (“Lied” or “Kunstlied” in Germany), which consists of a vocal artist and piano accompaniment. Version 4 is the only popular music version of “Serendipity,” composed by the singer-songwriter Chen Chiu-Hsia in her autobiographical

movie “Chiu-Hsia.” Although all four versions adopt the same lyrics from the poem “Serendipity,” each composer interprets and represents the metaphors and metaphorical expressions in the poem through different musical techniques. We will provide a detailed multimodal analysis of each version in the following sections.



5.2 Data Analysis

In the following section, we will illustrate our analysis based on the theoretical basis of multimodal fusion model. We propose that the cognitive mechanism of multimodal fusion evolves from the metonymic-metaphoric network, which combines the visual, aural, verbal mode, and conceptual level in the genre of poetry music as shown in Figure 13. Through the multimodal fusion, the metonymies and metaphors suggested by verbal, visual, and aural modes provide metonymic representations and metaphorical mappings with the conceptualization of the abstract concept underlying the poetry lyrics.

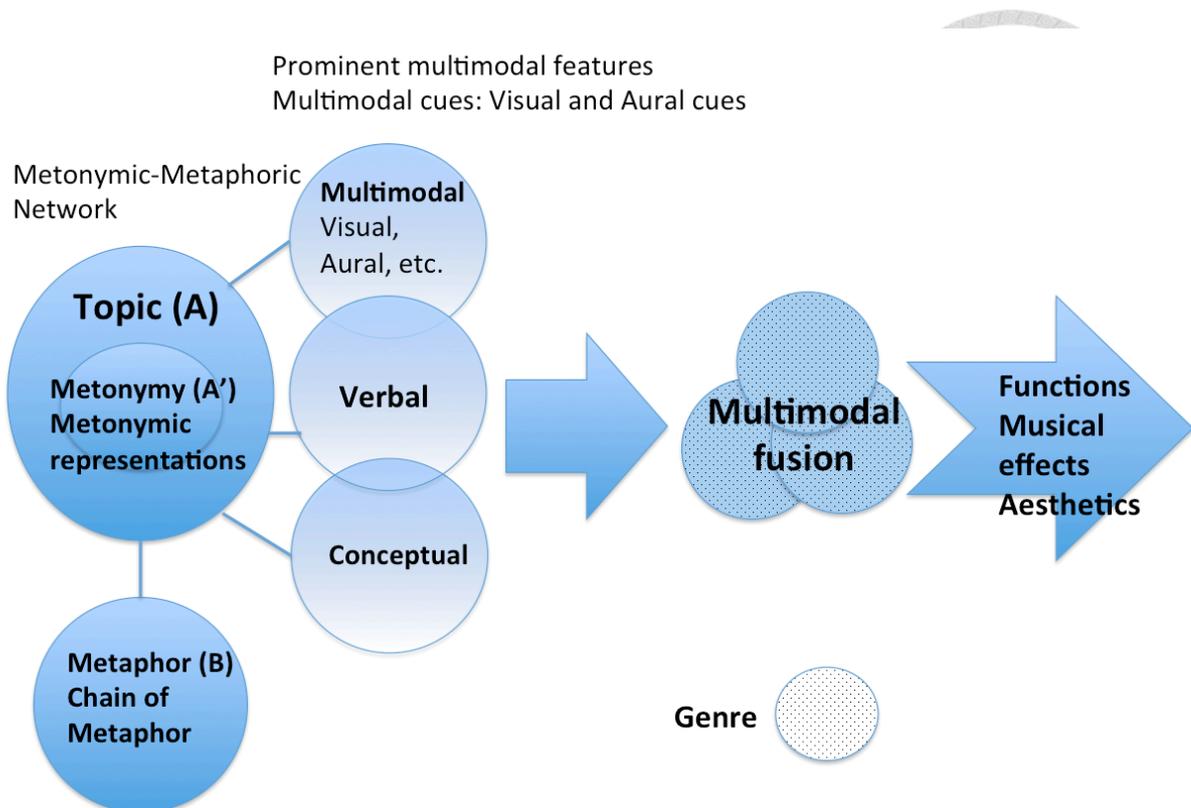
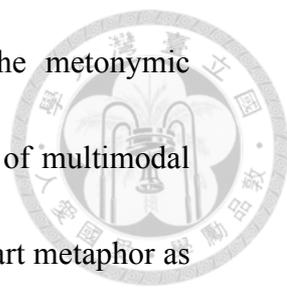


Figure 13. Multimodal Fusion in Poetry Music

In the verbal mode, the poetry lyrics convey the abstract concept of space, time, emotion, and feelings, etc. The concept underlies the poem has been conceptualized, visualized, and musicalized in four versions of songs through the “pictorial effect” on the score in the visual mode and the musical representations in the aural musical mode, creating different musical highlights and aesthetic effects. Through our analysis, we found metonymic representations of metaphors in examples of partial visualization of cloud image on the score representing CLOUD metaphor, aural simulation of wave rhythm representing SEA metaphor, etc. The metonymies representations



demonstrate the part-for-whole characteristics in building the metonymic association through visual, aural, verbal, and conceptual level of multimodal fusion, and show its indistinguishable feature with the counterpart metaphor as the metonymy-based metaphors in poetry music. Thus, we propose that multimodal fusion starting with metonymic-metaphoric network representing the concept that activate the conceptual level from poetry lyrics and then connecting the visual, aural, and verbal based on the metonymic associations and metaphorical mapping in different versions of songs.

To answer the research question concerning the transformation from poetry to music, we first identify and select the metaphors and metaphorical expressions in the poem²⁴, and then we incorporate the musical analysis based on the score of the four versions to examine how specific verbal metaphors and metaphorical expressions are represented in poetry lyrics combined with the musical mode. The following is the original Chinese version of “Serendipity” and our English translation.

²⁴ In the following metaphor-based analysis, we are aware that various approaches and perspectives of interpretation can be included in the analysis of the poetry “Serendipity” due to its rich meaning. We understand “Serendipity” can be viewed as the poetry of LOVE or other types of relationships in life. However, in order to incorporate the metaphor-based analysis with musical analysis on musical scores in the following analysis, we will mainly provide the analysis that mainly focuses on the metaphor in each sentence to select the main verbal metaphor and carefully examine how each verbal metaphor and its metaphorical expressions have been represented multimodally in the “musical sentence” of the musical score.



我是天空裡的一片雲，

I am a CLOUD in the sky,

偶然投影在你的波心，

by chance reflected on the wave of your heart

你不必訝異，

Don't be surprised,

無需歡喜，

Or too elated;

在轉瞬間消滅了蹤影，

In an instant I shall vanish without trace.

你我相逢在黑夜的海上，

We meet on the sea of dark night,

你有你的，我有我的方向，

You are on your way, I am on mine.

你記得也好，

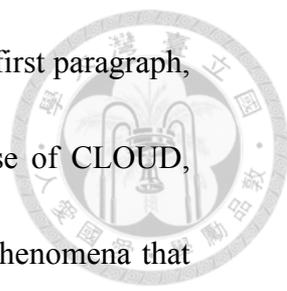
Remember if you will,

最好你忘掉，

Or, better still, forget

在這交會時互放的光亮

The glowing light exchanged in this encounter.



In this poem, the poet depicts a chance encounter in life. In the first paragraph, the poet uses a first person narrative and adopts the novel use of CLOUD, REFLECTION, and WAVE as metaphors, which are natural phenomena that keep moving and changing, to express his complex feelings with uncertainty, as the line “I shall vanish without trace” echoes such feelings and emphasizes the intangible instantaneous moment of the encounter. The major metaphors in the first paragraph are:

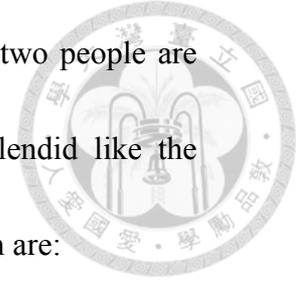
NARRATOR IS CLOUD

ENCOUNTER IS REFLECTION ON THE WAVE

The REFLECTION metaphor plays an important part in connecting the CLOUD and the WAVE metaphors. Imagining the scene of the reflection process, readers’ perspectives are also moving/changing and shifting from the upward CLOUD to the downward WAVE. Furthermore, the concept of “light” is essential in forming a reflection on the wave and echoing with the dark-light contrast in the second paragraph of the poem.

In the second paragraph, the poet uses the SEA, DARK NIGHT, and LIGHT metaphors to conceptualize the commencement and engaging process

of the encounter. The poet conveys the feeling that although two people are heading in different directions in life, their encounter is splendid like the glowing light in the dark night. The metaphors in this paragraph are:



ENCOUNTER LOCATION IS SEA OF DARK NIGHT

ENCOUNTER MOMENT IS GLOWING LIGHT

In the beginning of the second paragraph, the poet begins with the SEA metaphor in the dark of the night, which marks a contrast with the bright sky and reflection of the daytime in the previous paragraph. Also, the LIGHT metaphor in the end also brightens and lights up the dark night scene in this paragraph.

Thus, we believe the contrast of light and darkness is an essential element for interpreting the metaphors in this poem. Such contrast of brightness in the daytime and darkness at night has been highlighted through the major metaphors CLOUD, REFLECTION, SEA, and LIGHT, to convey the main theme, the narrator's emotional feelings toward the chance encounter.

The CLOUD, REFLECTION, SEA, and LIGHT metaphors are identified and selected as the main focus of the musical analysis. In the following

sections of our analysis, we will mainly focus on how these four metaphors and relevant metaphorical expressions in poetry are transformed and represented in the musical mode.



In the following, we will incorporate into our metaphor analysis the musical analysis to provide a detailed investigation on the scores (樂譜) of the four versions. In our analysis, the musical terms are explained in footnotes with definitions adapted from the reference of *A Dictionary of Music and Musicians* (Grove, 2009). In addition, although the English translation of the poem has already been presented in Section 3.1, we will provide our word-to-word English translation of the poem's lyrics in the score to facilitate the analysis in relation to the interrelationship between the metaphor in language mode and musical mode.

5.2.1 Version 1

In this version, the composer captures the concrete image of “reflection” by creating the contrast of space between the sky and the wave. As Figure 1 shows, when the poetry lyric begins with the I (NARRATOR) AM CLOUD metaphor, the lower (Fa) *leaps*²⁵ to the higher notes of (Re) and (Do) and continues with the nearby high notes to symbolize the cloud being high above

²⁵ *Leap*, “the skip between intervals for more than four steps,” for example, from Do to So.

the sky, and then it *leaps* to the lower (Fa) to prepare for the REFLECTION metaphor and proceeds with the lower notes of (Re) and (Do). The musical analysis of the score shows that the former higher notes and the latter lower notes are in fact similar in melody and tempo with the frequent use of (Re) and (Do). Such similarity in notes simulates the process of a cloud projecting and reflecting on the wave, which corresponds to the concrete image of the CLOUD and REFLECTION metaphors in the poem. As shown in Figure 14, the contrast of upper and lower space is clearly reflected through the shifting from higher to lower notes. Therefore, we suggest that the composing pattern of this version is to highlight the concrete image of the metaphors to simulate the visual imagination of reflection process with the space/perspective shift through the contrast of higher and lower notes.

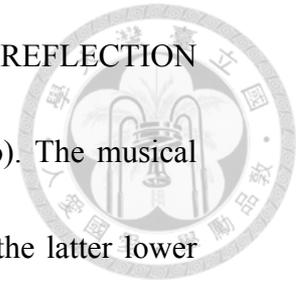


Figure 14. Version 1: Paragraph 1

In the second paragraph of poetry lyrics as shown in Figure 15, the composer uses *arpeggio*²⁶, ascending figures, in the piano accompaniment. In addition, lower dynamic (*mp*), lower pitch, and darker *timbre*²⁷, which is darker compared with the previous paragraph, combined with the ascending figures construct the scene of the SEA metaphor.

²⁶ *Arpeggio*, ‘the harp,’ “the employment in vocal or instrumental music of the notes of a chord in succession instead of simultaneously; also, in pianoforte music, the *breaking* or *spreading* of a chord, either upwards or downwards.”

²⁷ *Timbre*, ‘sound-colors,’ “now become the word ‘Quality.’ A sound said in fair English to possess three properties: Pitch, Intensity, Quality; respectively corresponding to the Frequency, the Amplitude, and the Form of the Sound wave.”

The image shows a musical score for a song. The top system features a vocal line and a piano accompaniment. The piano part has several arpeggiated chords circled in blue, with a blue box above them labeled 'Arpeggio(Ascending figures)'. The lyrics are in Chinese and English. The bottom system continues the piano accompaniment with more arpeggiated chords circled in blue. A watermark of Tsinghua University is visible in the top right corner.

Figure 15. Version 1: Paragraph 2-Part 1

Furthermore, as the score in Figure 15 shows, the image of *arpeggio*, the recurring ascending figures, are visually simulating the image of ocean waves, which creates an example of the pictorial effect shown on the score from our point of view.

In the final part, as shown in Figure 16, the dynamic increases gradually from (*p*) to (*mf*), and then it increases to (*f*) with the extending *tremolo*²⁸ that directly falls on the verb “glowing” in the poetry lyrics, and then the dynamic decreases in the end.

²⁸ *Tremolo*, “a figure consisting, on the pianoforte music it is a rapid alternation of the parts of divided chords, reproducing to a great extent a very fine effect.”

國立臺灣大學

你記得也好， 最好是忘
 You remember better forget

掉， 在這交會時， 互放的光亮。
 In this encounter exchanged glowing light

p *mf* *f* *mf* *8va*

Figure 16. Version 1: Paragraph 2-Part 2

We suggest the increasing dynamic in the music mode is associated with the metaphor of LIGHT that glows in the dark night. With the use of prominent loud volume (*f*) and motional extending *tremolo*, the composer emphasizes the “glowing” aspect of light as the instantaneous yet distinguished moment of the encounter.



5.2.2 Version 2

Before the poetry lyrics begin, the composer starts with a splendid and spectacular style of introduction with the largest dynamic (*ff*), a moderate pace (*Andante*), wide tone range, and repetitive rising-falling figures to achieve the “Motif for cloud.”

The image displays a musical score for the introduction of Version 2. The score is written for piano and is in 3/4 time, key of B-flat major. The tempo is marked 'Andante'. The dynamic is marked 'ff' (fortissimo). The score features a piano introduction with a forte (ff) dynamic. The melody is characterized by repetitive rising-falling figures, which are highlighted with blue circles and a speech bubble. A blue arrow points to the wide range of the melody, labeled 'Wide tone range'.

Figure 17. Version 2: Introduction

As Figure 17 shows, the CLOUD metaphor has been visually represented through the repetitive rising-falling figures in the notes, which we suggest create the pictorial effects that visually correspond to the concrete image of CLOUD based on the composer’s arrangement/intention. To be more specific,

we have surveyed the cloud image online and found the most highly used photo and pictorial cloud image as the shape of cloud in Figure 18, which might be the prototypical cloud image.

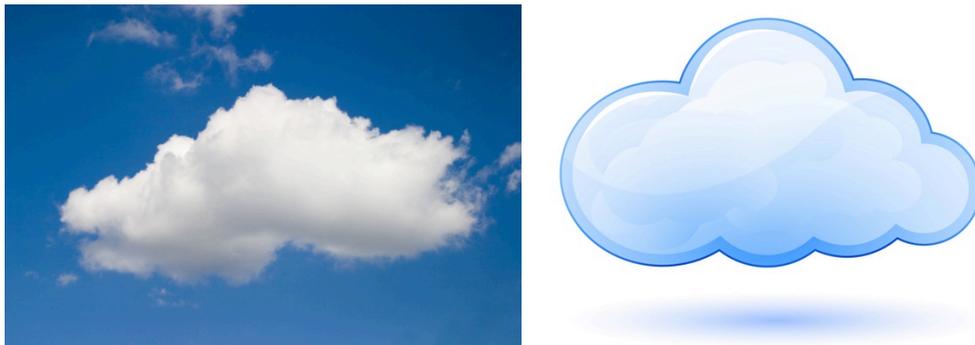


Figure 18. Photo Cloud Image and Pictorial Cloud Image²⁹

Comparing the cloud images in Figure 18 with the score in Figure 17, we found seven rising-falling figures that resemble the pictorial cloud on the score. The continuing musical flow of the rising-falling figures also resembles the indivisibly floating feature of cloud. This version provides an interesting example of how the concrete image of a metaphor can be visualized and musicalized in the multimodal genre of poetry music.

In the first paragraph of version 2 shown in Figure 19, the composer begins with the rising figure in large volume (*mf*) to reach the highest (Me) as

²⁹ The photo cloud image is from <http://labs.sogeti.com/wp-content/uploads/2014/03/cloud.jpg> and the pictorial cloud image is from <http://www.gamesparks.com/wp-content/uploads/2013/07/the-cloud.jpg>

reaching the sky in lyrics and then uses the falling figure to depict the CLOUD metaphor with ornate speedy falling figures in the piano accompaniment, preparing for the following REFLECTION metaphor.



The image displays two systems of musical notation. The first system features a vocal line in treble clef with lyrics in Chinese and English, and a piano accompaniment in bass clef. Annotations highlight a 'Rising figure' in the vocal line and a 'Falling figure' in the piano accompaniment, with a callout 'CLOUD' pointing to the piano part. The second system follows a similar structure with lyrics '偶爾投影在你的波心!' and 'By chance reflected on your wave heart', with annotations for 'Rising figure', 'Falling figure', and a callout 'REFLECTION lower pitch' pointing to the piano accompaniment.

Figure 19. Version 2: Paragraph 1

Then the composer again uses the rising and falling figures, which have a similar pattern and tempo to the previous ones, and the similar piano accompaniment in lower pitch to portray the reflection of “CLOUD” on the waves. The change from higher to lower pitch may also be associated with the

change of space and perspective shift. The nearly identical patterns of the rising figure, falling figure, and piano accompaniment in Figure 6 indicates that the composer uses the similar compositions to present the projected image of the CLOUD metaphor as the REFLECTION on the wave aurally but also visually in the musical mode. We found that this composer in particular values the pictorial effect and visual presentation of the notes while interpreting metaphor in the poetry lyrics.

In the second paragraph of version 2 as shown in Figure 7, the composer shifts to the minor key and uses the recurring *arpeggio* in the piano accompaniment with increasing-decreasing volume to construct the scene of the SEA metaphor.

Minor Key

Arpeggio

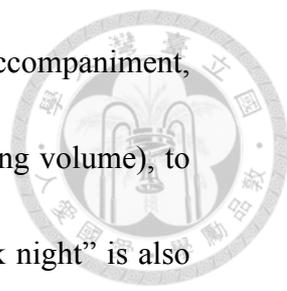
increasing-decreasing volume

mf

你 我相達在黑夜的海
You I meet dark night sea

mf

Figure 20. Version 2: Paragraph 2-Part 1



We suggest that the composer uses *arpeggio* in the piano accompaniment, playing in *crescendo* and *decrescendo* (increasing and decreasing volume), to simulate the image and rhythm of the SEA. The scene of “dark night” is also represented through the low pitch and dark *timbre* of *arpeggio* in the piano accompaniment. We propose such representation is extended from the music metaphor DARK IS LOW, since darker timbre is generally associated with lower pitch and key (Gislason, 2011: p. 35-36), thus the image of “dark night” has been musicalized through low pitch and dark timbre.

In the final part, as Figure 21 reveals, the last word “light” is reached with the highest pitch and longer duration, accompanied with the loud dynamic from (*f*) to (*ff*) and a wider range in the piano part, resulting in the bright timbre as a powerful ending.

你記得也好，最好你忘掉，在這交會時互放的光
 You remember better forget in this encounter exchanged glowing

light
 亮!

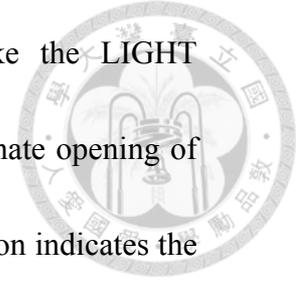
Highest pitch and longer duration

Wider range

Figure 21. Version 2: Paragraph 2-Part 2

In this version, the bright image of the LIGHT metaphor is highlighted and strengthened with the prominent musical features of high pitch and bright timbre. We believe this also extends from the music metaphor HIGH IS BRIGHT, since a bright timbre is generally represented through higher pitch and key (Gislason, 2011: p. 35-36), thus the bright image of LIGHT has been musicalized via high pitch and bright timbre. The composer places great emphasis on the LIGHT metaphor in the ending part as if telling the audience

that this chance encounter is splendid and glamorous like the LIGHT metaphor in the musical mode, which also echoes with the ornate opening of the introduction in the beginning. We believe such representation indicates the composer's coherent intention of representing the concrete image of metaphors in poetry with prominent visual and aural cues through the musical mode.



5.2.3 Version 3

In the first paragraph of version 3 as shown in Figure 22, the composer begins with *Legato*³⁰ in the piano accompaniment to maintain a well-connected rhythm and flow with lower dynamic (*mp*). When the poetry lyrics begin, the dynamic turns loud (*mf*) and the melody proceeds with the recurring notes of (La) and (So) and a narrow tone range with several one-step-intervals (like the interval between La and So) in representing the CLOUD and REFLECTION metaphors.

³⁰ *Legato*, 'connected;' "the sounds of each note of a phrase being sustained until the next is heard."

The image shows a musical score for a piece titled "Andante Con moto". It consists of four staves. The top staff is the vocal line, starting with a circled *mf* dynamic and the Chinese character "我" (I) under a note. A blue box labeled "So" points to a note in this staff. The second staff is the piano accompaniment, starting with a circled *mp* dynamic. The third staff is the vocal line with lyrics in Chinese and English: "是天空裏的一片雲 偶爾投影在你的波心" and "am sky in a cloud by chance reflect on your wave heart". A blue box labeled "La" points to a note in this staff. The bottom staff is the piano accompaniment. Annotations include "legato" with arrows pointing to the vocal line, and "Narrow tone range" with a blue box and arrows pointing to the vocal line's range. A circular watermark is visible in the background.

Figure 22. Version 3: Paragraph 1

Unlike versions 1 and 2, this version focuses on neither the contrast of high and low notes nor the pictorial effect and prominent musical features. The musical analysis in Figure 23 reveals the composing pattern in this version does not correspond to the concrete image of the CLOUD and REFLECTION metaphors. We suggest this version highlights the abstract metaphorical extension of the sentimental emotion that extends from CLOUD and REFLECTION metaphors. As the poetry lyrics say “don’t be surprised or too elated,” the melody begins with a relatively soft tune and narrow tone range to interpret the peaceful yet sentimental feelings toward the instant moment of encounter conveyed through the CLOUD and REFLECTION metaphors.



In the second paragraph of version 3 as shown in Figure 23, the combination of *agitato*³¹, loud dynamic (*f*), and chord repetition in the piano accompaniment constructs a rapid and powerful tune for the SEA metaphor.

Figure 23. Version 3: Paragraph 2-Part 1

To some extent, the repeating chord in the piano part in this paragraph simulates the sound of waves lapping on the shore, which is rhythmic and steady like the beating drum. The above analysis indicates that this composer tends to focus more on the sound related to the SEA metaphor rather than the pictorial effect or visual scene of its image.

³¹ *Agitato*, ‘agitated,’ ‘restless;’ “the adjective indicates the character of a movement and a rather rapid time.”

的 方向; 你記得也 好, 最 好 你 忘 掉
 way you remember better you forget

Rest

在 這 交 會 時 互 放 的 光 亮! 我
 in this encounter exchanged glowing light

Figure 24. Version 3: Paragraph 2-Part 2

In Figure 24, before the final line “the exchanged glowing light in the encounter,” a rest comes in to highlight the whole final line. The final word “light” has a longer duration, accompanied by the repeating chord in the piano part that marks the end. We suggest the composer tends to emphasize the whole final line due to its significance in understanding the metaphorical meaning of the life encounter as exchanged glowing LIGHT.

5.2.4 Version 4

Of the four songs, this is the only example of a popular song composed for the poem “Serendipity.” The genre of popular music provides an interesting contrast to examine how poetry lyrics are transformed and represented in different musical modes.



Figure 25. Version 4: Paragraph 1

“The Sighing Figure”

歡欣在轉瞬間 消滅了踪影 轉瞬間消滅了踪影 你我相
逢 在黑夜的海上 你有你的 我有我的方向 你記得也好 最好你
忘 掉 在交會時互放的光亮

Figure 26. Version 4: Paragraph 2

In this version, as shown in Figures 25 and 26, the composer’s intention is not to interpret each metaphor and metaphorical expression in distinct composing patterns or musical techniques. Instead, we suggest the composer views the abstract metaphorical extension of the metaphor, the emotions and feelings that underlie the metaphors in poetry lyrics, as the main focus by adopting a classic musical technique, the “*sighing figure*,”³² which is derived from downward falling figures with narrow intervals as shown with falling lines in

³² In music history, the *sighing figure*, which is derived from the downward string motif, has been a composing pattern widely used since the Baroque era. It has been generally used to represent the sentimental emotion in classical music and modern music.

Figure 26, to capture the sentimental and sorrowful emotions in the poetry lyrics throughout the song.



Overall, the metaphor and musical analysis of the four versions demonstrate the significant difference between the classical artistic poetry music of versions 1, 2, and 3, and the popular music of version 4. While the artistic poetry music reinterprets and simulates the concrete image of each metaphor from the language mode to the musical mode in detail, the popular song expresses the gestalt emotion and feelings of the poetry through a more down-to-earth and catchy melody. In the following section of discussion, we will present our classification of the four versions in answering the research question.

5.3 Discussion

The results of our analysis indicate that two significant aspects, the concrete image and abstract metaphorical extension of metaphors in poetry lyrics, are the main focus of distinct composing patterns in four musical versions of a poem. We found that different composing patterns and musical techniques are used to highlight concrete images or abstract metaphorical extensions of the metaphors or metaphorical expressions in poetry lyrics. For

instance, composers use prominent aural cues, such as increasing or decreasing dynamic, repetitive rising-falling figures, high or low pitch, high or low key/notes, speed rate, dark or bright timbre, and shift from major to minor key in order to highlight the “concrete image” of the CLOUD, REFLECTION, SEA, and LIGHT metaphors. On the contrary, when the composer’s main focus is on the abstract metaphorical extensions, the connotations or inferences extended from the metaphors, they may use balanced melodies, soft tunes, narrow tone range, or sighing figures to interpret the sentimental feelings and emotions underlying the metaphors in the poem. Based on the main composing pattern and musical technique used in each song, we propose that four versions of songs can be distributed on the continuum between concrete images and abstract metaphorical extensions, as shown in Figure 27.

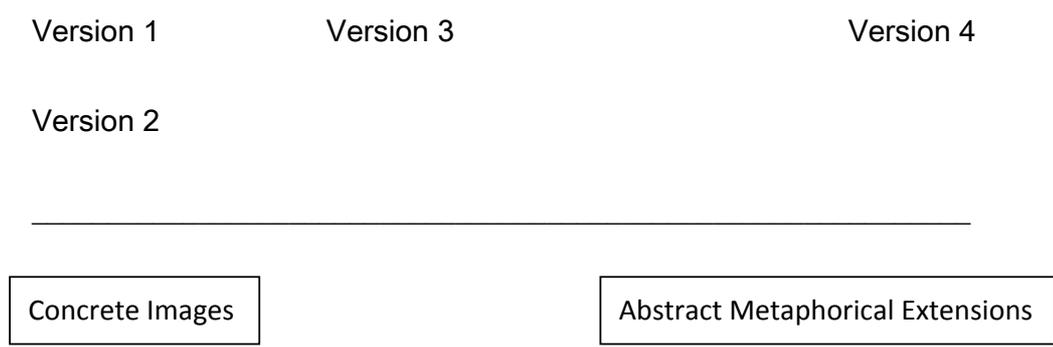


Figure 27. The Continuum between Concrete Images and Abstract Metaphorical Extensions

In general, this continuum provides a categorization of the concrete-abstract dimensions of composing patterns and musical techniques, indicating that language mode can be transformed into the music mode visually and aurally via each composer's highlights of two aspects, the concrete image and abstract metaphorical extension of metaphors. As Figure 27 shows, our analysis indicates that the concrete image of metaphors is highly valued and portrayed in versions 1 and 2, while it is partly portrayed in version 3.

To be more specific, versions 1 and 2 are considered the concrete versions that use prominent multimodal cues, including visual effect, such as pictorial effect or graphic notation on the score, and prominent musical features, such as change on dynamics, timbre, pitch, and figures, etc. to visually or aurally simulate the concrete image of the CLOUD, REFLECTION, SEA, and LIGHT metaphors; for instance, in versions 1 and 2, we found the pictorial effect of the CLOUD and SEA metaphors shown on the scores, which visually simulate the concrete image of cloud and ocean wave. The mirror-image of the REFLECTION metaphor (the cloud in the sky reflected on the wave) and space shift have been represented through *leap* (skip between the intervals) or the change from higher to lower pitch with similar patterns in the music mode; the image and rhythm of the SEA metaphor were

visualized and musicalized through *arpeggio* with increasing-decreasing volume. The glowing aspect of the LIGHT metaphor was highlighted with prominent musical features, such as *tremolo* with loud volume (*f*) or highest pitch with longer duration.

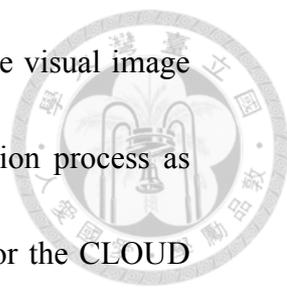
On the other hand, version 3 is considered a moderate version that includes both the concrete image and abstract metaphorical extensions in the song. In version 3, the song begins with a connected rhythm with a lower dynamic that does not correspond to the concrete image of the CLOUD and REFLECTION metaphors, yet the rhythm of the SEA metaphor was constructed with *agitato*, large dynamic, and chord repetition in the piano accompaniment. Unlike the previous versions, version 4 is the only song that mainly focuses on the abstract metaphorical extensions by adopting the musical technique of the sighing figure to capture the gestalt emotion and sentimental feelings conveyed through the metaphors in the poetry instead of portraying the concrete image of metaphors through prominent musical features. The continuum in Figure 27 demonstrates the significant difference between the classical artistic poetry music of versions 1, 2, and 3, and the popular music of version 4; versions 1, 2, and 3, as art songs, tend to interpret and depict the image of metaphor visually and aurally in the concrete

dimension, version 4, as a popular song, tends to express the gestalt emotion and metaphorical extensions that underlie the metaphors in the abstract dimension.



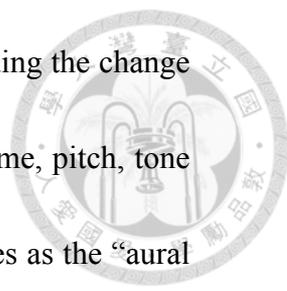
As a result, our analysis indicates that the metonymic-metaphoric network evolves through the “metonymic representations” of metaphor in poetry music. For example, the “pictorial effect” on the score assembles or simulates the concrete image of certain part of the source domains of the metaphor, for example, the visual shape or of cloud appears on the score to represent CLOUD metaphor. Besides the pictorial effect, the beating rhythm of the ocean waves occurs in the aural/musical mode to represent SEA metaphor. These examples have demonstrated the inseparable relationship between metonymy and metaphor in the multimodal genre of poetry music and displayed how part-for-whole metonymic representations of metaphor have been realized through the multimodal fusion of the visual, aural, and verbal mode.

Extending from the metonymic representations of metaphors, our analysis shows that the musical representations of metaphors are echoing and connecting with each other in the song, which we call the “chain of metaphor”. For example, the REFLECTION metaphor echoes the CLOUD metaphor with



similar notes in a lower pitch (version 1), creating a mirror-like visual image on the score, which visually and aurally simulates the reflection process as well. The splendid style of introduction (version 2) prepares for the CLOUD metaphor by creating the “Motif for cloud” with pictorial effect of the repetitive rising-falling figures that resembles the prototypical image of a cloud. The LIGHT metaphor (version 2) has been emphasized in the ending part with a splendid style of piano accompaniment that echoes with the introduction in the opening. As several examples show, we propose such an interactive and connecting relationship between metaphorical representations to be the “chain of metaphors” that helps to maintain the coherent and holistic characteristics of poetry music. We believe the “chain of metaphor” is more vividly and clearly displayed in the representations of multimodal metaphor in poetry music than the verbal metaphor in the language mode, which may be one of the significant features that distinguishes the difference between metaphor in multimodal (such as verbal-visual or verbal-aural mode) and monomodal mode.

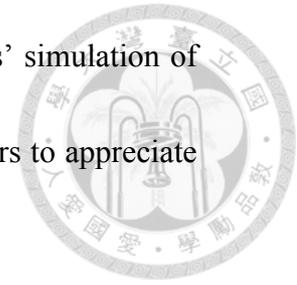
Furthermore, our findings have explored the embodied basis of multimodal fusion in poetry music, showing how the concrete image of metaphor and the abstract concept of emotion that underlies the metaphor have



been conceptualized through prominent musical features, including the change on rising and falling figures, major and minor key, speed, volume, pitch, tone range, timbre, etc. We consider these prominent musical features as the “aural cues” that activate listeners’ imagination and enable metaphor interpretation in poetry music.

Previous studies on embodiment and music provided ample evidence that the characterization of high and low musical “pitches” is basically metaphorical, as the conceptual metaphor, PITCH RELATIONSHIPS ARE RELATIONSHIPS IN VERTICAL SPACE, underlies the linguistic expressions of playing the higher or lower notes since ‘mapping up-down onto pitch allows us to import the concrete relationships through which we understand space into the domain of music, and thereby provide a coherent account of relationships between musical pitches’ (Zbikowski, 2000). Extending from this point, our analysis demonstrates that besides pitches, the concrete relationship of space can not only be mapped onto pitchers but also onto the prominent musical features, such as increasing and descending figures, lower and higher key, notes, and tone range in the musical domain. Our findings suggest that the high-low representations of the prominent musical features not only realize the conceptualization of space in poetry

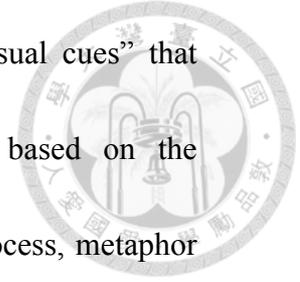
music but also function as the aural cues that activate listeners' simulation of the image of metaphors in poetry lyrics and thus enable listeners to appreciate the emotions conveyed through poetry music.



Extended from the embodied basis of the orientational metaphor HAPPY IS UP and music metaphor HIGH IS BRIGHT, LOW IS DARK, we also found the scale of emotion can be mapped onto the representations of prominent musical features in poetry music. For example, increasing dynamic, higher key, and bright timbre are often associated with positive emotions like expectation, excitement, and enthusiasm as shown in the representation of the LIGHT metaphor in versions 1 and 2. On the contrary, the lower dynamic, lower pitch, and dark timbre are often associated with more oppressed and sentimental emotions as shown in the representation of the SEA metaphor in versions 1 and 2.

In addition, we discovered that the pictorial effect on the score assembles or simulates the concrete image of certain source domains of the metaphor, such as the shape of cloud or ocean, which can be viewed as the “visual cues” to simulate the concrete image of the source domain being depicted and activating the performer/singer's imagination and interpretation of the metaphor (when they were performing with the notes).

To summarize, we claim that the “aural cues” and “visual cues” that underlie the composing patterns and musical techniques based on the composer’s arrangement may be linked to the imagination process, metaphor interpretation process, and multimodal aesthetic experience combining the verbal, visual, and aural modes while appreciating the emotion conveyed through poetry music.



Chapter 6 Audience Response to Multimodal Fusion in Poetry Music



In this chapter, we will present significant results of audience response to multimodal fusion in poetry music. We have conducted surveys with participants to investigate how they interpret metaphor and feel the musical effects and aesthetics created from multimodal fusion in poetry music. We will incorporate qualitative analysis and quantitative analysis from our surveys to provide a detailed account combining the general pattern and specific variations of the audience response.

6.1 Questionnaire Survey: Participants' Response

To conduct the survey, we designed the questionnaire to investigate how the “trained” participants process the transformation from poetry to music by recognizing the musical representations of “concrete images” or “abstract metaphorical extensions” of the metaphors in the poetry lyrics during the interpretation process. In total, 20 participants were interviewed in this survey. The participants are undergraduate students from National Taiwan University and Yuan Ze University, and they were trained with instructions on the basic concept of “concrete images” and “abstract metaphorical extensions” in music.

6.2 Survey Procedures



In this section, the survey procedures are introduced as follows:

First, metaphors are identified and highlighted in the poetry lyrics.

Second, the poetry music is divided into two parts, Part 1 and Part 2, to accurately evaluate the responses to each metaphorical expression and to avoid participants' tiredness caused by the length of the songs. The metaphors in Part 1 and Part 2 are listed in Table 10 as follows:

Table 10. Metaphors in “Serendipity”

	Metaphors
Part 1	Cloud (雲)
Part 1	Reflection (投影)
Part 2	Sea of dark night (黑夜的海上)
Part 2	Exchanged glowing light (交會的光亮)

Third, 4 versions of songs composed for “Serendipity” were edited into Part 1 and Part 2. Each part of each version had been played twice for the participants during the survey.

Fourth, the participants were asked to evaluate the degree of the

preference and matches of the metaphors in answering seven questions for

Part 1 and Part 2 in each version of the song.



The survey was conducted in Mandarin Chinese (For further details,

please see Appendix 2 for the original version of the poetry music survey

questionnaire). We will provide the translated English version of the main

seven questions in the following:

1. Based on your interpretation, does the music of “cloud” in “I am a cloud in the

sky” match its concrete image?

Yes, it extremely does. Yes, it does. Neutral No, it doesn't. No, totally not.

(Please explain how and why)

2. Based on your interpretation, does the music of “reflection” in “by chance

reflected on the wave of your heart” match its concrete image?

Yes, it extremely does. Yes, it does. Neutral No, it doesn't. No, totally not.

(Please explain how and why)

3. Based on your interpretation, does the music of “cloud” in “I am a cloud in the

sky” match its abstract metaphorical extension?

Yes, it extremely does. Yes, it does. Neutral No, it doesn't. No, totally not.

(Please explain how and why)

4. Based on your interpretation, does the music of “reflection” in “by chance reflected on the wave of your heart” match its abstract metaphorical extension?

Yes, it extremely does. Yes, it does. Neutral No, it doesn't. No, totally not.

(Please explain how and why)

5. Based on your interpretation, does this song appropriately convey the emotion in this poem?

Yes, it extremely does. Yes, it does. Neutral No, it doesn't. No, totally not.

(Please explain how and why)

6. Have you ever heard of this song?

Yes, I have. No, I haven't.

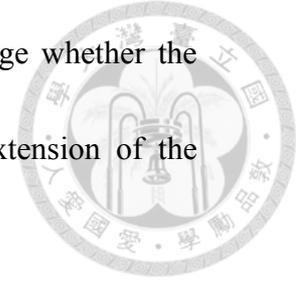
7. Do you like this song?

Yes, I like it very much. Yes, I do. Neutral No, I don't.
 No, totally not.

(Please explain why)

The above seven questions are the set of questions included with each part (Part 1 and Part 2) of the song. In total, four versions of songs (eight paragraphs) are included in the survey. The “metaphor” marked in shade will be switched accordingly in Part 1 and Part 2 as shown in Table 10. In each version of song,

the participants need to listen to each part of the song and judge whether the music matches the concrete image or abstract metaphorical extension of the “metaphor”.



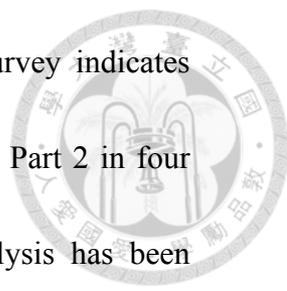
As a result, we have conducted the survey by interviewing 20 participants³³ to examine how they interpret metaphor in four versions of poetry music based on the poetry lyrics of “Serendipity”. In the next section, we will present the significant findings of the survey.

6.3 Results and Discussion

In this section, we will present the quantitative analysis to account for a general pattern from the statistical data, and we will further evaluate and examine the general findings and distribution of the statistical analysis with the qualitative analysis from our survey.

The statistical data of poetry music survey is conducted with SPSS analysis. The reliability statistics (Cronbach’s Alpha= 0.839) indicates that the reliability of survey is high based on the percentage of 83.9. This suggests that the participants’ answers are reliable and the design of the questionnaire is well arranged.

³³ Previous studies in the field of music and multimodality have mainly focused on the analysis and theoretical framework, however, very few studies have conducted interview surveys to collect audience response in this area of study. According to the reliability statistical data, the number of 20 participants has been considered a reliable amount of participants for the qualitative and quantitative results in this study.



First, the one-way ANOVA analysis of poetry music survey indicates significant differences among the eight paragraphs (Part 1 and Part 2 in four versions of songs) in the survey. Specifically, the SAS analysis has been conducted to evaluate how the participant responded to each question in the comparison among four versions of songs. In general, the SAS statistical analysis indicates several significant findings:

- (1) There are significant differences in the matching level of the concrete image of “CLOUD metaphor” among four versions of songs.
- (2) There are significant differences in the preference scores among four versions of songs.

In the following, we will discuss the above general findings of the statistical analysis and incorporate the quantitative analysis with the qualitative analysis from our survey.

6.3.1 Matching Level of the Concrete Image

In this survey, we ask the participants to evaluate the matching level of the concrete image of CLOUD, REFLECTION, SEA IN THE DARK NIGHT,

and LIGHT metaphor in the song. However, we only found significant differences in the matching level of the concrete image of “CLOUD” metaphor among four versions of songs. As summarized in Table 11, the statistical analysis indicates that P values are under 0.05. The results suggest that participants’ responses to the CLOUD metaphor have significant differences between version 1 and the other versions of songs. In particular, version 1 and version 3 has the mean difference of 3.01832, showing that the participants’ responses are the most different in version 1 and version 3.

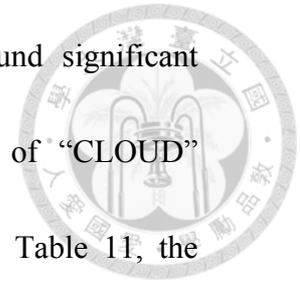
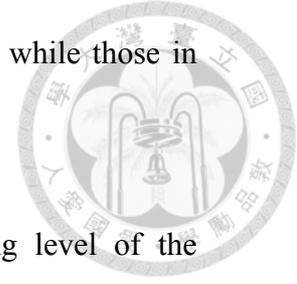


Table 11. The Statistical Data Summary of the Significant Differences in “CLOUD” Metaphor

Version 1	Version 2	Version 3	Version 4
P Value	0.046	0.003	0.0055
Mean Difference	2.0122	3.01832	2.8171

The statistical analysis may correlate with our analysis since we found that version 1 and version 3 may be distributed on the continuum between the concrete image and abstract metaphorical extension. Our analysis suggests that

the musical effects in version 1 highlight the concrete image, while those in version 3 focus more on abstract metaphorical extensions.

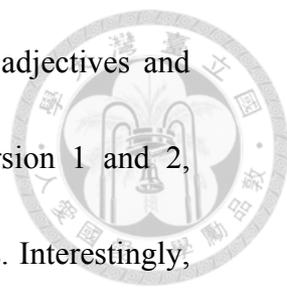


When we ask the participants to evaluate the matching level of the concrete image of CLOUD in the song and further describe the concrete image, the participants provide various detailed descriptions and interesting comments. Table 12 summarized participants' descriptions of the concrete image of CLOUD in four versions of songs based on the qualitative analysis of the survey.

Table 12. Participants' Descriptions of the Concrete Image of CLOUD

	Participants' description of the concrete image of CLOUD in the song
Version 1	<p>Soft, tender, unsteady, light, airy, floating, pure, continuing, heavy, the association between high note and sky, the association between high note and floating cloud up in the sky</p> <p>(柔軟，輕柔的，飄浮，飄忽不定，軟綿綿，輕輕的，輕盈，飄飄的，很飄，純淨，連續的，沉重的，高音聯想到天空，高音很像是雲飄在空中)</p>

<p>Version 2</p>	<p>Softer, lighter, tender, floating, smooth, silky, the cloud changes its shape at every second, piano accompaniment is like the change of cloud, the dynamic floating of cloud, a cloud floating on top of the sky, piece by piece (the shape of cloud)</p> <p>(更柔，更輕的感覺，柔軟，雲飄盪在空中，柔順連綿，雲的形狀變化萬千，鋼琴伴奏很像是雲的變化，動態地飄著移動的雲，天頂浮著一朵雲，一片一片的雲)</p>
<p>Version 3</p>	<p>Tender, leisurely, slowly, gentle, a piece of cloud, continuing, floating, peaceful, comfortable, balanced emotion</p> <p>(柔軟，悠閒的，緩慢的，溫柔的，一整片雲，連續的，飄著的雲，祥和的，舒服的，剛剛好不悲不喜的情緒)</p>
<p>Version 4</p>	<p>Lively, soft and silky, emotional, like the image of a god riding on the cloud and fog in the sky</p> <p>(生動活潑，柔軟綿延，情感豐富的，好像神在天上騰雲駕霧)</p>



As Table 12 shows, we found participants tend to use more adjectives and analogies to describe the concrete image of CLOUD in version 1 and 2, especially in the aspect of its shape, texture, weight, and status. Interestingly, the participants' responses echo with our analysis that the musical effects of version 1 and 2 focus more on the concrete image of metaphors than the abstract metaphorical extensions.

On the contrary, we found most participants give fewer descriptions of the image but describe more about the aspect of atmosphere, feelings, and emotion conveyed in version 3 and 4 (although we were asking them to comment on the concrete image). This also echoes with our analysis that we found the musical effects of version 3 and 4 highlight the abstract metaphorical extensions of feelings and emotion conveyed in the poetry.

To summarize, we believe the correspondence between the participants' responses and our analysis does not happen by chance. While the participants were interpreting the metaphors, they seemed to distinguish different musical effects from multimodal fusion of verbal and aural mode unconsciously and automatically. The audience responses provide us the chance to understand how interpretation may work at the communication among different versions of poetry music, the poem itself, and its listeners. On top of that, the

participants' responses have broadened our horizon by showing us different perspectives and rich interpretations beyond the limitations of our analysis.

However, more participants need to be examined to search for more relevant findings based on our analysis for further studies.



6.3.2 Preference among Four Versions of Songs

In this survey, we asked the participants to evaluate their preference for each version of song. The statistical data shows significant differences between version 4 and the other three versions in Part 2. As shown in Table 13, P values under 0.05 suggest that participants' preference scores show significant differences.

Table 13. The Statistical Data Summary of Preference Score

Version 4	Version 1	Version 2	Version 3
P Value	0.0008	0.0049	0.0085
Mean	3.429185	2.857654	2.667144
Difference			

The result may be attributed to the fact that version 4 is the only popular song which distinguishes itself from the style of art songs in version 1, 2, 3. Our



qualitative analysis reveals that more than 50 % of the participants prefer listening to version 4, describing that they “enjoy the melody”, “feel more relaxed”, “prefer catchy tunes”, “version 4 reminds me of the pop songs playing in my parents’ car”, “version 4 is like a movie soundtrack”, “it’s like the song playing for the soap opera”, “it has nostalgic and sentimental atmosphere”, etc. However, about 20 % of the participants criticize that they are “not touched by version 4” since it does not capture the meanings in the poetry as the previous versions do. Some participants even mentioned that the musical style of version 4 is “too commercialized and entertaining” thus it fails to convey the poetic feelings and atmosphere. These responses show that the correspondence between the connotation of the poetry lyrics (verbal mode) and music (aural mode) may indirectly affect the audience’s preference for the song. As a result, the significant differences between version 4 and the other versions reveal the fact that distinct musical style of pop song and art song each has its own supporters and fans.

6.3.3 Multimodal Fusion and Interpretation: Different Focuses on the



Multimodal and Conceptual level

Based on the above findings, we found that the participants are capable of distinguishing the musical effects of the concrete image and abstract metaphorical extensions through multimodal fusion in poetry music. When the participants interpret the metaphors in a more concrete version such as version 1 and 2, they naturally notice and describe the concrete aspect of shape, texture, weight, and status of the metaphor. On the other hand, when they interpret the metaphors in a more abstract version such as version 3 and 4, they unconsciously express the atmosphere, feelings, or emotion conveyed through the metaphor. Moreover, we found most participants are aware of the aural cues, including the high or low volume, pitch, and rising or falling figures while they pinpointed several prominent musical features when they were trying to describe the concrete image of the metaphors during the interviews.

In view of multimodal fusion in poetry music, we propose that the concrete image of the metaphor is closely related to the multimodal level (visual/aural mode), while the abstract metaphorical extensions of the

metaphor is related to the conceptual level. To clarify, we believe that four versions of poetry music clearly demonstrate different focuses on the multimodal level and conceptual level of multimodal fusion. Each version of song displays a unique combination of multimodal level and conceptual level in multimodal fusion, showing us how abstract concept and emotion in a poem may be conceptualized and construed in various ways with different focuses.

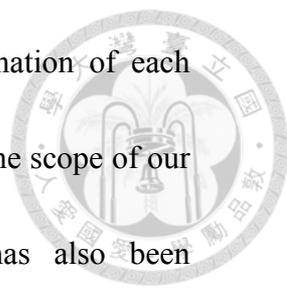
Moreover, the correspondence between the verbal mode and musical mode may be relevant to the audience's preference for the song. The audience preference scores sincerely reflects how the participants "feel" about the musical and aesthetics effects created from the multimodal fusion in different versions of songs through their interpretations on poetry music.

Besides noticing the concrete image interpreted through the music, we believe another visual aspect, the "pictorial effect" on the musical scores, should be included in participants' response in future studies. However, it remains difficulties in the current study since the majority of the participants cannot read the musical scores, which is hard to be trained or instructed in a short period.

Chapter 7 Conclusion



To conclude, this study proposes an interactive interpretation approach of “multimodal fusion model” that incorporates the conceptual level and multimodalities, including the verbal, visual, and aural modes due to the fact that the conceptual level, the verbal, visual, and aural modes are all encapsulated and integrated simultaneously in real multimodal discourse, such as in the genre of political cartoons and poetry music. In this study, multimodal fusion is an interactive interpretation model induced from the cross-comparison of the multimodal corpus from our political cartoon and poetry music data. Unlike previous studies of blending theory that mainly focused on the mental processing of the one-on-one correspondence, we have provided a modest proposal with more flexibility and possibilities to account for the complex relations and aspects of the unique one-on-many relationships between metaphor, metonymy, and multimodality in specific genre. Through our analysis, we have demonstrated how “multimodal fusion model” serves as the theoretical basis to analyze the general pattern and specific features underlie the complex relations between metaphor, metonymy, and multimodality.



Our analysis has been elaborated in the detailed examination of each piece of data (i.e. political cartoons and musical scores) under the scope of our model. The cross-comparison analysis among the data has also been conducted to provide the general account and distribution of the data in our model. In addition, we have illustrated how certain topic and concept (i.e. politics, emotion, space, etc.) are conceptualized and construed differently in the multimodal genre accompanied with the use of prominent multimodal cues (visual/aural/verbal cues) that facilitate multimodal realization and interpretation.

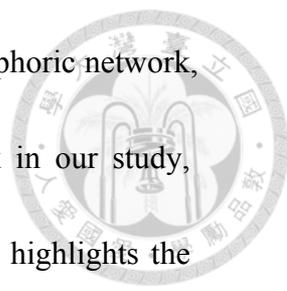
Our findings indicates that multimodal cues, such as verbal, visual, and aural cues, play significant parts in facilitating the interaction between different modes and the conceptual level, thus we claim that the multimodal cues need to be examined specifically in each mode while interpreting the multimodal data. Furthermore, our analysis suggests that the contrasting visual/aural features within the visual/aural mode reflected from the data, for example from the cartoon image or the musical score, are often the crucial elements that influence how certain topic or concept has been conceptualized in distinct media or versions of works.

In the following sections, we will recapitulate and summarize the main findings based on the evidence from political cartoon and poetry music.



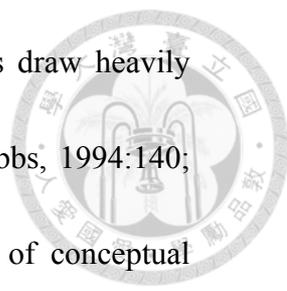
7.1 Multimodal Fusion in Political Cartoon

In this study, we have discovered several distinguishing features of political cartoons, summarized as follows. Through multimodal fusion model, we have shown the important role of “metonymy” in the multimodal context of political cartoons, demonstrating how metonymy and metaphors are interwoven in the process of multimodal fusion. Specifically, this study shows how metonymy-metaphor interaction has been realized in the multimodal genre of political cartoons through the cognitive mechanism “multimodal fusion”, which has evolved from metonymic-metaphoric networks. In related metonymic networks, relevant metonymies represented in visual images are found frequently and are productively used to conceptualize the main event through prominent visual features, such as a huge bull or a fierce bull conceptualized as the metonymy for the U.S. beef event. Most importantly, these metonymies often lead to the metaphorical mappings of conceptual scenarios related to POLITICS IS GAME and POLITICS IS WAR. Therefore, the metonymies and metaphors cannot be separated in the creation and



realization of political cartoons. In addition, in diversified metaphoric network, shown as the second type of metonymic-metaphoric network in our study, verbal mode serves as the prominent cue and crucial link that highlights the main focus and connects the main topic with visual image and conceptual scenario. We believe these two types of multimodal fusion evolving from metonymic-metaphoric network not only naturally underlie the genre of political cartoon but also the other multimodal genres combining visual, verbal, and conceptual mode.

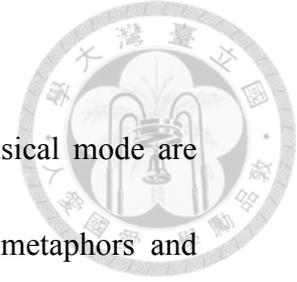
Moreover, this study shows that the cognitive mechanism of multimodal fusion with the fundamental basis of related metonymic network and diversified metaphoric network is a significant and recurrent representation technique in the unique genre of political cartoon. The representation of multimodal fusion demonstrates how abstract and complex political debates can be encapsulated and condensed through the combination of conceptual, visual, and verbal modes in political cartoons. We believe that the metonymic-metaphoric network of multimodal fusion naturally exists in the genre of political cartoons due to its cognitive function of encapsulating complex knowledge and critical messages efficiently with irony and humorous effect, which are the characteristics unique to this genre.



Though previous studies found many political metaphors draw heavily and systematically on the language of sports and warfare (Gibbs, 1994:140; Semino & Masci, 1996), very few studies explored the use of conceptual metaphors of politics in multimodal mode. By examining our multimodal corpus, we discovered that POLITICS AS GAME and POLITICS AS WAR are the most identified metaphor used in political cartoons. This study proves that the abstractness of politics is preferably and productively conceptualized through GAME and WAR domains in visual, verbal, and conceptual modes.

On top of that, we found the critical messages and distinct stances of the political cartoons in LT and UDN are highlighted and contrasted through prominent visual features shown in the participants in the cartoons. The contrast between prominent visual features, including comparisons and contrasts between facial expressions, depicted sizes, states, positions, and materials as well as highlighted by visual and verbal details, helps clarify the complex power relationships and the distinct political stances of the newspapers. In particular, we believe the prominence of size contrasting resonates with the primary metaphor, SIGNIFICANCE IS SIZE, by emphasizing the importance of certain figure or event through the larger size in political cartoons.

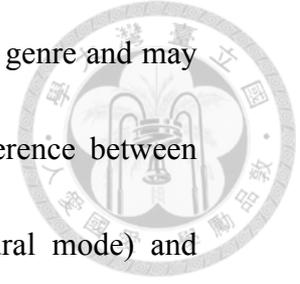
7.2 Multimodal Fusion in Poetry Music



By providing a detailed analysis on how verbal and musical mode are interactively combined and contrasted to represent specific metaphors and metaphorical expressions in poetry music, this study claims that the poem as poetry lyrics has been transformed into the musical mode visually and aurally via each composer's highlights of two aspects, the "concrete image" and the "abstract metaphorical extension" of metaphors in view of the composing patterns and musical techniques. To conclude, we found that four musical versions of the same poem are distributed on the continuum between the concrete and abstract dimensions of metaphorical representations in the musical mode, which marks the distinction between the genre of poetry music and popular song.

To recapitulate, the significant findings in poetry music are as follows. First, we discovered the interactively echoing relationship between metaphors in poetry music, the "chain of metaphors", which highlights and connects the concrete image of metaphors, such as CLOUD, REFLECTION, and LIGHT. The chain of metaphor has been reflected through the echoing musical features represented by each metaphor, creating a holistic representation of multimodal metaphor combining verbal and musical mode in poetry music. We suggest

that the chain of metaphor may exist mainly in the multimodal genre and may be one of the significant features that distinguishes the difference between metaphor in multimodal (such as verbal-visual or verbal-aural mode) and monomodal mode.



Second, we have found metonymic representations of metaphors in the examples of partial visualization of cloud image on the score representing CLOUD metaphor and aural simulation of wave rhythm representing SEA metaphor, etc. Our findings show that such metonymies representations demonstrate the part-for-whole characteristics in building the metonymic association through visual, aural, verbal, and conceptual level of multimodal fusion, which highlights the fact that metonymy is indistinguishable with its counterpart metaphor as the metonymy-based metaphors in poetry music. Our findings indicate that multimodal fusion starting with metonymic-metaphoric network representing the concept that activate the conceptual level from poetry lyrics and combine the visual, aural, and verbal mode based on the metonymic associations and metaphorical mapping in different versions of songs.

Third, viewing from the embodied perspective, our findings demonstrate that not only the concrete relationship of space but also the abstract attributes

of emotion can be mapped onto the representations of prominent musical features, such as increasing and descending figures, lower and higher keys/notes, and wide or narrow tone range in the music domain.



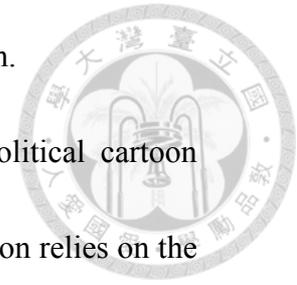
Fourth, this study indicates that the prominent musical features, such as aural cues, and pictorial effect, or graphic notation on the scores as visual cues, both facilitate metaphor realization and transformation from verbal mode to the verbal-musical mode in poetry music.

7.3 Audience Interpretation and Multimodal Fusion

In this study, we have incorporated the perspectives of analyzer and audience by examining the “audience responses” with our analysis of multimodal fusion model to extend the discourse-internal analysis to an interactional level of interpretation analysis on multimodality and metaphor. Questionnaire surveys and interviews have been conducted to investigate two main aspects (1) How the audience interprets the conceptualization of political issues concerning U.S. beef imports through multimodal fusion of verbal and visual modes in political cartoon? (2) How the audience interprets the transformation of metaphor and musical effects through multimodal fusion of verbal and aural/musical modes in poetry music? The quantitative and qualitative analysis of political cartoon survey are combined to account for the

general pattern and specific variations of audience interpretation.

Based on the quantitative and qualitative analysis of political cartoon surveys, we concluded that higher level of audience interpretation relies on the effective recognition, comprehension, and connection of the verbal mode, visual mode, and conceptual level of multimodal fusion in political cartoons. In other words, the results indicate that effective audience interpretation depends on the recognition in the visual mode (i.e. politicians' caricatures), the comprehension of the connotations in the verbal mode (i.e. the meaning of "CODEX"), and the connection between the multimodal level (visual and verbal modes) and the conceptual level in multimodal fusion. In addition, we found that audience interpretation on the conceptual scenario is influenced by prominent multimodal cues of the visual and verbal mode, which are inseparable with the conceptual level. The findings also indicate that different conceptual scenarios of the conceptual level may be activated and highlighted for different audience. Viewing from our analysis and audience response in political cartoon, we concluded that various conceptual scenarios are involved on the conceptual level of multimodal fusion, while audience may view specific conceptual scenario as more prominent and significant based on their interpretation.





Regarding the quantitative and qualitative analysis of poetry music surveys, we concluded that audience response supports our claim that each version of songs is in fact displaying the unique combination of multimodal level and conceptual level in multimodal fusion, demonstrating how the abstract concept in the poem can be conceptualized and construed in various ways with distinct focuses and musical effects. The results show that the audience is capable of distinguishing different musical effects of the concrete image of “CLOUD” metaphor in four versions of songs, which correlates with our analysis of the “concrete-abstract continuum”: when the audience were interpreting the metaphor in a more concrete version of song, they unconsciously described the concrete aspect of shape, texture, weight, and status of CLOUD metaphor, whereas when the audience were interpreting the metaphor in a more abstract version of song, they focused more on the abstract aspect of the atmosphere, feelings, and emotion conveyed through CLOUD metaphor. Moreover, we also found that the audience’s preference for certain song is relevant to the correspondence between the verbal mode and aural/musical mode. The results pinpointed that the audience interpretation survey on poetry music reflected how the audience “process” the musical effect and “appreciate” the aesthetic feelings through multimodal fusion in

different versions of songs composed for the poem that they have been familiar with.



7.4 Concluding Remarks

In conclusion, this study shows that the novel conceptualization and abstract reasoning of certain topic in the genre of political cartoons and poetry music has been realized through multimodal fusion on the fundamental basis of metonymic-metaphoric network connected by metonymic associations and metaphorical mappings. For further studies, we hope to expand the research scope in the following aspects:

- (1) To include more participants with different background for the in-depth analysis of audience response to multimodal data
- (2) To conduct surveys to interview the authors, creators, and performers of the multimodal genre, such as the composers, cartoonists, writers, and artists, etc. to provide a more dynamic analysis that accounts for the real world situation of multimodal interpretation process from the perspectives of the audience, author, and researcher
- (3) To collect data from other multimodal genres to further explore how the one-on-many relations and other complex relations between multimodality and conceptual level may be represented and realized in other genres

through multimodal fusion of different modalities and mediums

- (4) To develop the experiments conducted for multimodal analysis and audience interpretation from the perspective of interdisciplinary studies in the field of cognitive science and psychology

Through the investigation of the interaction between the interpretation process of audience response and multimodal fusion model, we hope to revise and refine the model of multimodal fusion through careful examination on data analysis, to broaden the scope of interdisciplinary figurative language studies on unique multimodal genre, and to shed light on the interface between figurative language and multimodality.

With a special concern on the critical effects and pragmatic functions created from multimodal fusion, we believe it would be insightful to explore other critical effects, pragmatic functions, and aesthetic purposes in future studies. We hope to shed light on interdisciplinary studies of the interface between the cognitive mechanism and pragmatic multimodal use of language, picture, music, and the aesthetic experience of the multimodality appreciation process.

Appendix 1: Word-to-word Translation of the Poem ‘Serendipity’:



我-是-天空-裡的-一片-雲，

I -am-sky-in-a-cloud

I am a cloud in the sky

偶然-投影-在-你的-波-心，

by chance-reflected-on-your-wave-heart

by chance reflected on the wave of your heart.

你-不必-訝異，

You-don't need to be-surprised

Don't be surprised,

無需-歡喜，

nor to be-elated

or too elated;

在-轉瞬間-消-滅了-蹤影，

In-an instant-vanish-without-trace

In an instant I shall vanish without trace.

你-我-相逢-在-黑夜-的-海上，

You-I-meet-on-dark night-sea

We meet on the sea of dark night,

你-有-你的，我-有-我的方向，

You-on-your, I-on-my way

You on your way, I on mine.

你-記得-也好，

You-remember-fine,

Remember if you will,

最好-你-忘掉 ,

better-you-forget

Or, better still, forget

在-這-交會時-互-放-的光亮

In-this-encounter-exchanged-glowing-light

The glowing light exchanged in this encounter.

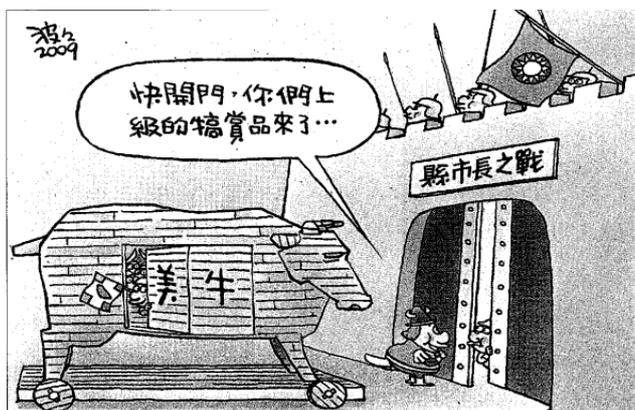


Appendix 2: Political Cartoon Survey Questionnaire



政治漫畫研究問券調查

(1)



1. 依照你的解讀，是否容易理解此幅漫畫中針對美牛事件的觀點？
 非常容易 容易 中等 困難 非常困難 (請簡述原因)
2. 依照你的解讀，你認為此幅漫畫可能刊登於哪家報紙？(請簡述原因)
 自由時報 聯合報 不知道
3. 承上，在你的解讀過程中，木馬屠城記(特洛伊戰爭 Troy)概念是否有出現在你的腦中？(請簡述原因)
 當然有 有 不知道 沒有 完全沒有



(2)



4. 依照你的解讀，是否容易理解此幅漫畫中針對美牛事件的觀點？(請簡述原因)

非常容易 容易 中等 困難 非常困難

5. 依照你的解讀，你認為此幅漫畫可能刊登於哪家報紙？(請簡述原因)

自由時報 聯合報 不知道

6. 承上，在你的解讀過程中，**登山/走山路**的概念是否有出現在你的腦中？(請簡述原因)

當然有 有 不知道 沒有 完全沒有



(3)



7. 依照你的解讀，是否容易理解此幅漫畫中針對美牛事件的觀點？(請簡述原因)

非常容易 容易 中等 困難 非常困難

8. 依照你的解讀，你認為此幅漫畫可能刊登於哪家報紙？(請簡述原因)

自由時報 聯合報 不知道

9. 承上，在你的解讀過程中，**馬戲團**的概念是否有出現在你的腦海中？(請簡述原因)

當然有 有 不知道 沒有 完全沒有

(4)



10. 依照你的解讀，是否容易理解此幅漫畫中針對美牛事件的觀點？(請簡述原因)

非常容易 容易 中等 困難 非常困難

11. 依照你的解讀，你認為此幅漫畫可能刊登於哪家報紙？(請簡述原因)

自由時報 聯合報 不知道

12. 承上，在你的解讀過程中，牛仔的概念是否有出現在你的腦中？(請簡述原因)

當然有 有 不知道 沒有 完全沒有

(5)



13. 依照你的解讀，是否容易理解此幅漫畫中針對美牛事件的觀點？(請簡述原因)

非常容易 容易 中等 困難 非常困難

14. 依照你的解讀，你認為此幅漫畫可能刊登於哪家報紙？(請簡述原因)

自由時報 聯合報 不知道

15. 承上，在你的解讀過程中，**孤島漂流**的概念是否有出現在你的腦中？(請簡述原因)

當然有 有 不知道 沒有 完全沒有



(6)



16. 依照你的解讀，是否容易理解此幅漫畫中針對美牛事件的觀點？(請簡述原因)

非常容易 容易 中等 困難 非常困難

17. 依照你的解讀，你認為此幅漫畫可能刊登於哪家報紙？(請簡述原因)

自由時報 聯合報 不知道

18. 承上，在你的解讀過程中，**打獵**的概念是否有出現在你的腦中？(請簡述原因)

當然有 有 不知道 沒有 完全沒有



(7)



19. 依照你的解讀，是否容易理解此幅漫畫中針對美牛事件的觀點？

(請簡述原因)

非常容易 容易 中等 困難 非常困難

20. 依照你的解讀，你認為此幅漫畫可能刊登於哪家報紙？(請簡述原因)

因)

自由時報 聯合報 不知道

21. 承上，在你的解讀過程中，**戰馬**的概念是否有出現在你的腦中？(請

簡述原因)

當然有 有 不知道 沒有 完全沒有

(8)



22. 依照你的解讀，是否容易理解此幅漫畫中針對美牛事件的觀點？(請簡述原因)

非常容易 容易 中等 困難 非常困難

23. 依照你的解讀，你認為此幅漫畫可能刊登於哪家報紙？(請簡述原因)

自由時報 聯合報 不知道

24. 承上，在你的解讀過程中，**拆解炸彈**的概念是否有出現在你的腦中？(請簡述原因)

當然有 有 不知道 沒有 完全沒有



(9)



25. 依照你的解讀，是否容易理解此幅漫畫中針對美牛事件的觀點？(請簡述原因)

非常容易 容易 中等 困難 非常困難

26. 依照你的解讀，你認為此幅漫畫可能刊登於哪家報紙？(請簡述原因)

自由時報 聯合報 不知道

27. 承上，在你的解讀過程中，三隻猴子(非禮勿聽,非禮勿視,非禮勿聞)的概念是否有出現在你的腦中？(請簡述原因)

當然有 有 不知道 沒有 完全沒有

(10)



28. 依照你的解讀，是否容易理解此幅漫畫中針對美牛事件的觀點？(請簡述原因)

非常容易 容易 中等 困難 非常困難

29. 依照你的解讀，你認為此幅漫畫可能刊登於哪家報紙？(請簡述原因)

自由時報 聯合報 不知道

30. 承上，在你的解讀過程中，補破網(台語厘語:意指修補難解的問題)的概念是否有出現在你的腦中？(請簡述原因)

當然有 有 不知道 沒有 完全沒有



31. 綜合以上政治漫畫的解讀經驗，你認為用娛樂性比喻代表政治議題的比例為何？

非常高 高 中等 低 非常低

32. 綜合以上政治漫畫的解讀經驗，你認為用戰爭性比喻代表政治議題的比例為何？

非常高 高 中等 低 非常低

受訪者基本資料

姓名：

性別：

年齡：

系所：

個人是否有政治興趣/政治漫畫相關背景：(請簡述之)

～謝謝您協助本研究問券調查～

Thank you very much!

Appendix 3: Poetry Music Survey Questionnaire



藝術歌曲問券調查

以下將依序播放以《偶然》譜成的歌樂作品，一共有**四種版本的音樂**：

- 請特別注意詩中的標粗網底楷體字「**雲， 投影在你的波心**」
- 請您在聽完各版本的音樂後，依您個人感受勾選符合您所感知的選項：

〔 第一段詩詞 〕

我是天空裡的一片**雲**，
偶爾**投影在你的波心**——
你不必訝異，更無須歡喜——
在轉瞬間消滅了蹤影。

版本一

1. 您認為「我是天空裡的一片**雲**」這句音樂中的「**雲**」符合其具體形象嗎？

非常符合 符合 中等 不符合 非常不符合 (請簡述原因)

2. 您認為「偶爾**投影在你的波心**」這句音樂中的「**投影在你的波心**」符合其具體形象嗎？

非常符合 符合 中等 不符合 非常不符合 (請簡述原因)

3. 您認為「我是天空裡的一片**雲**」這句音樂中的「**雲**」符合其抽象意涵嗎？

非常符合 符合 中等 不符合 非常不符合 (請簡述原因)

4. 您認為「偶爾**投影在你的波心**」這句音樂中的「**投影在你的波心**」

符合其抽象意涵嗎？

非常符合 符合 中等 不符合 非常不符合

(請簡述原因)

因)

5. 你認為這段歌曲是否適切地表達出這段詩傳達的情感？

非常適切 適切 中等 不適切 非常不適切 (請簡述原因)

因)

6. 你是否聽過這段歌曲？

是 否

7. 你喜歡這段歌曲嗎？

非常喜歡 喜歡 中等 不喜歡 非常不喜歡 (請簡述原因)

版本二

8. 您認為「我是天空裡的一片雲」這句音樂中的「雲」符合其具體形象嗎？

非常符合 符合 中等 不符合 非常不符合 (請簡述原因)

9. 您認為「偶爾**投影在你的波心**」這句音樂中的「**投影在你的波心**」符合其具體形象嗎？

非常符合 符合 中等 不符合 非常不符合 (請簡述原因)

10. 您認為「我是天空裡的一片雲」這句音樂中的「雲」符合其抽象意涵嗎？

非常符合 符合 中等 不符合 非常不符合 (請簡述原因)

11. 您認為「偶爾**投影在你的波心**」這句音樂中的「**投影在你的波心**」符合其抽象意涵嗎？

非常符合 符合 中等 不符合 非常不符合 (請簡述原因)

12. 您認為這段歌曲是否適切地表達出這段詩傳達的情感？



非常適切 適切 中等 不適切 非常不適切

(請簡述原因)

因)

13. 你是否聽過這段歌曲？

是 否

14. 你喜歡這段歌曲嗎？

非常喜歡 喜歡 中等 不喜歡 非常不喜歡 (請簡述原因)



版本三

15. 您認為「我是天空裡的一片雲」這句音樂中的「雲」符合其具體形象嗎？

非常符合 符合 中等 不符合 非常不符合 (請簡述原因)

16. 您認為「偶爾投影在你的波心」這句音樂中的「投影在你的波心」符合其具體形象嗎？

非常符合 符合 中等 不符合 非常不符合 (請簡述原因)

17. 您認為「我是天空裡的一片雲」這句音樂中的「雲」符合其抽象意涵嗎？

非常符合 符合 中等 不符合 非常不符合 (請簡述原因)

18. 您認為「偶爾投影在你的波心」這句音樂中的「投影在你的波心」符合其抽象意涵嗎？

非常符合 符合 中等 不符合 非常不符合 (請簡述原因)

19. 您認為這段歌曲是否適切地表達出這段詩傳達的情感？

非常適切 適切 中等 不適切 非常不適切 (請簡述原因)



因)

20. 你是否聽過這段歌曲？

是 否

21. 你喜歡這段歌曲嗎？

非常喜歡 喜歡 中等 不喜歡 非常不喜歡 (請簡述原因)

版本四

22. 您認為「我是天空裡的一片雲」這句音樂中的「雲」符合其具體形象嗎？

非常符合 符合 中等 不符合 非常不符合 (請簡述原因)

23. 您認為「偶爾**投影在你的波心**」這句音樂中的「**投影在你的波心**」符合其具體形象嗎？

非常符合 符合 中等 不符合 非常不符合 (請簡述原因)

24. 您認為「我是天空裡的一片雲」這句音樂中的「雲」符合其抽象意涵嗎？

非常符合 符合 中等 不符合 非常不符合 (請簡述原因)

25. 您認為「偶爾**投影在你的波心**」這句音樂中的「**投影在你的波心**」符合其抽象意涵嗎？

非常符合 符合 中等 不符合 非常不符合 (請簡述原因)

26. 您認為這段歌曲是否適切地表達出這段詩傳達的情感？

非常適切 適切 中等 不適切 非常不適切 (請簡述原因)

27. 你是否聽過這段歌曲？

是 否

28. 你喜歡這段歌曲嗎？

非常喜歡 喜歡 中等 不喜歡 非常不喜歡 (請簡述原因)



以下將依序播放以《偶然》譜成的歌樂作品，一共有**四種版本的音樂**，

- 請特別注意詩中的標粗網底楷體字「**黑夜的海上，互放的光亮**」
- 請您在聽完各版本的音樂後，依您個人感受勾選符合您所感知的選項：

〔第二段詩詞〕

你我相逢在**黑夜的海上**，
你有你的，我有我的，方向；
你記得也好，
最好你忘掉，
在這交會時**互放的光亮**！

版本一

29. 您認為「你我相逢在**黑夜的海上**」這句音樂中「**黑夜的海上**」符合其具體形象嗎？

非常符合 符合 中等 不符合 非常不符合 (請簡述原因)

30. 您認為「在這交會時**互放的光亮**」這句音樂中「**互放的光亮**」符合其具體形象嗎？

非常符合 符合 中等 不符合 非常不符合 (請簡述原因)

31. 您認為「你我相逢在**黑夜的海上**」這句音樂中「**黑夜的海上**」符

合其抽象意涵嗎？

非常符合 符合 中等 不符合 非常不符合 (請簡述原因)

因)

32. 您認為「在這交會時互放的光亮」這句音樂中「互放的光亮」符合其抽象意涵嗎？

非常符合 符合 中等 不符合 非常不符合 (請簡述原因)

33. 您認為這段歌曲是否適切地表達出這段詩傳達的情感？

非常適切 適切 中等 不適切 非常不適切 (請簡述原因)

因)

34. 你是否聽過這段歌曲？

是 否

35. 你喜歡這段歌曲嗎？

非常喜歡 喜歡 中等 不喜歡 非常不喜歡 (請簡述原因)

因)

版本二

36. 您認為「你我相逢在黑夜的海上」這句音樂中「黑夜的海上」符合其具體形象嗎？

37. 非常符合 符合 中等 不符合 非常不符合 (請簡述原因)

您認為「在這交會時互放的光亮」這句音樂中「互放的光亮」符合其具體形象嗎？

非常符合 符合 中等 不符合 非常不符合 (請簡述原因)

38. 您認為「你我相逢在黑夜的海上」這句音樂中「黑夜的海上」符

合其抽象意涵嗎？

非常符合 符合 中等 不符合 非常不符合

(請簡述原因)

因)

39. 您認為「在這交會時互放的光亮」這句音樂中「互放的光亮」符合其抽象意涵嗎？

非常符合 符合 中等 不符合 非常不符合 (請簡述原因)

40. 您認為這段歌曲是否適切地表達出這段詩傳達的情感？

非常適切 適切 中等 不適切 非常不適切 (請簡述原因)

因)

41. 你是否聽過這段歌曲？

是 否

42. 你喜歡這段歌曲嗎？

非常喜歡 喜歡 中等 不喜歡 非常不喜歡 (請簡述原因)

因)

版本三

43. 您認為「你我相逢在黑夜的海上」這句音樂中「黑夜的海上」符合其具體形象嗎？

非常符合 符合 中等 不符合 非常不符合 (請簡述原因)

44. 您認為「在這交會時互放的光亮」這句音樂中「互放的光亮」符合其具體形象嗎？

非常符合 符合 中等 不符合 非常不符合 (請簡述原因)

因)

45. 您認為「你我相逢在黑夜的海上」這句音樂中「黑夜的海上」符



合其抽象意涵嗎？

非常符合 符合 中等 不符合 非常不符合

(請簡述原因)

46. 您認為「在這交會時互放的光亮」這句音樂中「互放的光亮」符合其抽象意涵嗎？

非常符合 符合 中等 不符合 非常不符合 (請簡述原因)

47. 您認為這段歌曲是否適切地表達出這段詩傳達的情感？

非常適切 適切 中等 不適切 非常不適切 (請簡述原因)

48. 你是否聽過這段歌曲？

是 否

49. 你喜歡這段歌曲嗎？

非常喜歡 喜歡 中等 不喜歡 非常不喜歡 (請簡述原因)

版本四

50. 您認為「你我相逢在黑夜的海上」這句音樂中「黑夜的海上」符合其具體形象嗎？

非常符合 符合 中等 不符合 非常不符合 (請簡述原因)

51. 您認為「在這交會時互放的光亮」這句音樂中「互放的光亮」符合其具體形象嗎？

非常符合 符合 中等 不符合 非常不符合 (請簡述原因)

52. 您認為「你我相逢在黑夜的海上」這句音樂中「黑夜的海上」符合其抽象意涵嗎？

非常符合 符合 中等 不符合 非常不符合 (請簡述原因)



原因)

53. 您認為「在這交會時互放的光亮」這句音樂中「互放的光亮」符合其抽象意涵嗎？

非常符合 符合 中等 不符合 非常不符合 (請簡述原因)

54. 您認為這段歌曲是否適切地表達出這段詩傳達的情感？(請簡述原因)

非常適切 適切 中等 不適切 非常不適切

55. 你是否聽過這段歌曲？

是 否

56. 你喜歡這段歌曲嗎？

非常喜歡 喜歡 中等 不喜歡 非常不喜歡 (請簡述原因)

受訪者基本資料

姓名：

性別：

年齡：

系所：

是否有藝術歌曲相關背景：

～謝謝您！～

Appendix 4: Statistical Data of Questionnaire Surveys



1. Frequencies of distribution occur in “recognizing” the involvement of the conceptual scenarios of “Troy”

	Observed N	Expected N	Residual
1.00	7	6.8	.3
2.00	12	6.8	5.3
3.00	1	6.8	-5.8
4.00	7	6.8	.3
Total	27		

Test Statistics

	ID
Chi-square	9.000 ^a
df	3
Asymp. Sig.	.029

2. Frequencies of distribution occur in “recognizing” the involvement of the conceptual scenarios of “Mountain”

	Observed N	Expected N	Residual
1.00	1	5.4	-4.4
2.00	19	5.4	13.6
3.00	1	5.4	-4.4
4.00	3	5.4	-2.4
5.00	3	5.4	-2.4
Total	27		

Test Statistics

	ID
Chi-square	43.556 ^a
df	4
Asymp. Sig.	.000

3. Frequencies of distribution occur in “recognizing” the involvement of the conceptual scenarios of “Circus”

ID

	Observed N	Expected N	Residual
1.00	4	5.4	-1.4
2.00	17	5.4	11.6
3.00	2	5.4	-3.4
4.00	2	5.4	-3.4
5.00	2	5.4	-3.4
Total	27		

Test Statistics

	ID
Chi-square	31.704 ^a
df	4
Asymp. Sig.	.000

4. Frequencies of distribution occur in “recognizing” the involvement of the conceptual scenarios of “Cowboy”

	Observed N	Expected N	Residual
1.00	1	5.4	-4.4
2.00	17	5.4	11.6
3.00	2	5.4	-3.4
4.00	5	5.4	-.4
5.00	2	5.4	-3.4
Total	27		

Test Statistics

	ID
Chi-square	32.815 ^a
df	4
Asymp. Sig.	.000

5. Frequencies of distribution occur in “recognizing” the involvement of the conceptual scenarios of “Island”

	Observed N	Expected N	Residual
1.00	2	5.0	-3.0
2.00	14	5.0	9.0
3.00	3	5.0	-2.0
4.00	2	5.0	-3.0
5.00	4	5.0	-1.0
Total	25		

Test Statistics

	ID
Chi-square	20.800 ^a
df	4
Asymp. Sig.	.000

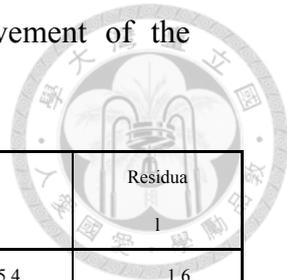
6. Frequencies of distribution occur in “recognizing” the involvement of the conceptual scenarios of “Bomb”

	Observed N	Expected N	Residual
1.00	6	6.8	-.8
2.00	18	6.8	11.3
3.00	1	6.8	-5.8
4.00	2	6.8	-4.8
Total	27		

Test Statistics

	ID
Chi-square	27.074 ^a
df	3
Asymp. Sig.	.000

7. Frequencies of distribution occur in “recognizing” the involvement of the conceptual scenarios of “Three Monkey”



	Observed N	Expected N	Residual
1.00	7	5.4	1.6
2.00	12	5.4	6.6
3.00	3	5.4	-2.4
4.00	4	5.4	-1.4
5.00	1	5.4	-4.4
Total	27		

Test Statistics

	ID
Chi-square	13.556 ^a
df	4
Asymp. Sig.	.009

8. Frequencies of distribution occur in “recognizing” the involvement of the conceptual scenarios of “Net”

	Observed N	Expected N	Residual
1.00	6	5.4	.6
2.00	11	5.4	5.6
3.00	2	5.4	-3.4
4.00	6	5.4	.6
5.00	2	5.4	-3.4
Total	27		

Test Statistics

	ID
Chi-square	10.222 ^a
df	4
Asymp. Sig.	.037

9. The matching level of the concrete image of CLOUD in four versions

i/j	1	2	3	4	5	6	7	8
1		-2.01222	-3.01832	-2.8171	-1.40855	-2.01222	-1.20733	-0.60366
		0.0460	0.0030	0.0055	0.1610	0.0460	0.2292	0.5470
2	2.012216		-1.00611	-0.80489	0.603665	0	0.804887	1.408551
	0.0460		0.3160	0.4221	0.5470	1.0000	0.4221	0.1610
3	3.018324	1.006108		0.201222	1.609773	1.006108	1.810995	2.41466
	0.0030	0.3160		0.8408	0.1095	0.3160	0.0721	0.0169
4	2.817103	0.804887	-0.20122		1.408551	0.804887	1.609773	2.213438
	0.0055	0.4221	0.8408		0.1610	0.4221	0.1095	0.0284
5	1.408551	-0.60366	-1.60977	-1.40855		-0.60366	0.201222	0.804887
	0.1610	0.5470	0.1095	0.1610		0.5470	0.8408	0.4221
6	2.012216	0	-1.00611	-0.80489	0.603665		0.804887	1.408551
	0.0460	1.0000	0.3160	0.4221	0.5470		0.4221	0.1610
7	1.20733	-0.80489	-1.81099	-1.60977	-0.20122	-0.80489		0.603665
	0.2292	0.4221	0.0721	0.1095	0.8408	0.4221		0.5470
8	0.603665	-1.40855	-2.41466	-2.21344	-0.80489	-1.40855	-0.60366	
	0.5470	0.1610	0.0169	0.0284	0.4221	0.1610	0.5470	

References



Agawu, K. (1992). Theory and practice in the analysis of the nineteenth-century 'Lied'. *Music Analysis*, 11 (1), 3-36.

Aksnes, H. (2002). *Perspectives of musical meaning: A study based on selected works by Geirr Tveitt*. Oslo: Unipub AS.

Atrey, P., Hossain, M., El Saddik, A., & Kankanhalli, M. (2010). Multimodal fusion for multimedia analysis: a survey. *Multimedia Systems*, 16, 345-379.

Bain, J., Chaban, N. & Kelly, S. (2012). Crumbling giant, rising dragon? Chinese news media cartoon reflections on the Eurozone debt crisis. *Communication, Politics & Culture*, 45, 217-240.

Baldry, A. and Thibault, P.J. (2006). *Multimodal transcription and text analysis: A multimedia toolkit and coursebook*. London: Equinox.

Barcelona, A. (2002). Clarifying and applying the notions of metaphor and metonymy within cognitive linguistics: an update. In R. Dirven & R. Pörings (Eds.), *Metaphor and metonymy in comparison and contrast* (pp. 207-278). Berlin: Mouton de Gruyter.

Bell, S. (2004). Drawing Tony. *The Guardian*, 17 July: 11

Benoit, W., & Klyukovski, A., McHale, J., & Airne, D. (2001). A fantasy theme analysis of political cartoons on the Clinton-Lewinsky-Starr affair.

Critical Studies in Media Communication, 18, 377–394.



Borgo, D. (2004). The play of meaning and the meaning of play in jazz.

Journal of Consciousness Studies, 11(3-4), 174-190.

Bostdorff, D.M. (1987). Making light of James Watt: A Burkean approach to the form and attitude of political cartoons. *Quarterly Journal of Speech*,

73, 43–59.

Bounegru, L., & Forceville, C. (2011). Metaphors in editorial cartoons representing the global financial crisis. *Visual Communication*, 10(2),

209-29.

Brower, C. (1997-1998). Pathway, blockage, and containment in density 21.5.

Theory and Practice, 22-23, 35-54.

Carroll, N. (1996). A note on film metaphor. In *Theorizing the Moving Image* (pp. 212-223). Cambridge: Cambridge University Press.

Chuck, G. (2004). *Musical meaning and cognitive operations of the embodied mind*. Unpublished PhD dissertation, University of Rochester, Eastman School of Music.

Cohn, N. (2007). A visual lexicon. *The Public Journal of Semiotics*, 1, 35–56.

Cohn, N. (2010). The limits of time and transitions: Challenges to theories of sequential image comprehension. *Studies in Comics, 1*, 127–147.

Connors, J.L. (1998) Hussein as enemy: The Persian Gulf War in political cartoons. *Press/Politics, 3*(3), 96–114.

Cook, N. (1998). *Analysing musical multimedia*. Oxford: Oxford University Press.

Coulson, S. (2001). *Semantic leaps: Frame-shifting and conceptual blending in meaning construction*. Cambridge: Cambridge University Press.

Coulson, S. (2002). *What's so funny? Conceptual blending in humorous examples*. Retrieved from <http://www.cogsci.ucsd.edu/~coulson/funstuff/funny.html>

Coulson, S., & Oakley, T. (2003). Metonymy and conceptual blending. In Panther, K. & L. L. Thornburg (Eds.), *Metonymy and pragmatic inferencing* (pp. 59-88). Amsterdam/Philadelphia: John Benjamins.

Coulson, S., & Oakley, T. (2005). Blending and coded meaning: Literal and figurative meaning in cognitive semantics. *Journal of Pragmatics, 37*(10), 1510-1536.

Coulson, S. (2006). Metaphor and conceptual blending. In K. Brown & A. H. Anderson (Eds.), *The encyclopedia of language and linguistics* (pp.



32–39). Oxford: Elsevier.

Cox, A. W. (2001). The mimetic hypothesis and embodied musical meaning.

Musicae Scientiae, 5 (2), 195-212.

Craig, D., Nersessian, N., & Catrambone, R. (2002). Perceptual simulation in

analogical problem solving. In L. Magnani & N. Nersessian (Eds.),

Model-based reasoning: Science, technology, values (pp. 167-190). New

York: Kluwer.

Croft, W. (2002). The role of domains in the interpretation of metaphors and

metonymies. In R. Dirven & R. Pörings (Eds.), *Metaphor and metonymy*

in comparison and contrast (pp. 161-205). Berlin: Mouton de Gruyter.

Davies, L. J. (1995). The multidimensional language of the cartoon: A study in

aesthetics, popular culture, and symbolic interaction. *Semiotica*, 104,

165–211.

Deignan, A. (2005). *Metaphor and corpus linguistics* (Vol. 6).

Amsterdam/Philadelphia: John Benjamins.

Diamond, M. (2002). No laughing matter: Post-September 11 political

cartoons in Arab/Muslim newspapers. *Political Communication*, 19,

251–272.

Dirven, R. (2002). Metonymy and metaphor: different mental strategies of

conceptualization. In R. Dirven & R. Pörings (Eds.), *Metaphor and metonymy in comparison and contrast* (pp. 75-111). Berlin: Mouton de Gruyter.



Edwards, J. L. (1997). *Political cartoons in the 1988 presidential campaign: Image, metaphor, and narrative*. New York/London: Garland.

Edwards, J. L. (2001). Running in the shadows in campaign 2000. *American Behavioral Scientist*, 44, 2140–2151.

Edwards, J. L., & Winkler, C. K. (1997). Representative form and the visual ideograph: The Iwo Jima image in editorial cartoons. *Quarterly Journal of Speech*, 83, 289–310.

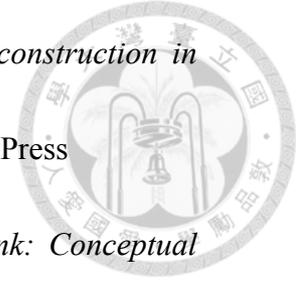
El Refaie, E. (2003). Understanding visual metaphor: The example of newspaper cartoons. *Visual Communication* 2(1), 75–96.

El Refaie, E. (2009a). Metaphor in political cartoons: Exploring audience responses. In C. Forceville & E. Urios-Aparisi (Eds.), *Multimodal metaphor* (pp.173–196). Berlin: Mouton de Gruyter.

El Refaie, E. (2009b). Multiliteracies: how readers interpret political cartoons. *Visual Communication*, 8(2), 181-205.

Fauconnier, G. (1985). *Mental spaces: Aspects of meaning construction in natural language*. Cambridge, MA: MIT Press.

Fauconnier, G. (1994). *Mental spaces: Aspects of meaning construction in natural language*. Cambridge, UK: Cambridge University Press



Fauconnier, G., & Turner, M. (2002/2008). *The way we think: Conceptual blending and the mind's hidden complexities*. New York: Basic Books.

Feld, S. (1981). Flow like a waterfall: The metaphors of Kaluli musical theory. *Yearbook for Traditional Music*, 13, 22-47.

Feld, S. (1982). *Sound and sentiment: Birds, weeping, poetics, and song in Kaluli expressions*. Philadelphia, PA: University of Pennsylvania Press.

Fernandez-Duque, D., & Johnson M. (1999). Attention metaphors: How metaphors guide the cognitive psychology of attention. *Cognitive Science* 23: 83-116.

Fillmore, C. J., (1985). Frames and the semantics of understanding. *Quaderni di Semantica* 6, 222-255.

Fillmore, C. J., (2006). Frame semantics. In D. Geeraerts (Ed.), *Cognitive linguistics: Basic Readings* (pp. 373-400). Berlin: Mouton de Gruyter.

Fillmore, C. J., & Atkins, B. T., (1992). Toward a frame-based lexicon: the semantics of RISK and its neighbors. In A. Lehrer & E. F. Kittay (Eds.), *Frames, fields, and contrasts: New essays in semantic and lexical organization* (pp. 75—102). NJ: Hillsdale.

Forceville, C. (1994). Pictorial metaphor in advertisements. *Metaphor and Symbolic Activity*, 9 (1), 1–29.



Forceville, C. (1996). *Pictorial metaphor in advertising*. London/New York: Routledge.

Forceville, C. (2000). Compasses, beauty queens and other PCs: Pictorial metaphors in computer advertisements. *Hermes, Journal of Linguistics*, 24, 31-55.

Forceville, C. (2006). Non-verbal and multimodal metaphor in a cognitivist framework: agendas for research. In G. Kristiansen, M. Achard, R. Dirven, & F. Ruiz de Mendoza Ibanes (Eds.), *Cognitive linguistics: Current applications and future perspectives* (pp. 379–402). Berlin/New York: Mouton de Gruyter.

Forceville, C. (2007). Multimodal metaphor in ten Dutch TV commercials. *Public Journal of Semiotics 1*, 19-51.

Forceville, C. (2008). Pictorial and multimodal metaphor in commercials. In E. F. McQuarrie & B. J. Phillips (Eds.), *Go figure! New directions in advertising rhetoric* (pp. 272-310). Armonk, NY: ME Sharpe.

Forceville, C. (2012 a). Creativity in pictorial and multimodal advertising metaphors. *Discourse and creativity*, 113-132.

Forceville, C. (2012 b). 14 The journey metaphor and the SourcePathGoal schema in Agnès Varda's autobiographical gleaning documentaries. *Beyond cognitive metaphor theory: Perspectives on literary metaphor*, 76.

Forceville, C. (2014). Relevance theory as model for analysing visual and multimodal communication. *Visual Communication*, 51-70. Berlin: Mouton de Gruyter.

Forceville, C., & Urios-Aparisi, E. (Eds.). (2009). *Multimodal metaphor*. Berlin, Germany: Mouton de Gruyter.

Gibbs, R. W. (1994). *The poetics of mind: Figurative thought, language, and understanding*. Cambridge: Cambridge University Press.

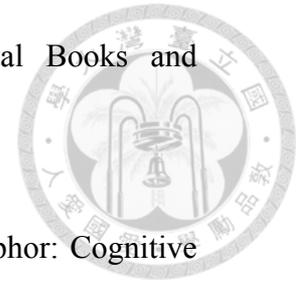
Gibbs, R. W., Jr. ed. (2008). *The Cambridge handbook of metaphor and thought*. Cambridge: Cambridge University Press.

Gibbs, R. W., Jr., & Colston, H. (1995). The psychological reality of Image Schemas and their transformations. *Cognitive Linguistics*, 6, 347-378.

Gibbs, R. W., Jr., & Colston, H. (2012). *Interpreting figurative meaning*. New York, NY: Cambridge University Press.

Gibbs, R. W., Jr., & Steen, G. J., eds. (1999). *Metaphor in cognitive linguistics*. Amsterdam: John Benjamins.

Gislason, S. (2011). *The sound of music*. Persona Digital Books and Environmed Research Inc.



Gleason, D. W. (2009). The visual experience of image metaphor: Cognitive insights into imagist figures. *Poetics Today*, 30, 423–470.

Grady, J., Oakley, T., & Coulson, S. (1999). Conceptual blending and metaphor. In R. Gibbs (Ed.), *Metaphor in cognitive linguistics* (pp. 421-440). Amsterdam/Philadelphia: John Benjamins.

Green, A. J. (2009). Mapping North America: Visual representations of Canada and the United States in recent academic work and editorial cartoons. *American Review of Canadian Studies*, 37, 130–149.

Groarke, L. (2002). Towards a pragma-dialectics of visual argument. In F. van Eemeren (Ed.), *Advances in pragmadialectics* (pp. 137–151). Amsterdam, The Netherlands: Sic Sat.

Grove, G. (2009). *A dictionary of music and musicians. A.D. 1450-1880*. Cambridge: Cambridge University Press.

Hsu, H.C., & Su, L. I. (2014). Love in disguise: Incongruity between text and music in song. *Journal of Pragmatics*, 62, 136-150.

Jakobson, R. (1956). Two aspects of language and two types of aphasic disturbance. In R. Jakobson & M. Halle (Eds.), *Fundamentals of*

language (pp. 53-83). S-Gravenhage: Mouton.



Johnson, M. (1987). *The body in the mind: The bodily basis of meaning, imagination, and reason*. Chicago: University of Chicago Press.

Johnson, M. (1993). *Moral imagination: implications of cognitive science for ethics*. Chicago: University of Chicago Press.

Johnson, M. (2005). The philosophical significance of Image Schemas. In B. Hampe and J. E. Grady (Eds.), *From perception to meaning. Image schemas in cognitive linguistics*, (pp. 15-33). Berlin/New York: Mouton de Gruyter.

Kaplan, S. T. (1990). Visual metaphors in the representation of communication technology. *Critical Studies in Mass Communication*, 7 (1), 7-47

Kelley-R., S., & Westgate, V. (2007). Blaming Bush. *Journalism Studies*, 8, 755-77

Kövecses, Z. (2005). *Metaphor in culture: Universality and variation*. Cambridge: Cambridge University Press.

Kövecses, Z. (2010). *Metaphor: A practical introduction* (2nd ed.). Oxford: Oxford University Press.

Kramer, L., (1984). *Music and poetry: The nineteenth century and after*.

Berkeley: University of California Press.

Kuo, C-T. (2002). *The study of newspaper images of the first ladies*. Master's Thesis. Tamkang University.



Lakoff, G. (1993). The contemporary theory of metaphor. In A. Ortony (Ed.), *Metaphor and thought* (pp. 202–251). Cambridge: Cambridge University Press.

Lakoff, G. (2004). *Don't think of an elephant! Know your values and frame the debate: The essential guide for progressives*. Vermont: Chelsea Green.

Lakoff, G., & Johnson, M. (1980/2003). *Metaphors we live by*. Chicago, IL: University of Chicago Press.

Lakoff, G., & Johnson M. (1999). *Philosophy in the Flesh: The embodied mind and its challenge to western thought*. New York: Basic Books.

Lakoff, G., & Nuñez, R. (2000). *Where mathematics comes from: How the embodied mind brings mathematics into being*. New York: Basic Books.

Lakoff, G., & Turner, M. (1989). *More than cool reason: A field guide to poetic metaphor*. Chicago, IL: University of Chicago Press.

Langer, S. K., (1953). *Feeling and form: A theory of art developed from philosophy in a new key*. Scribner, New York.

Larson, S. (1997-1998). Musical forces and melodic patterns. *Theory and Practice*, 22-23, 55-71.



Lin, M. H. (1987). *The international news analysis in the national newspapers* 1980-1986. Master's Thesis. Chinese Culture University.

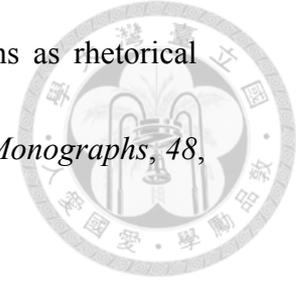
Lin, Y. Y. (2009). *The Chien-Ming Wang phenomenon: A critical metaphor model analysis of newspaper discourse in Taiwan*. Master's Thesis. National Taiwan University.

Lin, T.Y.Y., & Chen, I.H. (2014). How semantics is embodied through visual representation: Image Schemas in the art of Chinese calligraphy. *Proceedings of the 38th Annual Meeting of the Berkeley Linguistics Society*, 328-337.

Lin, T.Y.Y., & Chiang W.Y. (2015). Multimodal fusion in analyzing political cartoons: Debates on U.S. beef imports into Taiwan. *Metaphor and Symbol*, 30 (2), 137-161.

Lumsden, L. (2010). Striking images: Visual rhetoric and social identify in the radical press, 1903–1917. *Visual Communication Quarterly*, 17, 225–240.

McLuhan, M. (1964). *The medium is the message*. Hardwired, San Francisco, 8-9.



Medhurst, M. J., & Desousa, M. A. (1981). Political cartoons as rhetorical form: A taxonomy of graphic discourse. *Communication Monographs*, 48, 197–236.

Messaris, P. (1997) *Visual persuasion: The role of images in advertising*. Thousand Oaks, CA: Sage.

Mittelberg, I., & Waugh, L. R. (2009). Metonymy first, metaphor second: A cognitive-semiotic approach to multimodal figures of thought in co-speech gesture. *Multimodal metaphor*, 329-356.

Morrison, S. S. (1992). The feminization of the German Democratic Republic in political cartoons 1989-90. *The Journal of Popular Culture*, 25(4), 35–52.

Müller, C., & Cienki, A. (2009). Words, gestures, and beyond: Forms of multimodal metaphor in the use of spoken language. *Multimodal metaphor*, 297-328.

Oakley, T., & Hougaard, A. (Eds.). (2008). *Mental spaces in discourse and interaction* (Vol. 170). Amsterdam/Philadelphia: John Benjamins Publishing.

Parrill, F., & Sweetser, E. (2004). What we mean by meaning: Conceptual integration in gesture analysis and transcription. *Gesture*, 4(2), 197-219.

Peirsman, Y. & Geeraerts, D. (2006). Metonymy as a prototypical category.

Cognitive Linguistics 17, 269-316.



Philippe, R. (1982). *Political graphics: Art as weapon*. Oxford: Phaidon Press.

Radden, G. (2000). How metonymic are metaphors. In R. Dirven & R. Pörings

(Eds.), *Metaphor and metonymy at the crossroads: a cognitive perspective*, (pp. 93-108). Berlin: Mouton de Gruyter.

Saslaw, J. K. (1997-1998). Life forces: Conceptual structures in Schenker's

free composition and Schoenberg's musical idea. *Theory and Practice*, 22-23, 17-33.

Schilperoord, J., & Maes, A. (2009). Visual metaphorical conceptualization in

editorial cartoons. In C. Forceville, & E. Urios-Aparisi (Eds.), *Multimodal metaphor* (pp. 213–240). Berlin: Mouton de Gruyter.

Schilperoord, J. (2013). Raising the issue: A mental-space approach to Iwo

Jima-inspired editorial cartoons. *Metaphor & Symbol*, 28(3), 185-212.

Schoenberg, A. (1983). *Structural functions of harmony* (rev. ed.) London:

Faber and Faber.

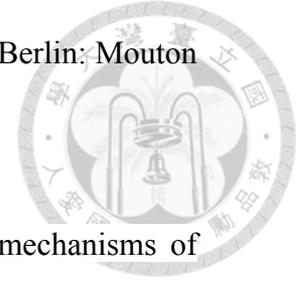
Semino, E., & Masci, M. (1996). Politics is football: metaphor in the discourse

of Silvio Berlusconi in Italy. *Discourse & Society*, 7(2), 243-269.

Seymour-Ure, C. (1998). Preface. In J. Newton (Eds.) *Political cartoons of*



- 1998, (p.3). Canterbury: Centre for the Study of Cartoons and Caricature.
- Spitzer, M. (2004). *Metaphor and musical thought*. Chicago, IL: University of Chicago Press.
- Sweetser, E. (1990). *From etymology to pragmatics. Metaphorical and cultural aspects of semantic structure*. Cambridge: Cambridge University Press.
- Teng, N. Y. (1999). The language of thought and the embodied nature of language use. *Philosophical Studies*, 94, 237–251.
- Teng, N. Y. & Sun, S. (2002). Grouping, simile, and oxymoron in pictures: A design-based cognitive approach. *Metaphor and Symbol*, 17(4): 295–316.
- Templin, C. (1999) Hillary Clinton as threat to gender norms: Cartoon images of the First Lady, *Journal of Communication Inquiry* 23(1), 20–36.
- Tsai, C. C. (2006). *Stylish chameleon: An exploratory examination on Taiwan Apple's Daily*. Master's Thesis. National Taipei University.
- Turner, M. (1991). *Reading minds: The Study of English in the Age of Cognitive Science*. Princeton, NJ: Princeton University Press.
- Urios-Aparisi, E. (2009). Interaction of multimodal metaphor and metonymy in TV commercials: Four case studies. In C. Forceville, & E.



- Urios-Aparisi (Eds.), *Multimodal metaphor* (pp. 95–117). Berlin: Mouton de Gruyter.
- Veale, T., Feyaerts, K., & Brône, G. (2006). The cognitive mechanisms of adversarial humor. *Humor: International Journal of Humor Research*, 19 (3), 305-339.
- Vultee, F. (2007). Dr. FDR and Baby War: The world through Chicago political cartoons before and after Pearl Harbor. *Visual Communication Quarterly*, 14, 158–175.
- Walker, M.E. 2000. Movement and metaphor: Towards an embodied theory of music cognition and hermeneutics. *Bulletin of the Council for Research in Music Education*, 145 (Summer), 27-42.
- Whittock, T. (1990). *Metaphor and film*. Cambridge: Cambridge University Press.
- Winter, S. (2001). *A clearing in the forest: Law, life, and mind*. Chicago: University of Chicago Press.
- Yu, N. (2011a). A decompositional approach to metaphorical compound analysis: The case of a TV commercial. *Metaphor and Symbol*, 26, 243-259.
- Yu, N. (2011b). Beijing Olympics and Beijing opera: A multimodal metaphor

in a CCTV Olympics commercial. *Cognitive Linguistics*, 22, 595-628.

Zbikowski, L. M. (1991). Large-scale rhythm and systems of grouping.

Unpublished PhD dissertation, Yale University, New Haven, CT.

Zbikowski, L. M. (1998). Metaphor and music theory: Reflections from cognitive science. *Music Theory Online*, 4 (1), January.

(<http://smt.vesb.edu/mto/issues/mto.98.4.1.html>)

Zbikowski, L. M. (1999). The blossoms of 'Trockne Blumen': Music and text in the early nineteenth century. *Music Analysis*, 18, 307-345.

Zbikowski, L. M. (2002). *Conceptualizing music: cognitive structure, theory, and analysis*. Oxford: Oxford University Press.

Zbikowski, L. M. (2008). Metaphor and music. In R. W. Gibbs, Jr. (Eds.), *The Cambridge handbook of metaphor and thought* (pp. 502-524). Cambridge: Cambridge University Press.

Zbikowski, L. M. (2009). Music, language, and multimodal metaphor. In C. J. Forceville and E. Urios-Aparisi (Eds.), *Multimodal metaphor* (pp. 359-381). Berlin: Mouton de Gruyter.