

**A MULTIMODAL DISCOURSE ANALYSIS
OF FEMALE K-POP MUSIC VIDEOS**

**A MULTIMODAL DISCOURSE ANALYSIS
OF FEMALE K-POP MUSIC VIDEOS**

by

Chris K Brady

A dissertation submitted to the
College of Arts and Law
of the University of Birmingham
in part fulfillment of the requirements
for the degree of

Master of Arts

In

Applied Linguistics

This dissertation consists of approximately **14,716** words

Supervisor: Dr. Alexanne Don

ELAL
College of Arts & Law
University of Birmingham
Edgbaston, Birmingham B15 2TT
United Kingdom

September 2015

ABSTRACT

This dissertation employs multimodal discourse analysis to investigate the semiotic choices made in two sets of K-pop music videos, using Kress and van Leeuwen's (2006) framework for the grammar of visual design as well as van Leeuwen's framework for pitch (1999) to investigate whether the two sets of videos serve to satisfy a male viewer or not.

The analysis showed that the two sets of videos differed in that the *SNSD* videos *did* serve to satisfy a male viewer, whilst the *2NE1* videos *did not*. Whilst the two *SNSD* videos empowered a male viewer through representing the girls as infantilized and through positioning the male viewer as if he were the male on screen, the two *2NE1* videos did not satisfy or empower the male viewer but instead avoided or challenged the viewer through the use of oblique angles, particular types of clothing, and the actions of the girls themselves.

ACKNOWLEDGEMENTS

I would like to thank my tutor for her help and advice during this dissertation, and I would like to thank my parents for their love and support.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION.....	12
1.1 Rationale for the study.....	12
1.2 Narrowing the scope of enquiry.....	12
1.3 Formulating the research question.....	13

CHAPTER 2

ANALYTICAL FRAMEWORKS AND METHODOLOGY.....	14
2. Visual and musical analytical frameworks and their application.....	14
2.1 Framework for visual analysis.....	14
2.1.1 <i>Experiential</i> metafunction.....	15
2.1.2 <i>Interpersonal</i> metafunction.....	20
2.2 Framework for musical analysis.....	25
2.3 The application of the frameworks.....	28
2.3.1 The Texts.....	28
2.3.2 Visual analysis: The use of the <i>Multimodal Analysis Video</i> software.....	28
2.3.3 Musical analysis.....	31

CHAPTER 3

A MULTIMODAL DISCOURSE ANALYSIS OF THE MUSIC VIDEOS FOR <i>GEE</i> (2009) AND <i>I GOT A BOY</i> (2013) BY <i>SNSD</i> , AND <i>LONELY</i> (2011) AND <i>I AM THE BEST</i> (2011) BY <i>2NE1</i>	33
3. Presentation of findings.....	33
3.1 <i>SNSD</i>	33
3.1.1 Analysis of <i>Gee</i> (2009).....	33
3.1.1.1 Interpersonal analysis - State transition machine findings.....	33
3.1.1.2 Shots involving <i>offer, oblique, eye</i>	36
3.1.1.3 Shots involving <i>demand, close, frontal, eye</i>	38
3.1.1.4 Musical analysis.....	40
3.1.2 Analysis of <i>I Got A Boy</i> (2013).....	45
3.1.2.1 Experiential analysis - Frequency count findings.....	45

3.1.2.2 <i>Matched Disconnected Transactional Actions</i>	47
3.1.2.3 <i>Transactional Reaction Processes</i> involving the male as <i>phenomenon</i>	50
3.1.2.4 Musical analysis.....	54
3.2 <i>2NE1</i>	56
3.2.1 Analysis of <i>Lonely</i> (2011).....	56
3.2.1.1 Experiential analysis – Frequency count findings.....	56
3.2.1.2 Interpersonal analysis – State transition machine findings.....	58
3.2.1.3 Avoidance of the viewer.....	60
3.2.1.4 Musical Analysis.....	62
3.2.2 Analysis of <i>I Am The Best</i> (2011).....	64
3.2.2.1 Experiential analysis – Frequency count findings.....	64
3.2.2.2 <i>Circumstance: Means</i>	66
3.2.2.3 <i>Reaction Processes</i>	68
3.2.2.4 Interpersonal analysis – State transition machine findings.....	69
3.2.2.5 Shots involving <i>close, frontal, eye</i>	70
3.2.2.6 Musical Analysis.....	71
3.3 Summary of findings.....	73
 CHAPTER 4	
CONCLUSION.....	75
 REFERENCES.....	77

LIST OF FIGURES

Figure 2.1 Example of an <i>Action Process</i> taken from the video for <i>I Am The Best</i> (2011) at 2:34-2:34.....	16
Figure 2.2 Example of a <i>Reaction Process</i> taken from the video for <i>I Got A Boy</i> (2013) at 0:10-0:10.....	16
Figure 2.3 Example of a <i>Circumstance of Accompaniment</i> taken from the video for <i>Gee</i> (2009) at 1:43-1:43.....	18
Figure 2.4 Example of a <i>Circumstance of Setting</i> taken from the video for <i>I Got A Boy</i> (2013) at 1:21-1:21.....	18
Figure 2.5 Example of a <i>Filmic Event</i> taken from the video for <i>I Am The Best</i> (2011) at 2:17-2:17.....	20
Figure 2.6 Example of a <i>demand</i> act taken from the video for <i>I Got A Boy</i> (2013) at 0:27-0:28.....	21
Figure 2.7 Example of an <i>offer</i> act taken from the video for <i>Lonely</i> (2011) at 0:05-0:10.....	21
Figure 2.8 List of subject-initiated choices for the systems of DYNAMIC CONTACT (GAZE), DYNAMIC SOCIAL DISTANCE and DYNAMIC ATTITUDE.....	24
Figure 2.9 List of camera-initiated choices for the systems of DYNAMIC SOCIAL DISTANCE and DYNAMIC ATTITUDE.....	25
Figure 2.10 A system network for describing melodies (van Leeuwen 1999: 119).....	26
Figure 2.11 Screenshot of the <i>Multimodal Analysis Video</i> software systems library showing multimodal systems (left) and system choices (right).....	29

Figure 2.12 Screenshot of GUI on <i>Multimodal Analysis Video</i> software showing player window [1], transcription window with time stamped dialogue [2], selected system choice [3], available system choices [4], system strips based on the personalised list of systems chosen in the library [5], and time stamped nodes showing selected system choices [6].....	30
Figure 3.1 State transition diagram for <i>Gee</i> using the systems of STATIC CONTACT AND STATIC ATTITUDE.....	34
Figure 3.2 State Transition Diagram for <i>Gee</i> using the systems of STATIC CONTACT, STATIC ATTITUDE and STATIC SOCIAL DISTANCE.....	35
Figure 3.3 Melody of <i>Gee</i> [0:40-0:50] – (Repeated at 0:50-1:00 and 1:40-2:00).....	41
Figure 3.4 Melody of <i>Gee</i> [1:00-1:19] – (Repeated at 2:00-2:19).....	43
Figure 3.5 Shot from the video for <i>Gee</i> (1:04-1:05) that works as a visual complement to the use of the 2 nd note of the scale at notes 20 and 26-28 in figure 3.4.....	45
Figure 3.6 Three shots taken from the video for <i>I Got A Boy</i> at 0:21, 0:22 and 0:25.....	48
Figure 3.7 Three consecutive shots taken from the video for <i>I Got A Boy</i> at 4:21-4:23, 4:23-4:23 and 4:24-4:24.....	49
Figure 3.8 Two consecutive shots taken from the video for <i>I Got A Boy</i> at 2:24-2:24 and 2:25-2:25.....	52
Figure 3.9 Two consecutive shots taken from the video for <i>I Got A Boy</i> at 4:14-4:14 and 4:15-4:15.....	53
Figure 3.10 Two consecutive shots taken from the video for <i>I Got A Boy</i> at 4:16-4:17 and 4:17-4:17.....	53
Figure 3.11 Melody of <i>I Got A Boy</i> [3:07-3:20].....	54

Figure 3.12 Shot taken from the video for <i>I Got A Boy</i> at 3:08-3:08 in which a visual complement to the descent in pitch exhibited by notes 3 to 7 in figure 3.16 occurs.....	55
Figure 3.13 State transition diagram for <i>Lonely</i> using the systems of STATIC CONTACT and STATIC ATTITUDE.....	59
Figure 3.14 Three shots taken from the video for <i>Lonely</i> at 0:13-0:13, 0:37-0:37 and 2:29-2:29.....	61
Figure 3.15 One shot taken from the video for <i>Lonely</i> at 2:54-2:54.....	61
Figure 3.16 Melody of <i>Lonely</i> [0:20-0:38] – (Repeated at 1:43-2:01).....	63
Figure 3.17 Shot taken from the video for <i>I Am The Best</i> at 2:28-2:28.....	67
Figure 3.18 State transition diagram for <i>I Am The Best</i> using the systems of STATIC CONTACT, STATIC ATTITUDE and STATIC SOCIAL DISTANCE.....	69
Figure 3.19 Melody of <i>I Am The Best</i> [0:09-0:23] – (Repeated at 1:24-1:38 and 2:24-2:38).....	72

LIST OF TABLES

Table 3.1 Results from the state transition diagram for <i>Gee</i> for the systems of STATIC CONTACT AND STATIC ATTITUDE.....	34
Table 3.2 Results from the state transition diagram for <i>Gee</i> showing choices from the system of STATIC CONTACT combined with the choices of <i>close</i> , <i>frontal</i> and <i>eye</i>	36
Table 3.3 Table showing instances of CTAs co-occurring with the interpersonal system choices of offer, oblique and eye in the video for <i>Gee</i>	37
Table 3.4 Selection of <i>demand</i> , <i>close</i> , <i>frontal</i> , <i>eye</i> shots featuring gestures in the video for <i>Gee</i>	39
Table 3.5 Selection of <i>demand</i> , <i>close</i> , <i>frontal</i> , <i>eye</i> shots featuring hats being worn in the video for <i>Gee</i>	39
Table 3.6 Two shots from the video for <i>Gee</i> that work as visual complements to the decreased pitch range exhibited by notes 1-20 in figure 3.3.....	42
Table 3.7 Two shots from the video for <i>Gee</i> that work as visual complements to the beginning of two separate descents in pitch exhibited by notes 1 and 33 in figure 3.4.....	44
Table 3.8 Frequency of <i>participants</i> in the video for <i>I Got A Boy</i>	46
Table 3.9 Frequency of <i>action processes</i> in the video for <i>I Got A Boy</i>	46
Table 3.10 Frequency of <i>reaction processes</i> in the video for <i>I Got A Boy</i>	47
Table 3.11 Table showing MDTAs in the video for <i>I Got A Boy</i>	48
Table 3.12 <i>Transactional Reaction Processes</i> featuring male as <i>phenomenon</i>	51

Table 3.13 Frequency of Participants in the video for <i>Lonely</i>	56
Table 3.14 Frequency of <i>Action Processes</i> in the video for <i>Lonely</i>	57
Table 3.15 Frequency of <i>Reaction Processes</i> in the video for <i>Lonely</i>	57
Table 3.16 Results from the state transition diagram for <i>Lonely</i> for the systems of STATIC CONTACT AND STATIC ATTITUDE.....	59
Table 3.17 Frequency of <i>participants</i> in the video for I Am The Best (2011).....	65
Table 3.18 Frequency of <i>Action Processes</i> in the video for I Am The Best (2011).....	65
Table 3.19 Frequency of <i>Reaction Processes</i> in the video for I Am The Best (2011).....	65
Table 3.20 Objects used as <i>Circumstance of Means</i> and their frequency.....	66
Table 3.21 Results from the state transition diagram for <i>I Am The Best</i> for the systems of STATIC CONTACT, STATIC ATTITUDE and STATIC SOCIAL DISTANCE.....	70

CHAPTER 1

INTRODUCTION

1.1 Rationale for the study

Korea has traditionally been, and still is, a highly gender divided society. Moon argues that the transformation of Korea into a modern industrialized state reinforced traditional Confucian patriarchal gender roles, forming a capitalist industrialized patriarchy (Moon 2002; cited in Jung 2011). What this has meant is that much of Korea's modernization has been instigated and controlled by men, while Korean women have been relegated to the status of "passive tools or men's subordinates" (Jung 2011: 26).

K-pop, of course, is no different in this sense. The founders of the three biggest talent agencies are all male, and it is all too often males who have the final say in the production of K-pop. As Jang and Kim point out, "idol groups are still subject to Confucian (patriarchal) managerial principles within their own production teams" (2013: 99; my brackets).

The idea for this dissertation was borne out of my fascination with the K-pop industry and the way it presents its female stars. Being aware of the fact that the representation of femininity and female sexuality is part of a marketing strategy initiated by K-Pop entertainment agencies, I felt compelled to investigate how this was achieved (Jung and Hirata 2012). I also wanted to investigate if there was any sign of change in the way the women in K-pop were being presented, or whether it was all just shamelessly made by men, for men. Having lived in South Korea for the past eight years of my life, I was also interested to learn more about this, at times peculiar, aspect of Korean culture.

1.2 Narrowing the scope of enquiry

One of the first tasks was to decide which K-pop artists to look at, and the best place to start was with the K-pop entertainment agencies. I knew that *YG* Entertainment has a reputation for being the most music-oriented label, and its relationship with international

producers and musicians has been described as friendly (Choi 2011). I also knew that, by contrast, the other of the big two entertainment agencies, *SM Entertainment*, has gained the negative reputation of being an “idol-manufacturing company” (Choi 2011: 71).

Although both of these agencies adopt vertical integration in their organization, they differentiate themselves from each other through cultivating distinct brand images. As Shin and Kim state, “each house internally developed a distinct and internally consistent packaging strategy for their entertainers” (2013: 266) Thus, whilst *SM* artists are trained to be “proper, perfect and elite-like”, artists trained at *YG* are trained to express “individuality” as part of a brand image that places rebelliousness at the forefront (Kim 2012; cited in Shin and Kim 2013: 266). Each of these labels has cultivated its brand, then, through manipulating the *images* of its entertainers, and this was precisely what I wished to investigate.

1.3 Formulating the research question

After choosing the most successful girl group from each label, I formulated the following hypothesis:

The semiotic choices made in *SM Entertainment*’s SNSD music videos¹ serve to satisfy a heterosexual male viewer, whereas the semiotic choices made in *YG Entertainment*’s 2NE1 videos do not serve to satisfy a heterosexual male viewer².

I decided that the best way to investigate this was through conducting a multimodal analysis of the videos. Additionally, given that they are music videos, investigating the music was necessary too. The use of multimodal analysis is, in fact, perfectly suited for such an investigation as it allows for the complementarity of audio and visual elements present in K-pop videos to be taken into account. The following chapter details the frameworks that were chosen and how they were applied in order to test the hypothesis.

¹ Henceforth *videos*

² From here on *male* will stand in for *heterosexual male*. Whilst a study looking at the position of homosexual male viewers in relation to K-pop is no doubt interesting, it is beyond the confines of this study.

CHAPTER 2

ANALYTICAL FRAMEWORKS AND METHODOLOGY

2. Visual and musical analytical frameworks and their application

The purpose of this dissertation was to investigate whether the semiotic choices made in two of the music videos by the K-pop group *SNSD* and two of the music videos by the K-pop group *2NE1* served to satisfy a male viewer or not. The framework that was used to investigate the visual aspects of each of the videos was based on Kress and van Leeuwen's grammar of visual design (2006), whilst the framework that was used to analyse the music for each of the videos was based on the approach outlined by van Leeuwen in his book, *Speech, Music, Sound* (1999). Specific elements of these two frameworks were utilised, and in some cases modified, in order to test the hypothesis set out in this dissertation. The elements that were selected from these two frameworks will be explained in detail in the first two sections of this chapter before an explanation of how the study was conducted is presented in the final section.

2.1 Framework for visual analysis

Kress and van Leeuwen's (2006) social semiotic approach to the analysis of images is grounded in Halliday's approach to language as a social semiotic (2004). In this approach, language is seen as being "one of the semiotic systems that constitute a culture" understood "by reference to its place in the social process", and "modelled as a resource for making meaning that has evolved, and is organized in response to the three functions ('metafunctions') it serves in society" (Halliday 1978; cited in Djonov and Zhao 2014: 3).

According to Halliday (2004), the three *metafunctions* language serves in society are based upon:

- representation and the way experience is portrayed in the world – the **experiential** metafunction;
- social relations in the world and how they are enacted – the **interpersonal** metafunction;
- the organisation of meaning within text – the **textual** metafunction.

These three *metafunctions* also form the basis for Kress and van Leeuwen's (henceforth KvL) grammar of visual design (2006). However, in KvL's approach they are adapted in order to reflect the fact that images "can 'say' (some of) the same things as language – in very different ways" (KvL 2006: 50).

In KvL's approach, the *Experiential* metafunction concerns "the representation of interactions and conceptual relations between the people, places and things depicted in images" (KvL 2006: 114). The *Interpersonal* metafunction in their approach, meanwhile, concerns the "complex set of relations that can exist between images and their viewers" (KvL 2006: 175).

Given that the hypothesis tested in this dissertation concerns the way the girls are presented in the videos in relation to whether or not such a presentation could be said to serve a male viewer, investigating the girls interactions within the world of the video and their interactions with the viewer were deemed to be the most worthwhile avenues of enquiry. To this end, only the *Experiential* and *Interpersonal* metafunctions were employed in order to test the hypothesis.

2.1.1 *Experiential* metafunction

KvL's conception of the *Experiential* metafunction in images, as it relates to narrative processes, concerns the representation of objects and the representation of relationships between those objects (2006). In their approach, the relationships between objects are either realised by vectors formed by movement (*Action Processes*) or vectors formed by the direction in which the eyes of an object are looking (*Reaction Processes*). An

example of an *Action process* can be seen in figure 2.1, and an example of a *Reaction process* can be seen in figure 2.2.



Figure 2.1 Example of an *Action Process* taken from the video for *I Am The Best* (2011) at 2:34-2:34



Figure 2.2 Example of a *Reaction Process* taken from the video for *I Got A Boy* (2013) at 0:10-0:10

The nature of the vector determines the classification of the objects (or *participants*) involved in the transaction. *Participants* involved in *Action Processes* are referred to as *actors* and the recipient of their action is referred to as a *goal*. *Participants* involved in *Reaction Processes* are referred to as *reacters* and the recipient of their look is referred to as a *phenomenon*. In the *Action Process* in figure 2.1, we can see that the *actor* is the girl

and the *goal* is the glass cabinet. In the *Reaction Process* in figure 2.2, we can see that the *reacter* is the girl on the right and the *phenomenon* is the girl on the left.

In addition to classifying the *participants* and the type of process, the nature of the circumstances in which the action is completed was also identified. In KvL's model, there are three types of *Circumstance* (2006). The ***Circumstance of Means*** concerns the object with which the action was carried out. It forms part of the vector and, as such, can be realised by parts of the body such as a finger or hand. We can see in figure 2.1 that the *Circumstance of Means* is the baseball bat. It completes the vector formed between the girl and the glass cabinet.

The second type of *Circumstance* is the ***Circumstance of Accompaniment***. This concerns who is present during the action, but not involved in any way with it, which is to say that they have “no vectorial relation with other participants” (KvL 2006: 75). We can see an example of this in figure 2.3, where the girl on the far left clearly has no vectorial relationship with the other participants.

The final type of *Circumstance* is the ***Circumstance of Setting***. This concerns contrasts between the foreground and the background of a narrative structure in terms of darkness and lightness, colour saturation, detail (in painting) or focus (in photography). Figure 2.4 shows an example of a *Circumstance of Setting* that is defined by the lightness of the background.



Figure 2.3 Example of a *Circumstance of Accompaniment* taken from the video for *Gee* (2009) at 1:43-1:43



Figure 2.4 Example of a *Circumstance of Setting* taken from the video for *I Got A Boy* (2013) at 1:21-1:21

In KvL's model the *Action* and *Reaction* processes are also further subdivided in relation to whether they are directed at someone or something (***Connected Transactional Actions/Reactions***) or whether they are not directed at anyone or anything (***Non-Transactional Actions/Reactions***). Transactions involving a vector and a *goal* with no *actor*, meanwhile, are referred to as ***Events***.

However, although most of the classifications included in KvL's approach to the analysis of images could be readily applied to the analysis of film, in certain instances modifications were necessary, for as van Leeuwen (1996) points out, moving images differ from still images in terms of permitting motion, allowing for the editing of sequences of images, and for permitting the combination of the moving images with other modes such as music.

In particular, the ability to edit sequences of images together means transactions can take place over multiple shots, with the participants involved being shown in separate shots. They are therefore disconnected, and the maker of the film can either present these two disconnected shots as matched or as unmatched depending on how much he or she wishes us to identify with either one of the participants involved. This applies to both *Action* and *Reaction Processes*. With this in mind, the system choices of ***Matched*** and ***Unmatched Disconnected Transactional Actions/Reactions*** were devised for the analysis of the videos in this dissertation.

One final addition was the inclusion of what I have termed a ***Filmic Event***. Van Leeuwen (1996) refers to them as being simply ***Events***, but in order to distinguish them from processes where there are only *vectors* and *goals*, the term ***Filmic Event*** was devised. ***Filmic Events*** are processes of pure movement such as rippling water that involve no *participants* at all. The video for *I Am The Best* by 2NE1 features 20 of these ***Filmic Events***, and an example of one can be seen in figure 2.5, where the black diamond moves round by itself.

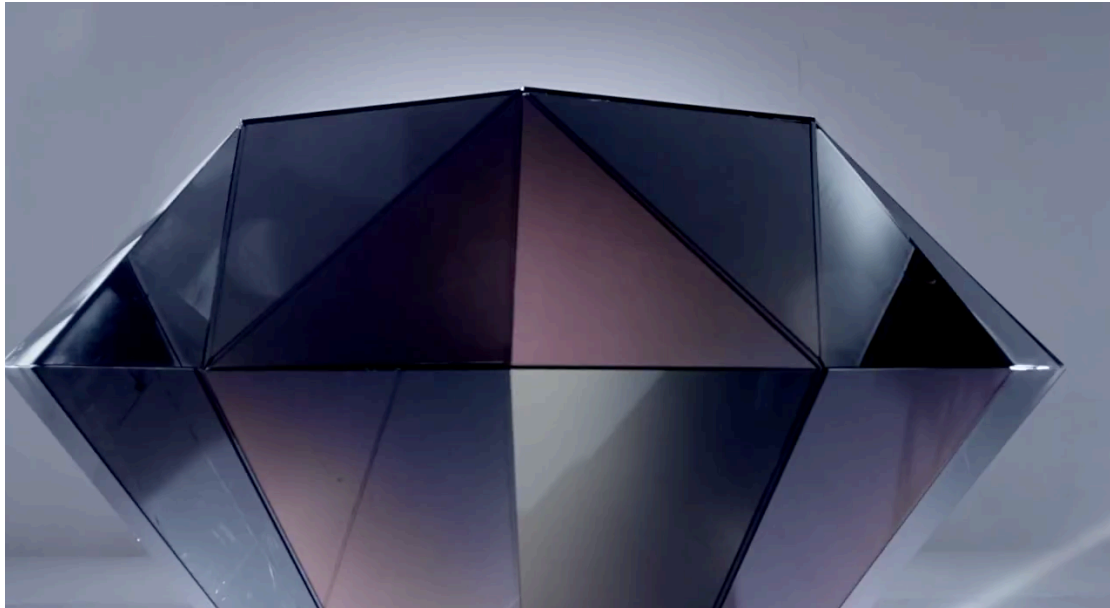


Figure 2.5 Example of a *Filmic Event* taken from the video for *I Am The Best* (2011) at 2:17-2:17

2.1.2 Interpersonal metafunction

KvL's conception of the Interpersonal metafunction in images concerns the interaction between the represented participants and the viewer of the image (2006). According to their approach, there are three main ways in which interactive meanings are made.

The first way is through the system of **CONTACT**. This can either involve the represented participant looking at the viewer (*demand*), or not (*offer*). *Demand* acts involve something being done to the viewer via the image. A *demand* act can also be combined with other gestures such as smiling, pouting, or pointing at the viewer. We can see an example of this in figure 2.6.



Figure 2.6 Example of a *demand* act taken from the video for *I Got A Boy* (2013) at 0:27-0:28

Offer acts, however, reverse the process so that the viewer becomes the subject of the look rather than the object. This creates a sense of detachment, reducing the role of the viewer to that of an onlooker, whilst the participants themselves are offered to the viewer as specimens to be observed or otherwise as “objects of contemplation” (KvL 2006: 119). We can see an example of this happening in figure 2.7.



Figure 2.7 Example of an *offer* act taken from the video for *Lonely* (2011) at 0:05-0:10

In my analysis, combinations of *offer* and *demand* acts were also permitted for situations when more than one participant was in the shot.

The second way interactive meanings are made is through the system of **SOCIAL DISTANCE**. This relates to the distance between the represented participant and the viewer. In this dissertation, the three basic distances of *close*, *medium* and *long* were used to classify the shots, with close shots involving the participant's head taking up approximately half of the screen, medium shots showing the body of the participant approximately down to the knees and long shots showing anything more than that.

KvL (2006) point out that the distance of shots relates to Hall's concept of proxemics, where the distance we keep from each other determines and is determined by the social relations that exist between us (1966; cited in KvL 2006). Extreme close shots thus create a sense of intimacy between the viewer and the represented participant, ensuring the participant is portrayed as if they were our lover or close friend, while long shots are impersonal, portraying participants as if they were strangers we might pass on the street.

The final way interactive meanings are made is through the system of **ATTITUDE**, which can be subjectively or objectively expressed. In this dissertation, I focused exclusively on subjective ATTITUDE. The system of subjective ATTITUDE involves choices associated with the angle (horizontal and vertical) at which the participants are portrayed. Frontal angles communicate involvement with the participants, while oblique angles communicate detachment from them. High angles ensure the viewer has power over the participants, whilst the inverse is true for low angles. Eye level angles, meanwhile, stand for equality (KvL 2006). A further angle, *Backview*, was also used for participants showing their backs to the viewer. Combinations of angles were also permitted for when there was more than one represented participant.

KvL point out that the choice of angle betrays the degree of alignment the image producer wishes the viewer to have with the represented participants, hence adding an ideological dimension to the predefined image relationship the viewer has with the represented participant (2006). In figure 2.7 (above) we can see an example of an oblique angle being used, signalling that we are not supposed to identify with the girl on the screen. From the

perspective of a study looking at whether or not videos have been designed with a male viewer in mind, considering angles was of particular relevance.

However, looking at the image of the girl in figure 2.7 only tells us so much. What is not taken into account is that the fact that she is turning towards the viewer. Hence, the framework for interactive meanings outlined above had to be adapted to account for the movement that occurred in the videos analysed in this dissertation. Van Leeuwen (1996) states that the dynamicization of the Interpersonal is achieved through either movement of the subject or movement of the camera. With this in mind, I first of all divided the three systems above according to whether they were STATIC or DYNAMIC. Crucially, when STATIC systems were chosen it did not imply that there was a lack of movement on screen, but rather that the movement on screen was not significant enough to result in a transition from one of the pre-defined categories to another.

Once the newly created systems of DYNAMIC SOCIAL DISTANCE and DYNAMIC ATTITUDE had been created, they were further subdivided according to whether the choices they comprised were camera-initiated or subject-initiated. Choices from the system of DYNAMIC CONTACT, meanwhile, can only be realised by the subject and thus it was coded as being subject-initiated from the start.

Finally, choices were then devised to account, as best as possible, for the movements realised within these DYNAMIC systems. The selection of choices created to account for subject-initiated movement can be seen in figure 2.8, whilst the selection of choices created to account for camera-initiated movement can be seen in figure 2.9.

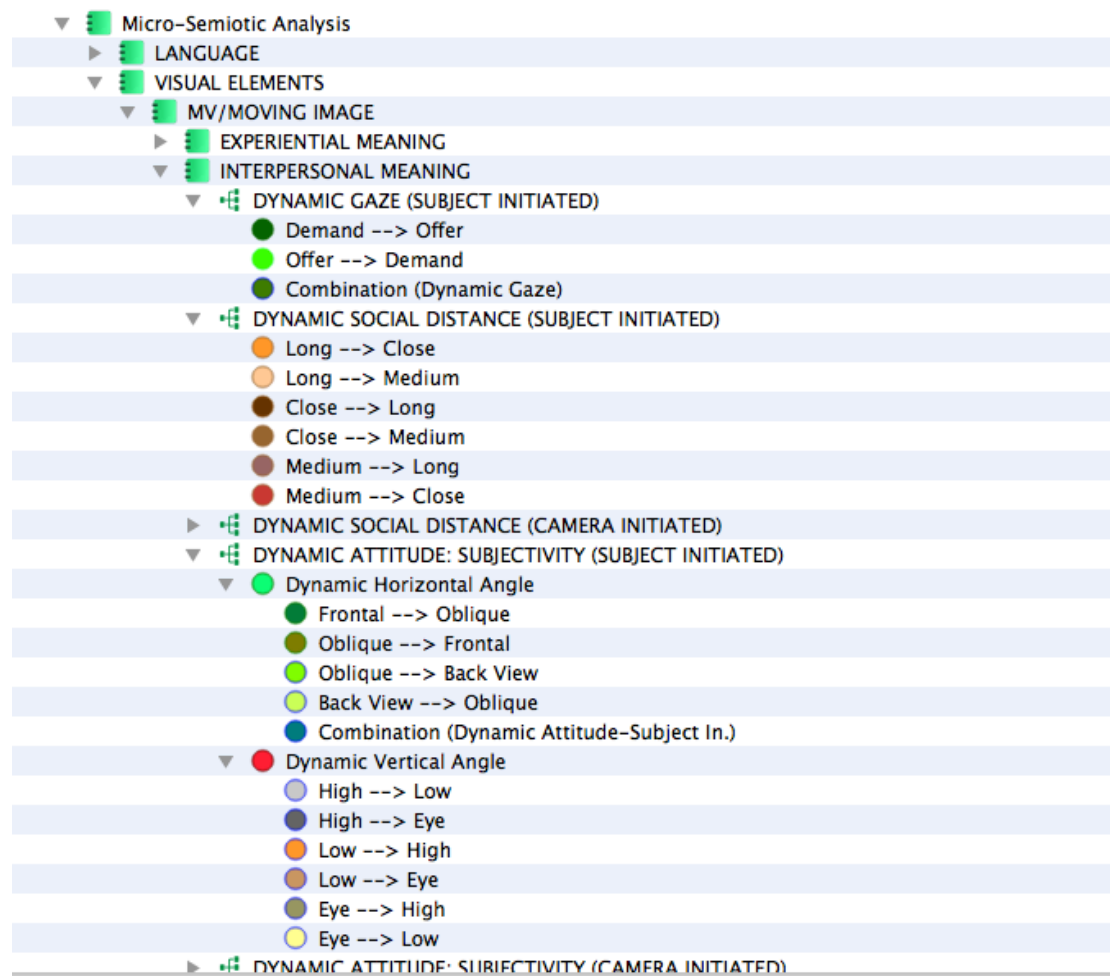


Figure 2.8 List of subject-initiated choices for the systems of DYNAMIC CONTACT (GAZE), DYNAMIC SOCIAL DISTANCE and DYNAMIC ATTITUDE

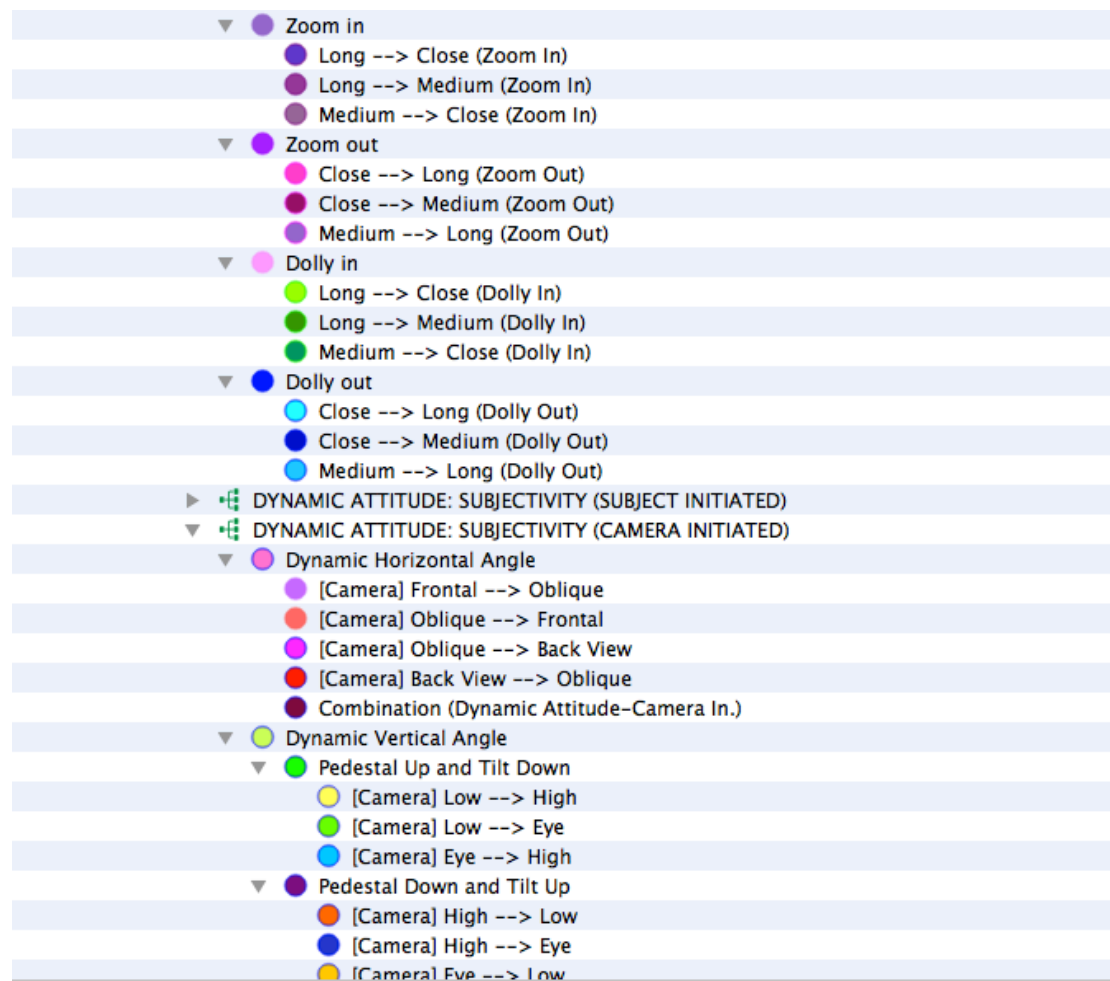


Figure 2.9 List of camera-initiated choices for the systems of DYNAMIC SOCIAL DISTANCE and DYNAMIC ATTITUDE

In this dissertation, it was important not just to consider the visual aspects of the video, but to consider the musical aspects too. The next section details the framework that was used for the analysis of the music in this dissertation.

2.2 Framework for musical analysis

The framework upon which the analysis of the music in this dissertation was based was van Leeuwen's *Speech, Music, Sound* (1999). However, Machin's book, *Analysing Popular Music* (2010) was also consulted during the analysis of the music. The main area of focus in this dissertation was pitch and melody. Van Leeuwen's system network for describing melodies sets out clearly the elements that were considered in my analysis (1999). This system network can be seen in figure 2.10.

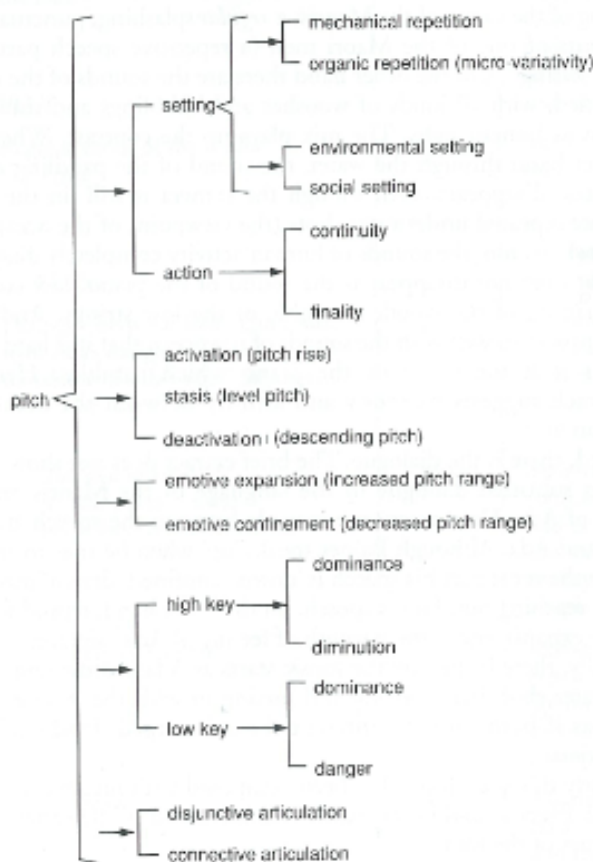


Figure 2.10 A system network for describing melodies (van Leeuwen 1999: 119)

The main elements that were analysed were the pitch movement, the pitch range, the pitch level and the articulation of the melody, as these four elements are “frequently used for the depiction of the actions and qualities of people, places and things”, which was deemed to be highly relevant to a study that considers the way females are represented in music videos (van Leeuwen 1999: 112). However, elements of *setting* and *action* were considered in certain cases too. I will now explain these four elements in turn.

The pitch movement concerns rises and falls of pitch, combinations of both or, equally, the lack of movement that results in pitch stasis. Van Leeuwen argues that rises in pitch correspond to greater vocal effort and hence he refers to rises in pitch as *activation*, where “the more the pitch rises the more active and interactive the participants involved in its production and reception will be” (1999: 111). Falls in pitch, he argues, correspond to a decrease in vocal effort and hence realise *deactivation*, where participants are “brought into some state of non-activity – relaxation, contemplation and so on” (van Leeuwen

1999: 111). Cooke, meanwhile, suggests that descending melodies characterise incoming emotion and a sense of holding feelings in, where ascending melodies characterise the expression of outward emotion (1959; cited in Machin 2010).

According to van Leeuwen, the pitch range is associated with the extension of emotion (1999). He states that an increased pitch range allows for a greater expression of attitudes and feelings, whilst a decreased pitch range signifies emotional containment, which can signal a lack of energy or be a sign of deliberate containment. As Machin succinctly puts it, “a large pitch range means letting more energy out, whereas a small pitch range means holding more energy in” (2010: 102).

The pitch level can be either high key or low key and it can be combined with loudness or softness to create different effects. *Dominance*, argues van Leeuwen, is characterised by loudness and a high pitch level, and it is used most often by men to dominate (1999).

When women go high and loud, however, it often results in the “shrill and strident fishwife stereotype” (van Leeuwen 1999: 134). When a high pitch level is combined with softness, it results in *diminution*. This is due to the association of high-pitched sounds with small people and things. However, van Leeuwen also points out that “pitch level is strongly gendered” (1999: 111). Hence, a low pitch level combined with loudness realises *dominance* for women, whilst a low pitch level combined with softness imbues the woman with a sense of *danger*.

The *articulation* of the pitch, meanwhile, can either be *disjunctive*, in which case each of the sounds is emphatic and separate from one another, or it can be *connective*, in which case the sounds all flow into one another.

The following section will explain how both the visual and musical frameworks outlined so far in this chapter were utilised in order to collect the data and test the hypothesis.

2.3 The application of the frameworks

2.3.1 The Texts

The four videos chosen for analysis comprise two videos made for the group *SNSD*, who are signed to the entertainment agency *SM Entertainment*, and two videos made for the group *2NE1*, who are signed to the entertainment agency *YG Entertainment*. As mentioned in chapter 1, these two labels differ markedly in terms of the style of their output. The four videos were chosen on the basis of them being the top two most viewed videos for each group on *Youtube*³. As of July 22nd 2015, *Gee* by *SNSD* had garnered 142,887,729 views, *I Got A Boy* by *SNSD* had garnered 135,846,2145 views, *I Am The Best* by *2NE1* had garnered 127,846,135 views and *Lonely* by *2NE1* had garnered 46,743,183 views.

2.3.2 Visual analysis: The use of the *Multimodal Analysis Video* software

After the videos were chosen, they were uploaded to the *Multimodal Analysis Video* software. System choices based on the framework of visual analysis outlined in sections 2.1.1 and 2.1.2 were then created and added to the existing library of system choices already on the software. A screenshot of this library can be seen in figure 2.11. On the left is the personalised list of multimodal systems, whilst on the right is a list of available system choices, which could be added to the personalised list of multimodal systems on the left whenever necessary.

³ The second most viewed *2NE1* video was the video for their song *Lollipop*, but this was a collaboration with the boy group *Big Bang*, so it was not included.

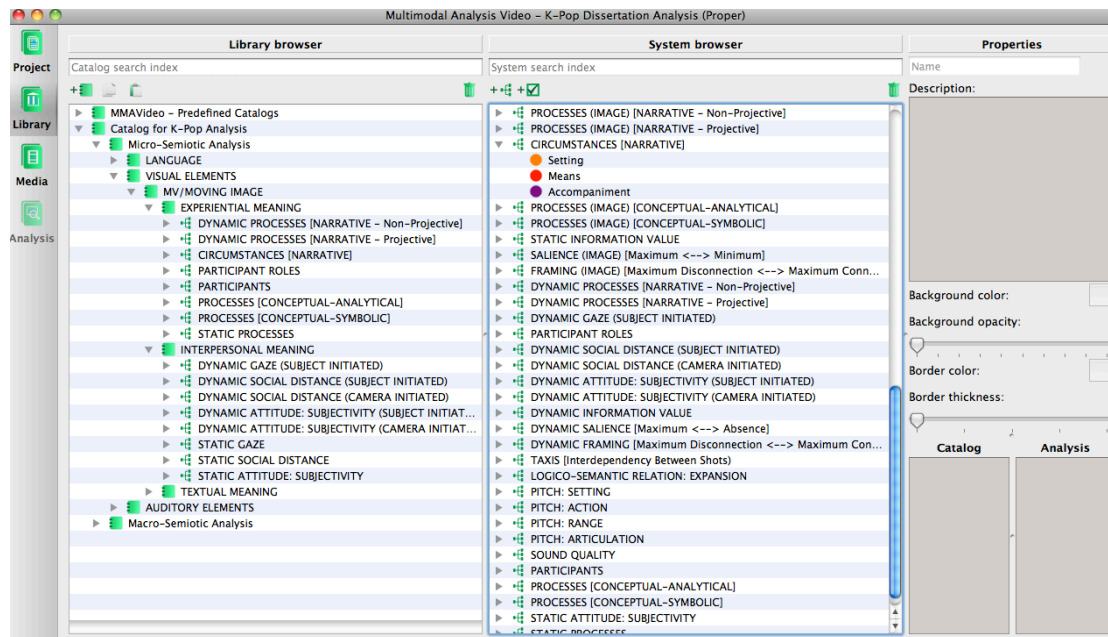


Figure 2.11 Screenshot of the *Multimodal Analysis Video* software systems library showing multimodal systems (left) and system choices (right).

Once the systems had all been added, the video could then be tagged using time stamped nodes that showed the selected system choices. The screenshot in figure 2.12 shows the guided user interface (GUI) that was used to analysis the music videos.

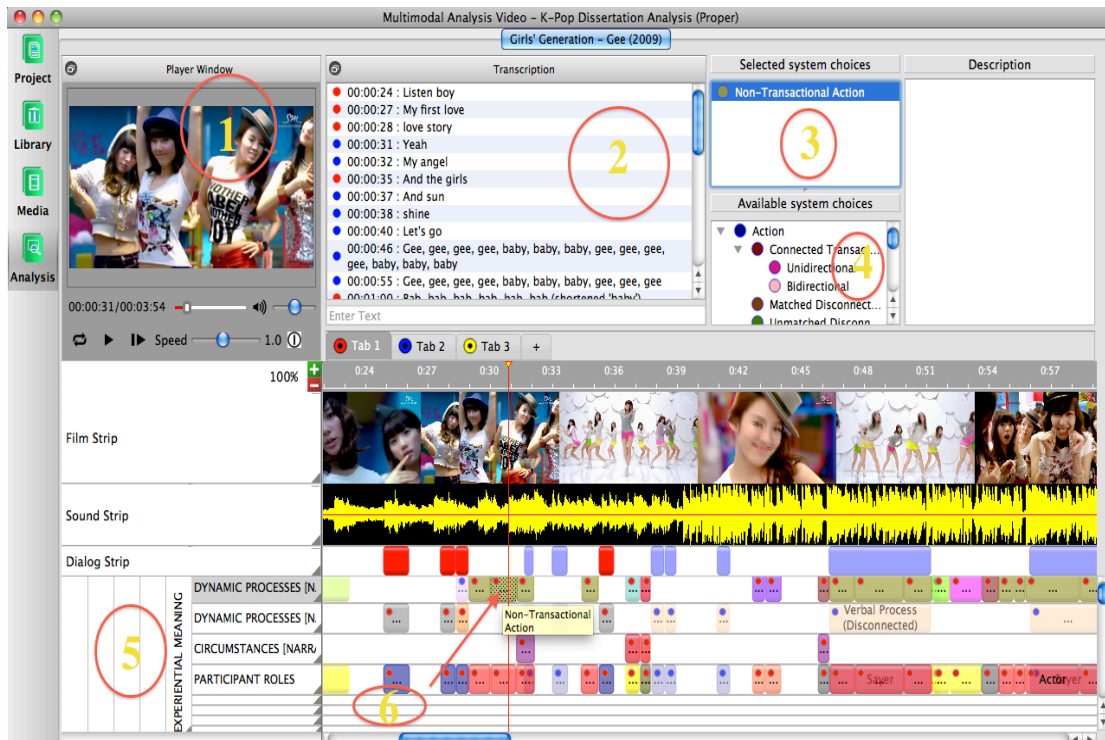


Figure 2.12 Screenshot of GUI on *Multimodal Analysis Video* software showing player window [1], transcription window with time stamped dialogue [2], selected system choice [3], available system choices [4], system strips based on the personalised list of systems chosen in the library [5], and time stamped nodes showing selected system choices [6].

Each of the individual systems [5] had its own strip that ran for the duration of the video. The nodes [6] were then positioned along this strip to reflect a certain choice made from the available system choices [4] at any given time in the video. The nodes could be stretched to match the duration of the choice in relation to the action taking place in the video. All of the system choices were inputted as time stamped nodes along the system strip for each video. The use of multiple tabs also permitted the transcription of simultaneous events in the video. This process was completed for both the experiential and interpersonal analyses. The topographical layout of the nodes allowed for comparisons of the choices in the experiential and interpersonal analyses to be made through simply moving the scroll bar on the right of the screen up or down.

When all of the choices for both the experiential and interpersonal analyses had been time stamped for all four of the videos, the analytical data could be exported to *Excel* for further analysis and tabulation. The software also includes a state transition machine.

This meant combinations of choices could be visualised as states and represented as a relative percentage of the overall video time. However, due to the simultaneity of choices often occurring in the experiential analysis for the four videos, multiple tabs had to be used. The simultaneity of action thus rendered any data showing combinations of action choices as a percentage of overall video time inaccurate. Hence, the state transition machine was only used for the interpersonal analysis of the videos, while a frequency count was chosen for the experiential analysis.

One of the great benefits of using the software was that it ensured that the more detailed investigations of the text, at least in the Interpersonal analysis, were data driven. It also facilitated the investigations of the videos through allowing one to “annotate the data within the same environment in which one accesses the data, via an annotation interface, and store these analyses in a database” (Smith et al. 2011: 361).

2.3.3 Musical analysis

For each of the four songs, musical phrases were identified and selected. The beginning and end of each phrase was determined by significant changes in the song, such as a change of singer or a change in the melodic pattern. All of the phrases that were selected occurred more than once in each song. However, in the case of the song *I Got a Boy* by SNSD, this did not apply. This song has a highly complex structure. Aside from the chorus and one other phrase, all of the phrases in this song occur only once. The majority of the phrases in this song, therefore, were randomly selected.

Once a shortlist of phrases for each song had been chosen, the sheet music for each song was obtained⁴. The sheet music for each of the four songs is arranged for piano. The sheet music for the song *Gee* by SNSD also includes vocals. In order to identify the

⁴ The sheet music for *Gee* (2009) was obtained from: <https://sonetown.wordpress.com/soshi-aholic/snsd-download-2/snsd-piano-sheets/>

The sheet music for *I Am the Best* (2011) was obtained from: <http://pianohousevn.net/sheet-piano/sheet-piano-i-am-best-2ne1.html>

The sheet music for *Lonely* (2011) was obtained from: <http://pianohousevn.net/sheet-piano/sheet-piano-lonely-2ne1.html>

The sheet music for *I Got a Boy* (2013) was obtained from: <http://www.jgmb829.com/sheets>
All of the sheet music was obtained on August 5th 2015.

melody line for each selected phrase, I listened to the songs whilst following the notation on the sheet music. The predominant key of each song was determined so that each note of the melody could be located numerically within the scale of that key and then plotted on a graph. Any accidental notes that occur in the melody lines of the songs are highlighted on the graphs⁵.

The values located on the vertical axis of the graphs represent the position of the notes on the scale. To determine the upper and lower limits of the vertical axis on the graphs, the highest and lowest notes in each song were identified. In addition, the vertical axes on the graphs for all of the phrases within each song were kept the same even if the song temporarily moved into a different key⁶. This was done in order to allow comparisons to be made between the phrases of a song.

The horizontal axis of the graph demonstrates the progression of the melody, with each value on it representing a note. In all of the graphs, the horizontal axis intersects with the lower 7th value on the vertical axis.

Finally, for each song, the identity of the tonic note (e.g. C♯) was recorded. This was done in order to allow comparisons to be made between the four songs.

In this chapter I have outlined the frameworks upon which the analysis in this dissertation was based. I have also explained how the frameworks were applied in the analysis of the four videos. The following chapter presents the findings of these analyses.

⁵ Accidental notes are notes that do not occur in the scale of the main key of the song. Where the note of the scale has been sharpened, it is shown on the pitch diagrams in green. Where it has been flattened, it is shown in grey.

⁶ Where substantial sections of a song moved into a different key, the colour system shown in footnote five was also used to express the new key signature so as to maintain consistency in the plotting of the melodic patterns.

CHAPTER 3

A MULTIMODAL DISCOURSE ANALYSIS OF THE MUSIC VIDEOS FOR *GEE* (2009) AND *I GOT A BOY* (2013) BY *SNSD*, AND *LONELY* (2011) AND *I AM THE BEST* (2011) BY *2NE1*

3. Presentation of findings

This chapter presents the findings of a multimodal discourse analysis of four female K-pop videos. The analysis shows how choices from the visual and auditory semiotic modes as well as the meaning potential afforded by the combination of both served to satisfy a male viewer in the case of the *SNSD* videos, but that in the case of the *2NE1* videos the semiotic choices that were made in the visual and auditory semiotic modes did not do this. In this chapter, each of the four videos will be considered in turn, starting with the two videos by *SNSD*. A short summary of the findings from the four videos will be presented at the end of the chapter.

3.1 *SNSD*

3.1.1 Analysis of *Gee* (2009)

3.1.1.1 Interpersonal analysis - State transition machine findings

The interpersonal analysis for *Gee* was initially guided by the results that were produced from the state transition machine. The first state transition diagram for *Gee*, which shows the combination of choices across the systems of STATIC CONTACT and STATIC ATTITUDE, can be seen in figure 3.1. Table 3.1 below shows the top five combinations of choices, which are also shown in the centre of the circle in figure 3.1.

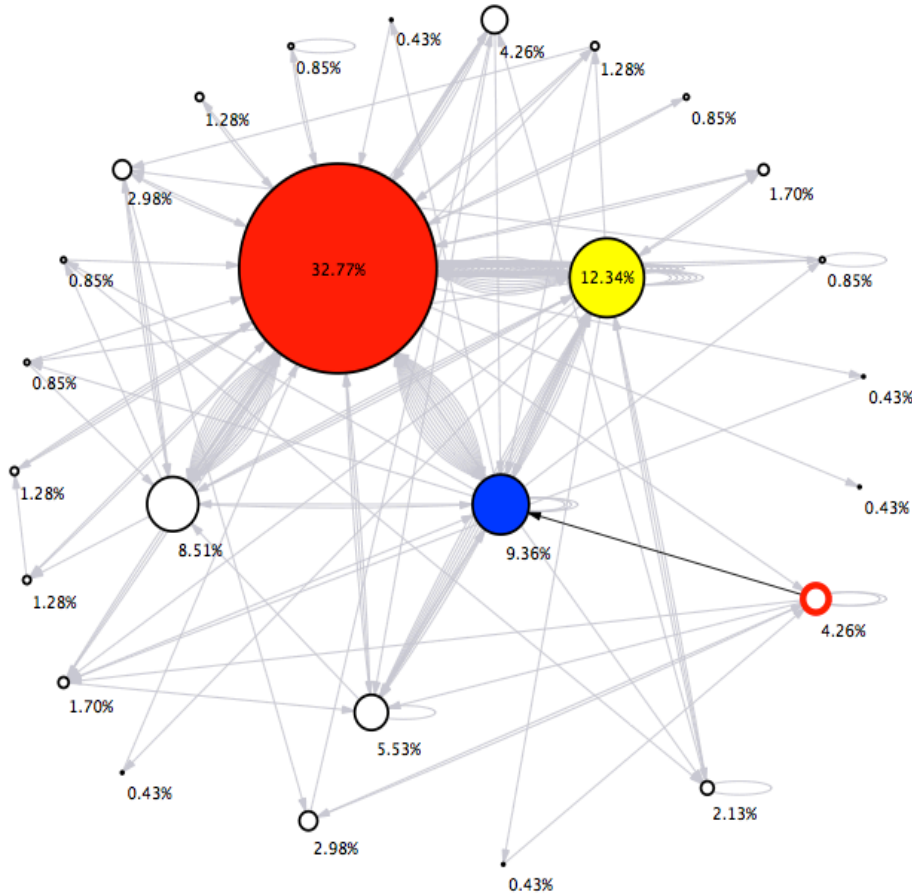


Figure 3.1 State transition diagram for *Gee* using the systems of STATIC CONTACT AND STATIC ATTITUDE

Percentage of video time realised by combination (Time realised by combination is shown in brackets)	Combination of choices
32.77% (1:16)	Demand, Frontal, Eye
12.34% (0:28)	Offer, Frontal, Eye
9.36% (0:21)	Offer, Oblique, Eye
8.51% (0:19)	Demand, Frontal, Low
5.53% (0:12)	Frontal, Eye

Table 3.1 Results from the state transition diagram for *Gee* for the systems of STATIC CONTACT AND STATIC ATTITUDE

The 2nd and 3rd most common combinations of choices, shown in yellow and blue in figure 3.1 were selected for further investigation. The most common combination of choices, shown in red in figure 3.1, meanwhile, was further investigated using a state transition diagram for all of the static systems. This is shown in figure 3.2 and table 3.2, where we can see that the combination of *demand*, *frontal* and *eye* choices most frequently occurs with the choice of *close shot* from the system of STATIC SOCIAL DISTANCE. This is shown in figure 3.2 in red, and for comparison the combination of *close*, *frontal* and *eye* with *offer* instead of *demand* is shown in blue. The difference between these two combinations shows that frontal interaction with the viewer at close range is one of the stylistically salient features of this video. Hence, this particular combination of *demand*, *close*, *frontal*, and *eye* was also chosen for further analysis.

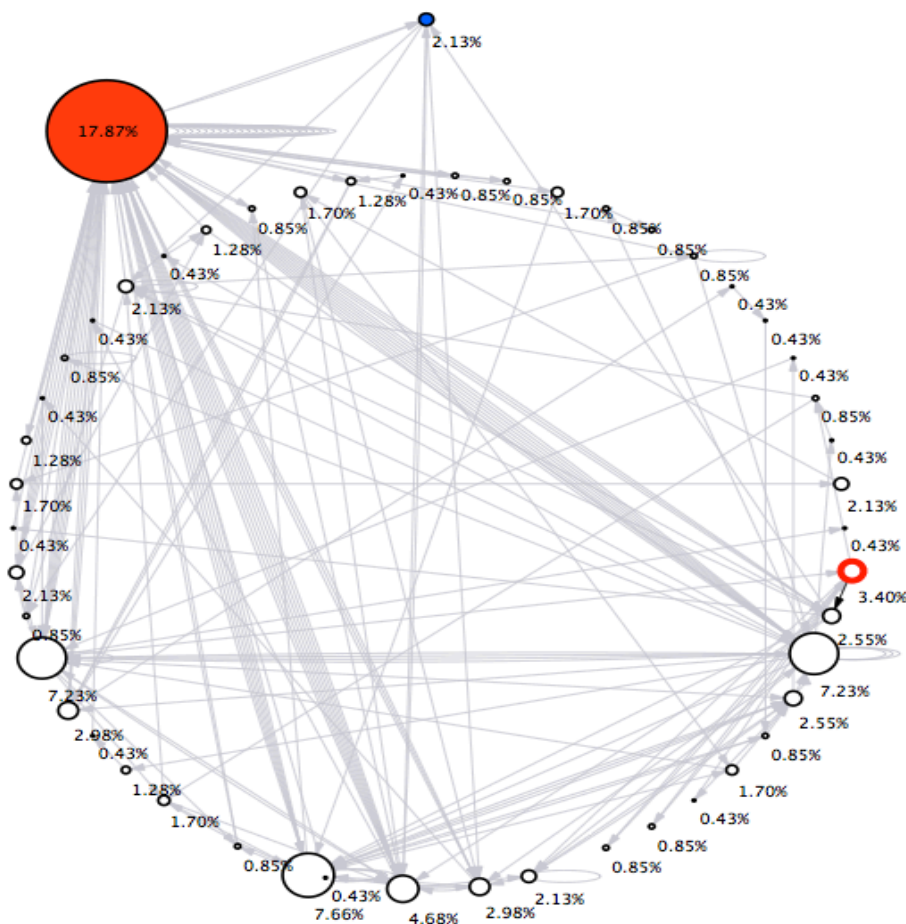


Figure 3.2 State Transition Diagram for *Gee* using the systems of STATIC CONTACT, STATIC ATTITUDE and STATIC SOCIAL DISTANCE

Percentage of video time realised by combination (Time realised by combination is shown in brackets)	Combination of choices
17.87% (0:41)	Demand, Close, Frontal, Eye
2.13% (0:04)	Offer, Close, Frontal, Eye

Table 3.2 Results from the state transition diagram for *Gee* showing choices from the system of STATIC CONTACT combined with the choices of *close*, *frontal* and *eye*.

3.1.1.2 Shots involving *offer*, *oblique*, *eye*

During the experiential analysis it was found that there were only 32 instances out of 190 transactions where group members featured as Actors in *connected transactional actions* (henceforth CTA). Interestingly, during the interpersonal analysis, it was found that 13 of these 32 CTAs co-occurred with the interpersonal system choices of *offer*, *oblique* and *eye*. Table 3.3 shows the times of each of these shots with a description of the actions taking place and a still of the video taken from each shot.

These 13 instances come from a particular selection of narrative scenes. The narrative of the video concerns the mannequins we see at the beginning of the video coming to life and engaging in a world of play in a clothes store. One of the central arguments I will make in the analysis of this video is that this ludic world of play is part of a discourse of infantilization that is designed for the pleasure of a male viewer.

One way this world of play is established is through the inclusion of make-believe actions. This can be observed in some of the shots in table 3.3. At 1:14-1:15, for example, we can see the girls⁷ *pretending* to punch the central girl. At 1:16-1:17, we can see the central girl making a gesture with her fingers on her cheeks *as if* crying.

⁷ *Girl* in this dissertation is used to stand in for *girl-group member*

Time of shot and CTA (Description of action shown in brackets)	Still from video	Time of shot and CTA (Description of action shown in brackets)	Still from video
0:37-0:37 (Girl on right writes on blackboard)		1:55-1:56 (Central girl puts a hat on)	
0:45-0:46 (Central girl bites sunglasses)		1:55-1:56 (Girl on left fans herself) *(Same shot as before)	
1:14-1:15 (Girls point and pretend to punch central girl)		1:59-2:00 (Girl centre right pushes girl left with hands connected)	
1:16-1:16 (Central girl points at her cheeks as if crying)		2:00-2:00 (Girl centre right pulls girl towards left of camera) *(Same shot as before)	
1:16-1:16 (Girls point at central girl) *(Same shot as before)		2:00-2:00 (Girl far right holds hand of other girl) *(Same shot as before)	
1:42-1:43 (Girls run towards girl on right)		2:05-2:06 (Girl on left bangs the wall playfully with her fists)	
1:53-1:53 (Central girl touches her cheek with palm of hand)			

Table 3.3 Table showing instances of CTAs co-occurring with the interpersonal system choices of offer, oblique and eye in the video for *Gee*

At 1:55-1:56 the girl on the left fans herself without the actual use of a fan, and at 2:05-2:06, the girl on the left playfully bangs the wall with her fists, again alluding to what might well be a more aggressive act of banging the wall in reality, but which here, in this video, never quite transcends the world of play.

Interestingly, the girls in these scenes are all presented to us from an oblique angle, and as Seabrook points out, this is no accident (2012). Discussing the manual of cultural technology devised by the founder of *SM Entertainment*, in which every detail concerning the way the stars signed to the label should be presented, Seabrook highlights the fact that the manual includes specific details concerning “the camera angles to be used in the videos” (Ibid). According to KvL, when participants are presented from an oblique angle, they become detached from the world of the viewer (2006). Here we can see that, as viewers, we are not encouraged to identify with the girls during these scenes, but are instead, through the choice of angle, positioned to merely observe them and their actions as a spectacle.

3.1.1.3 Shots involving *demand, close, frontal, eye*

54 shots out of the total 202 shots that were transcribed featured the combination of *demand, close, frontal* and *eye* choices. In 47 of these shots, there was just one girl present.

Many of the actions that take place during these interactive shots also pertain to a discourse of infantilization. Much of the action consists of playful gestures or facial expressions that are designed to communicate to the viewer simple instinctual feelings such as surprise, excitement, happiness or tiredness. What is evident is that the girls belong to a clear-cut juvenescent world that is without complexity. We can see some examples of this in table 3.4. At 1:00-1:01 and 1:25-1:26, the feeling is clearly one of surprise, at 1:37-1:37 a *bedtime* gesture is made and at 1:51-1:52 the girl makes herself smile with her index finger.




Time	Still from video	Time	Still from video
1:00-1:01		1:37-1:37	
1:25-1:26		1:51-1:52	

Table 3.4 Selection of *demand, close, frontal, eye* shots featuring gestures in the video for *Gee*

The sense of make-believe that runs throughout this video is also reinforced through the choice of hats that are worn by the girls during these *demand, close, frontal, eye* shots. Some shots featuring these hats can be seen in Table 3.5.

Time	Still from video	Time	Still from video
0:40-0:40		1:29-1:29	
1:18-1:19		2:04-2:05	

Table 3.5 Selection of *demand, close, frontal, eye* shots featuring hats being worn in the video for *Gee*

According to Owyong, the shape of accessories such as hats and the membership logos on them can help to communicate meanings interpersonally (2009). In the case of the hats worn in the shots at 0:40-0:40 and 1:29-1:29, we can see from their shape and design that they are navy hats and that the hat at 1:29-1:29 also has a membership logo with the word *Captain* written on it that underlines this point. Ordinarily, such a hat would confer upon its wearer a degree of respectability in relation to his or her rank in the navy. In the video, though, it is divorced from its usual context. The girls in the video wear the hat without having truly earned the right to do so and, thus, on the girls, the hat signifies the identity of the navy without any of the attached responsibility it ordinarily encompasses. This kind of role-playing is popular with children, and in this context it helps present the girls as being unthreatening, powerless and childlike, which feeds a male viewer's fantasy of superiority and control.

This sense of control and superiority is also affirmed through certain types of gesture that are directed at the viewer. At 0:40-0:40 in table 3.5, for example, we can see the girl saluting the viewer, at 1:18-1:19 we can see the girl winking, while at 1:29-1:29 we can see the girl pointing. This table is by no means exhaustive though. In the data it was discovered that 42 out of the 54 instances of *demand*, *close*, *frontal*, *eye* involved gesturing of some sort.

3.1.1.4 Musical analysis

So far only the meaning potential of the visual semiotic choices in this video have been discussed. However, much of the meaning of the song is also created through the auditory semiotic choices that are made in the video as well as the way in which these auditory choices sometimes interact with the choices made in the visual semiotic.

The first phrase under consideration runs from 0:40-0:50, and it is immediately repeated to form the verse. This verse occurs again at 1:40-2:00. A diagram of the melody for this phrase can be seen in figure 3.3.

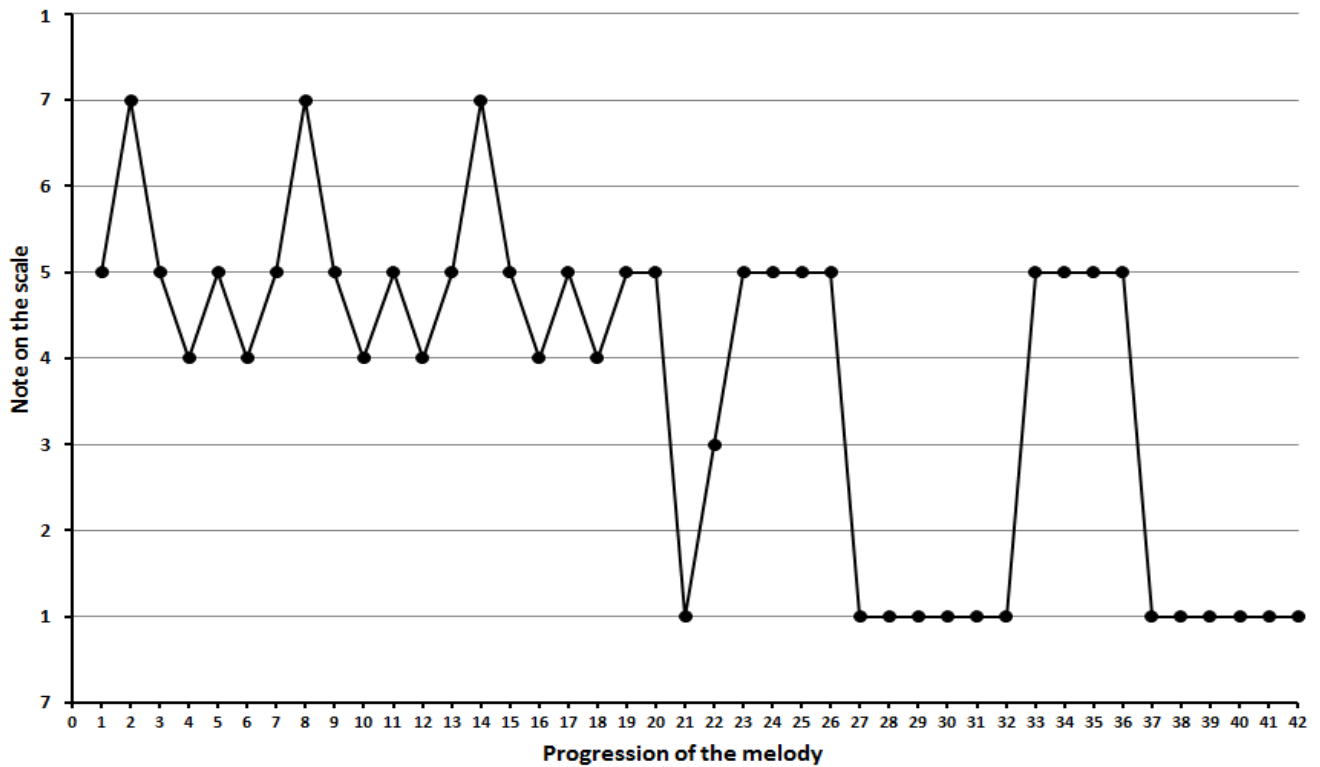


Figure 3.3 Melody of *Gee* [0:40-0:50] – (Repeated at 0:50-1:00 and 1:40-2:00)

Notes 1-20 (shown on the horizontal axis) constitute a decreased pitch range, which in van Leeuwen’s framework indicates *emotive confinement* (1999). Machin suggests that decreased pitch ranges are also associated with a holding in of emotions and modesty (2010). There are two scenes that occur during notes 1-20 that demonstrate how this holding in of emotion that takes place in the music is also realised and enhanced by the visuals. In the first scene, the girl in the centre looks at the viewer then quickly closes her eyes, pulling her arms towards her chest at the same time as if pulling in her emotion, while the girl on the right can be seen biting her nails. In the second scene, the girl pulls her raised arms close to her chest while clenching her teeth. These two scenes are shown in table 3.6.



Time of shot	Still from video	Time of shot	Still from video
0:44-0:45		0:52-0:54	

Table 3.6 Two shots from the video for *Gee* that work as visual complements to the decreased pitch range exhibited by notes 1-20 in figure 3.3

Holding in of emotion is traditionally associated with a discourse of passive femininity as opposed to a more active outgoing masculinity, which tends to be characterised by stepped increases in pitch. As mentioned in section 3.1.1.3, the powerlessness suggested by the way the girls act throughout the video, which is realised through a discourse of infantilization, ensures the assumed male viewer feels in control. Here, through the use of a decreased pitch range coupled with two scenes of the girls drawing their arms in to their chests, clenching their teeth and biting their nails, we can see how this sense of control and superiority is multimodally reaffirmed for the assumed male viewer in the face of passive and powerless girls.

The second phrase to be considered occurs from 1:00-1:19 and is repeated at 2:00-2:19. A diagram of the melody for this phrase can be seen in figure 3.4. In this phrase, there are 3 instances of stepped descending pitch. They occur at notes 1-5, 10-14 and 33-39. The phrase begins with 2 consecutive instances of stepped descending pitch. This, according to Cooke, is indicative of inward emotion (1959; cited in Machin 2010). In the video, the first note of the first instance of this stepped descending pitch (note 1) co-occurs with a surprised reaction by one of the girls coupled with a placing of her hands on her chest (1:00-1:01). This complements the inward emotion represented by the descent in pitch and further emphasises the powerlessness of the girls. A further example of this can also be seen occurring in the video (1:07-1:08) at the moment the first note of the third descent in pitch occurs (note 33).

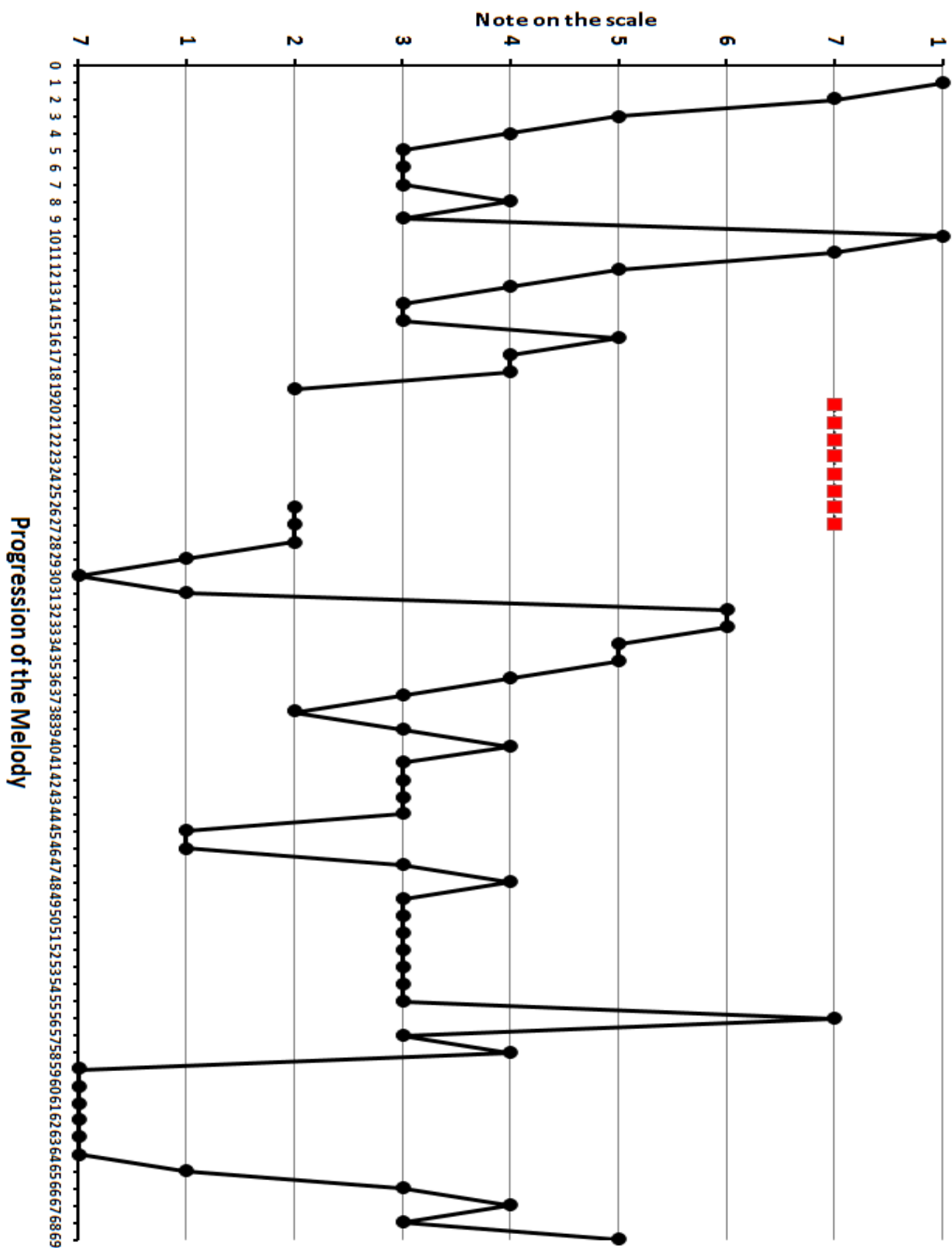


Figure 3.4 Melody of *Gee* [1:00-1:19] – (Repeated at 2:00-2:19)



Time of shot	Still from video	Time of shot	Still from video
1:00-1:01		1:07-1:08	

Table 3.7 Two shots from the video for *Gee* that work as visual complements to the beginning of two separate descents in pitch exhibited by notes 1 and 33 in figure 3.4

Here the girl in the centre claps both hands together as if praying or wishing for something, which is indicative of a mental process and, hence, inward emotion. Both of these examples can be seen in table 3.7.

Note 20 in figure 3.4 falls on the 2nd note of the scale. This is the first time that this note on the scale occurs in this phrase. Notwithstanding what is arguably a separate melodic phrase at notes 21-28 (shown in red in figure 3.4), you are effectively left on that note of the scale until note 26, where the same note of the scale is then repeated three more times before finally resolving around the tonic at the first note of the scale beginning at note 29. Through falling on the 2nd note of the scale, notes 20 and 26-28 express an extended need to resolve back to the tonic. This need to resolve followed by the return to the tonic expresses what van Leeuwen calls *continuity* (1999). Machin argues that this creates a feeling of limbo or entrapment (2010). In the video, at 1:04, we can clearly see that this feeling of limbo and entrapment is represented visually, as the girl, appearing to be distraught and concerned, twiddles her fingers and looks down to the ground (see figure 3.5).



Figure 3.5 Shot from the video for *Gee* (1:04-1:05) that works as a visual complement to the use of the 2nd note of the scale at notes 20 and 26-28 in figure 3.4

Representing this girl as trapped and distraught helps to empower the male through encouraging him to come to the rescue of the girl at her time of need. Hence the need to resolve back to the tonic is visually matched with the need to resolve the girl's distress. Furthermore, this brief and simple scene of distress reinforces the discourse of infantilization through complementing the other scenes of the video in which instinctual childlike feelings of happiness and surprise are expressed by the girls (see section 3.1.1.3).

In this analysis of *Gee* I have argued that the overall representation of the girls in the video is governed by a discourse of infantilization that satisfies a male viewer through offering up subjects who pose no threat or challenge to his assumed authority.

3.1.2 Analysis of *I Got A Boy* (2013)

3.1.2.1 Experiential analysis - Frequency count findings

The experiential analysis for *I Got A Boy* was initially guided by a frequency count of the *participants* and *processes*. A breakdown of the frequency for each of the *participants* can be seen in Table 3.8. A breakdown of the *action processes* and *reaction processes*

can be seen in Tables 3.9 and 3.10 respectively⁸. In addition to *action* and *reaction processes*, there was also 1 *filmic event*, 2 *mental processes* and 68 *verbal processes*.

Participant Role	Frequency
Actor	213
Goal	53
Reacter	89
Phenomenon	41

Table 3.8 Frequency of *participants* in the video for *I Got A Boy*

Action Processes	Frequency (Ordered from most to least)
NTA (Non-Transactional Action)	160
CTA (Connected Transactional Action)	51
MDTA (Matched Disconnected Transactional Action)	3
UMDTA (Unmatched Disconnected Transactional Action)	0
EVENT	0

Table 3.9 Frequency of *action processes* in the video for *I Got A Boy*

⁸ At 0:38-0:38 and 0:38-0:39, two types of *reaction* were coded for each: 1 CTR (Parallel Reacter) and 1 CTR (Reacter in background). Furthermore, one of the CTAs in table 3.9 was *bi-directional*.

Reaction Processes	Frequency (Ordered from most to least)
NTR (Non-transactional Reaction)	48
CTR (Connected Transactional Reaction) [Reacter in background]	21
CTR (Connected Transactional Reaction) [Parallel Reactor]	14
MDTR (Matched Disconnected Transactional Reaction)	5
CTR (Connected Transactional Reaction) [Reacter in foreground]	3
UMDTR (Unmatched Disconnected Transactional Reaction)	0

Table 3.10 Frequency of *reaction processes* in the video for *I Got A Boy*

Interestingly, throughout the video, a male was periodically present. The male was *actor* in 12 instances, *goal* in 4 instances, *reacter* in 2 instances and *phenomenon* in 12 instances.

In this analysis, I will focus on the scenes involving the male in order to show how the girls in the video are presented in relation to him and how this has a bearing on the way the viewer is positioned in relation to the girls.

3.1.2.2 *Matched Disconnected Transactional Actions*

One of the central arguments I will make in this analysis of *I Got A Boy* is that the viewer is repeatedly invited to virtually step into the shoes of the male character in the video, hence, willy-nilly, becoming the *boy* referred to in the title of the song. That the viewer is indeed positioned as a male is made clear very early on, where we are provided with a point of view (POV) shot through the door viewer shortly after we see the male trying the door handle (See Figure 3.6).



Figure 3.6 Three shots taken from the video for *I Got A Boy* at 0:21, 0:22 and 0:25

We can see how such positioning of the viewer is also achieved, albeit slightly less directly, through the three MDTA shots in table 3.11. There are only three instances of MDTAs being used in the video, but they all feature the male as *actor* and one of the girls as *goal*.







Duration of MDTA (Description of action shown in brackets)	Shot 1	Shot 2
2:18-2:20 (Male passes the girl a teddy bear)		
3:28-3:31 (Male makes the girl smile with his thumb)		
4:21-4:23 (Male puts his hand on the girl's hands, which are on her lap)		

Table 3.11 Table showing MDTAs in the video for *I Got A Boy*

On completion of each of the actions in the 2nd shot shown in table 3.11, the representation of the male is reduced to an absolute bare minimum. In the first instance all we have is a little show of hair in the top left hand corner, whilst in the second and third shots we only see a hand. The minimal presentation of the male in the first shot of each of these three transactions, meanwhile, provides us with just enough information to know that it is a male. KvL state that filmic disconnection can “radically disconnect *actors* from the *goals* of their actions” (2006: 259). Here, through disconnecting the *actor* and the *goal* using MDTA shots, and by reducing the representation of the male to an absolute bare minimum, the viewer is made to feel almost as if it is them completing the action rather than the male on the screen, leading us to feel as if we are the male responsible for the girl’s happiness in the first two examples.

Furthermore, in the third example from table 3.11, the girl’s reaction to the male’s hand being placed on hers is delayed. In between the completion of the action and her reaction to it, a short *demand, close, frontal, eye* shot is presented in which the same girl from the scene winks at the viewer. The progression of these three shots can be seen in figure 3.7.



Figure 3.7 Three consecutive shots taken from the video for *I Got A Boy* at 4:21-4:23, 4:23-4:23 and 4:24-4:24

The reaction in the 2nd shot, being abstracted from the diegesis of the scene and filmed with *demand, close, frontal, eye*, is suggestive of a brief internal thought on the part of the girl. Through interacting with the viewer and winking though, it seems to pander to a male fantasy of how a man would wish a girl to react to him putting his hand on hers. The third shot then represents her external reaction, which is presented to the viewer for contemplation by way of an Offer. Thus the (male) viewer is made to feel in control of

the action in the first scene, rewarded for that action in the second scene, and is then able to observe the result in the third scene. That this all happens in a matter of seconds is characteristic of the medium of music video, where meanings are made in a matter of seconds in between long periods of choreographed dancing.

3.1.2.3 Transactional Reaction Processes involving the male as *phenomenon*

Whilst the male is only the *goal* of the girl's actions 4 times in the video, he is the object of their looks in 12 transactions, which amounts to more than a quarter of the total of the 41 *phenomena* that feature throughout the video. What this shows is that whilst the girls may rarely act upon the man, they frequently do react to him. A list of the *Transactional Reaction Processes* featuring the male as *phenomenon* along with stills from the video can be found in Table 3.12.

In table 3.12 we can see that 8 out of the 12 instances involve the *reacter*, who is female in every case, being placed in the background. Furthermore, the two MDTR shots at 1:40-1:40 and 3:28-3:31 also feature the *reacter* in the background. This means that, as viewers, we are facing the girls and not facing the man in 10 out of 12 of these shots where the male is *phenomenon*, which ensures that we can clearly see the girls reactions and expressions in response to the male. Our view of the male in all of these processes, though, is always obstructed in some way.

This is achieved in four of the shots, at 1:38-1:39, 4:15-4:16, 4:16-4:17 and 4:31-4:32, by only showing the male's back or the back of his head. Whilst in the four shots at 2:19-2:19, 2:20-2:21, 4:41-4:41 and 4:42-4:42, we can only see one part of his body. In the shot at 2:24-2:24, the back of his head is shown completely out of focus. We only, in fact, even know it is him at all through anaphoric reference to the opening scene in which he rings the doorbell (0:10-0:10). In the MDTR shot at 1:40-1:40, the *phenomenon* of the shot is suggested textually through connection to the last but one shot, where we see the male (at 1:38-1:39) sat with his back facing us. The MDTR shot at 3:28-3:31, meanwhile, shows only a fraction of the male's hair and his arm and hand in the previous shot to which it is connected.













Time of Transactional Reaction	Reacter	Type of Transactional Reaction	Still from video
1:38-1:39	1 Female	CTR (Reacter in background)	
1:40-1:40	1 Female	MDTR	
2:19-2:19	1 Female	CTR (Reacter in background)	
2:20-2:21	1 Female	CTR (Reacter in background)	
2:24-2:24	1 Female	CTR (Reacter in background)	
3:28-3:31	1 Female	MDTR	
4:15-4:15	1 Female	CTR (Parallel Reactor)	
4:15-4:16	1 Female	CTR (Reacter in background)	
4:16-4:17	1 Female	CTR (Reacter in background)	
4:31-4:32	1 Female	CTR (Reacter in background)	
4:41-4:41	1 Female	CTR (Parallel Reactor)	
4:42-4:42	1 Female	CTR (Reacter in background)	

Table 3.12 Transactional Reaction Processes featuring male as *phenomenon*

Out of these 12 *Transactional Reactions* featuring the male as *phenomenon*, only 2 involve *parallel reactors*. Tellingly, in both of these shots, where the male and the female are side by side, the male is excised from the shot via careful framing. We can see how only the bottom corner of his face is showing at 4:15-4:15, and we can see how only his outstretched arm is presented in shot 4:41-4:41 in table 3.12.

The purpose of obstructing the view of the male in all of these shots is to downplay his agency and role in the private one-on-one scenes he participates in with the girls in order that we, the viewers, may feel as if we are close enough to be the one who is being looked at and reacted to by the girls. Additionally, it also allows the focus of our attention to be directed towards the girls themselves, hence satisfying the male gaze. This is facilitated by the use of *offer* acts in all of the shots presented in table 3.12. Thus, we are close enough to feel as if we are the male, whilst also remaining distant enough to enjoy the benefit of being able to stare at the girls' faces.

The notion that we are supposed to feel as if we are the represented male is also reinforced through certain *demand* acts that occur before and after some of the *Transactional Reaction Processes* that feature the male as *phenomenon*. In shot 2:24-2:24 we see the girl opening the door to the male. In the following shot, we are presented with the same girl staring into our eyes as part of a *demand* act. Through presenting the same girl who was in the previous shot with the same haircut, and through keeping the colour and lighting scheme the same, we are made to feel as if we have been directly transported into the role of the male. This can be seen in figure 3.8.

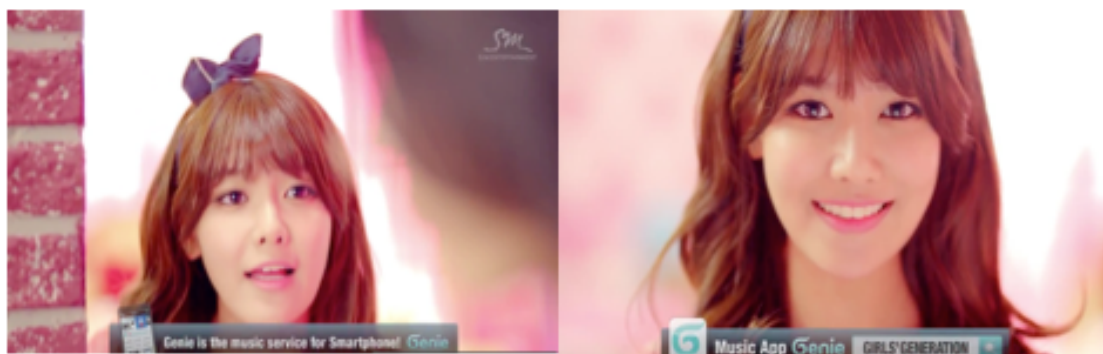


Figure 3.8 Two consecutive shots taken from the video for *I Got A Boy* at 2:24-2:24 and 2:25-2:25

The three *Transactional Reaction Processes* occurring at 4:15-4:15, 4:15-4:16 and 4:16-4:17, meanwhile, are sandwiched between two separate *demand* acts that both involve the girl in question pointing directly at the viewer. The *demand* act that occurs before the first *Reaction Process* at 4:15-4:15 can be seen alongside that process in figure 3.9, whilst the *demand* act that occurs after the last *Reaction Process* at 4:16-4:17 can be seen alongside that process in figure 3.10.



Figure 3.9 Two consecutive shots taken from the video for *I Got A Boy* at 4:14-4:14 and 4:15-4:15

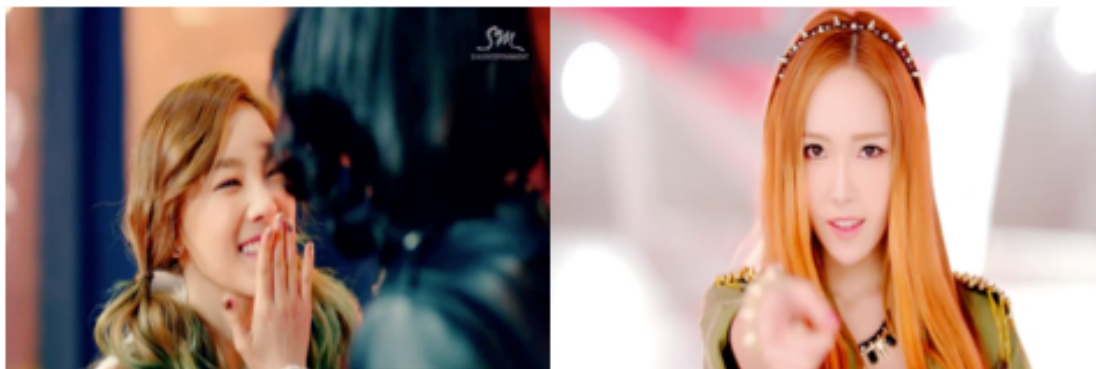


Figure 3.10 Two consecutive shots taken from the video for *I Got A Boy* at 4:16-4:17 and 4:17-4:17

In figure 3.9, it is as if she is saying that the guy you are about to see is actually you, the viewer. In this way the gesture works cataphorically, anticipating the scenes to come. The reverse of this then happens in figure 3.10, where we can see that the girl on the right once again points at the viewer, but this time it is as if she is saying that that guy *was* you. Hence, this gesture works anaphorically, referring back to the scenes before it.

In a number of the *Transactional Reaction Processes* featuring the male as *phenomenon* we have seen how our attention is drawn to the girls' faces and reactions through placing them in the background, facing the viewer, whilst the male is barely presented at all, often facing his back to the viewer or otherwise being presented in soft-focus or partially excised from the frame. As has already been pointed out, this serves two purposes. Firstly, it helps to downplay the role of the male so that the male viewer may feel as if he is the one being looked at by the girls, and secondly it satisfies a male viewer through drawing attention to the girls' faces. However, there are also aspects of the music that could be said to satisfy a male viewer.

3.1.2.4 Musical analysis

The phrase shown in figure 3.11 occurs only once and runs from 3:07-3:20.

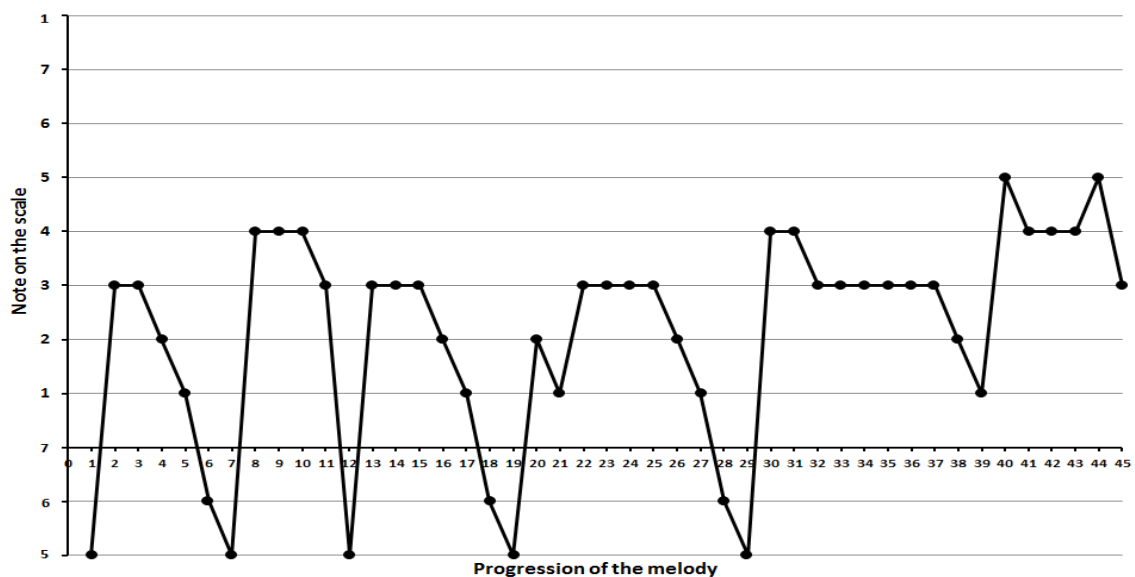


Figure 3.11 Melody of *I Got A Boy* [3:07-3:20]

The range of pitch for this phrase is large, which, van Leeuwen says, “allows us to give vent to strong feelings” (1999: 106). The fact that this phrase is characterised by three large stepped descents in pitch though, which all range from the third note of the scale down to the fifth (on the vertical axis) and occur from the 3rd note to the 7th, the 15th note to the 19th and the 25th note to the 29th (on the horizontal axis), means that the overriding

expression of the phrase concerns “incoming emotion”, which is indicative of reflection and contemplation (Cooke 1959; cited in Machin 2010: 101). This seems to suggest a private contemplative reaction to the acquisition of the boy that is alluded to in the title. Furthermore, the way the notes sensuously flow into each other during this phrase, exhibiting smooth *connective articulation*, adds to the gentle sense of contemplation and inner reflection already established by the stepped descents in pitch.

This world of inner reflection and privacy can be seen being represented visually in the video at 3:08-3:08 (see figure 3.12), whereby the girls whisper in each others ears at the exact same moment the first descent occurs from the 3rd note on the horizontal axis to the 7th in figure 3.11.



Figure 3.12 Shot taken from the video for *I Got A Boy* at 3:08-3:08 in which a visual complement to the descent in pitch exhibited by notes 3 to 7 in figure 3.16 occurs

Creating this mood of privacy and contemplation through this particular phrase of the music as well as through the visuals emphasises a passive model of femininity. Such a model of femininity naturally makes a male viewer feel more in control of the girls on the screen, empowering him in the process. As we shall see in the following section, this stands in stark contrast to the model of femininity favoured by *2NE1*.

In this section on the videos of *SNSD*, I have demonstrated how the video for *Gee* empowers a male viewer through presenting the girls as childish and infantilized, whilst the video for *I Got A Boy* is able to satisfy a male viewer through placing him in the position of the male character, who we barely ever see.

In the next section I will consider the two videos by the group *2NE1*, showing how the semiotic choices made in their construction do not serve to satisfy a male viewer.

3.2 *2NE1*

3.2.1 Analysis of *Lonely* (2011)

3.2.1.1 Experiential analysis – Frequency count findings

A frequency count of each of the *participants* and *processes* in the video for *Lonely* was first of all calculated. A breakdown of the frequency for the *participants* can be seen in table 3.13, a breakdown of the *Action Processes* can be seen in table 3.14⁹, and a breakdown of the *Reaction Processes* can be seen in table 3.15. In addition to *Action* and *Reaction Processes*, there are also 32 *Verbal Processes*.

Participant Role	Frequency
Actor	64
Goal	10
Reacter	61
Phenomenon	6

Table 3.13 Frequency of Participants in the video for *Lonely*

⁹ Two of the CTAs in table 3.14 are *bi-directional*

Action Processes	Frequency (Ordered from most to least)
NTA (Non-Transactional Action)	55
CTA (Connected Transactional Action)	11
EVENT	1
MDTA (Matched Disconnected Transactional Action)	0
UMDTA (Unmatched Disconnected Transactional Action)	0

Table 3.14 Frequency of *Action Processes* in the video for *Lonely*

Reaction Processes	Frequency (Ordered from most to least)
NTR (Non-transactional Reaction)	55
CTR (Connected Transactional Reaction) [Parallel Reacter]	6
CTR (Connected Transactional Reaction) [Reacter in background]	0
CTR (Connected Transactional Reaction) [Reacter in foreground]	0
MDTR (Matched Disconnected Transactional Reaction)	0
UMDTR (Unmatched Disconnected Transactional Reaction)	0

Table 3.15 Frequency of *Reaction Processes* in the video for *Lonely*

We can see from table 3.13 that there are very few *goals* and very few *phenomena* in this video. In terms of CTA's, we can see that in table 3.14 there are only 11 instances. Furthermore, it was found that only 5 of these are actually completed by members of the group. In terms of NTAs though there is a much higher figure, with 49 out of the 55 instances of NTA shown in table 3.14 being completed by the girls in the group. Overall, 45 out of the total 55 NTAs are realized by walking. Aside from walking though, very little action happens at all. The only other notable action concerns the girls moving their heads, which happens 10 times, and turning their bodies round, which happens 5 times. There are also two instances where the girls stand up. By way of comparison to the other

two videos that have so far been discussed it is also interesting to note that here in this video there are no instances whatsoever of dancing. Of the 55 instances of NTRs that occur in this video, meanwhile, it was found that 48 are by the girls in the group.

What these findings show is that the girls in the video are very disconnected from their immediate environment, which befits the title of the song. However, as we will come to see, they are also very disconnected from the viewer.

3.2.1.2 Interpersonal analysis – State transition machine findings

The state transition diagram for *Lonely*, which shows the combination of choices across the systems of STATIC CONTACT and STATIC ATTITUDE, can be seen in figure 3.13. Table 3.16 shows the top three combinations of choices, which are also shown in the centre of the diagram in figure 3.13.

These findings show that there is very little engagement with the viewer in terms of eye contact and in terms of the direction of the girls' bodies, which invariably face away from the viewer. Upon further investigation it was also found that all 8 instances of *offer, eye* co-occur with the system of (subject initiated) DYNAMIC ATTITUDE concerning the horizontal angle. These choices from the system of (subject initiated) DYNAMIC ATTITUDE play a greater role, then, in this video than in the previous two videos we have looked at.

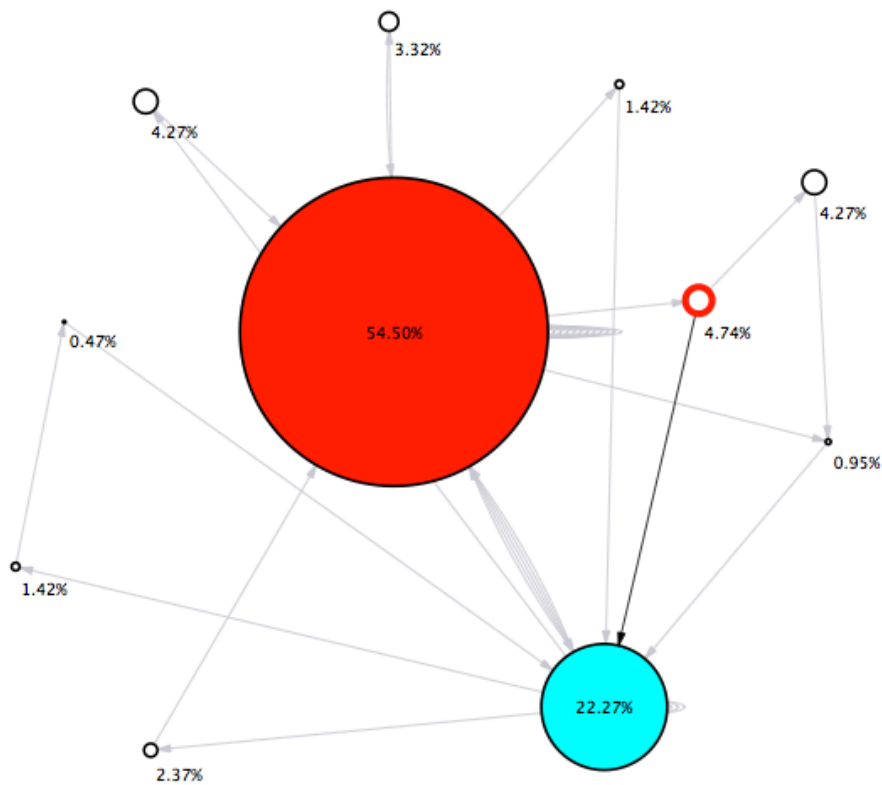


Figure 3.13 State transition diagram for *Lonely* using the systems of STATIC CONTACT and STATIC ATTITUDE

Percentage of video time realised by combination (Time realised by combination is shown in brackets)	Combination of choices
54.50% (1:54)	Offer, Oblique, Eye
22.27% (0:46)	Offer, Eye
4.74% (0:09)	Offer, Back-view, Eye

Table 3.16 Results from the state transition diagram for *Lonely* for the systems of STATIC CONTACT AND STATIC ATTITUDE

3.2.1.3 Avoidance of the viewer

Two of the most interesting discoveries in the analysis of *Lonely* concerned the way in which the girls appeared to avoid any interaction with the viewer. This was primarily achieved through an almost total absence of *demand* acts and through the use of an oblique angle. As KvL state, “the difference between the oblique and frontal angle is the difference between detachment and involvement” (2006: 136). They also state that representing *participants* using *offer* acts reduces the viewer’s role to being that of an “invisible onlooker”, hence reinforcing the sense of detachment already established by the use of an oblique angle (KvL 2006: 119). What is interesting to consider in the context of *Lonely* is the use of DYNAMIC ATTITUDE concerning the horizontal angle, which in 12 out of 14 instances in the video was initiated by the girls themselves, and the fact that the only two *demand* acts in the whole video were from the system of DYNAMIC CONTACT. In contrast to the use of STATIC CONTACT AND STATIC ATTITUDE (concerning the horizontal angle), where only the straightforward option of engagement with or detachment from the viewer exists, the inclusion of movement in these two systems allows for an element of *resistance* to be introduced into the relationship the girls have with the viewer.

One example of this resistance to interact with the viewer can be seen in the first instance of DYNAMIC CONTACT, which occurs with frontal angle at 0:39-0:44 in the video. The *demand* act issued by the girl in the shot is sustained for just two seconds, (from 0:39-0:41), before she then looks off screen to her left. She only turns her head again once the camera has moved into an oblique position and even then she keeps her eyes closed.

The virtual absence of the use of frontal angle also reveals the girls’ unwillingness to engage with the viewer. In all but one case, the frontal angle can only be found occurring during instances of subject initiated DYNAMIC ATTITUDE. There are 9 instances where frontal angle is involved in an instance of subject initiated DYNAMIC ATTITUDE, but only three of these instances, (at 0:10-0:20, 0:36-0:37 and 2:20-2:29), actually end with the girl in frontal position. The other 6 instances either involve the girl turning away from the viewer, or otherwise turning to face the viewer before then turning

away. Furthermore, of the three instances in which the girl does turn towards the viewer, in each case her head is either lowered or turned away from the camera. We can see the three examples of this occurring in figure 3.14 below.



Figure 3.14 Three shots taken from the video for *Lonely* at 0:13-0:13, 0:37-0:37 and 2:29-2:29

There is, in fact, only one instance in the whole video where a frontal angle without any movement occurs. This one instance occurs at 2:52-2:56 and it is also the only shot in the video where absolutely no identifiable objects can be observed. Hence, the one time we are invited in to their world, so to speak, the world presented is an abstracted one. We can see this clearly in figure 3.15.



Figure 3.15 One shot taken from the video for *Lonely* at 2:54-2:54

A resistance to engage with the viewer can also be observed in terms of SOCIAL DISTANCE. In five instances of DYNAMIC SOCIAL DISTANCE, only one involves moving towards the girl (at 2:32-2:41). The other four instances (at 0:50-0:59, 2:20-2:29

[in which both the subject and the camera move], and 3:16-3:20), involve the distance increasing between the girl and the viewer.

Furthermore, out of 13 close shots in the video, only 4 (at 0:47-0:50, 1:57-1:59, 2:01-2:11 and 2:29-2:32), involved an absence of movement, with the other 9 shots comprising 5 instances of DYNAMIC ATTITUDE and 4 instances of DYNAMIC SOCIAL DISTANCE. Hence, any sense of intimacy that might have otherwise been created between the viewer and the girl is mitigated to some extent by movement. This stands in stark contrast to the proliferation of *demand*, *close*, *frontal*, *eye* shots found in the video for *Gee*, where intimacy and engagement with the viewer was a necessary part of satisfying the male gaze. This video, by comparison, offers no such engagement with the viewer. There is no sense of any attempt to please, satisfy or entertain any viewer here, male or otherwise.

The only time, in fact, that any significant engagement with the viewer does occur is when a *demand* act is issued in the final shot, which lasts for 9 seconds. Here the four girls of the group are represented together for the first time, looking over their shoulders at the viewer. KvL point out that the degree of involvement or attachment that is communicated by the use of the horizontal angle can interact with the use of *demand* and *offer* acts in interesting ways (2006). They argue that it creates a double message: “although I am not part of your world, I nevertheless make contact with you, from my own, different world” (KvL 2006: 138). I would argue that in this final shot of *Lonely*, the girls uniting for the first time combined with the movement of their bodies and their heads towards the viewer suggests a degree of defiance and solidarity. To be sure, they are turning to make contact with us from their world, where it is not so much a case of ‘I Got A Boy’, but more a case of ‘together we can survive without one.’

3.2.1.4 Musical Analysis

This sense of solidarity established in the final shot is also present in one of the phrases of the song that was analysed. This phrase runs from 0:20-0:38 and can be seen in figure 3.16.

In this phrase, the fall to the lowest note at the 14th, 17th and 43rd notes on the horizontal axis coincides with a switch from singing in unison to singing solo. Unison singing, therefore, exists throughout most of the phrase. This is different to the visuals, where the girls' loneliness rather than their togetherness is emphasised throughout. In this way the music helps to reinforce a sense of solidarity between the girls that is not realised visually until the final shot.

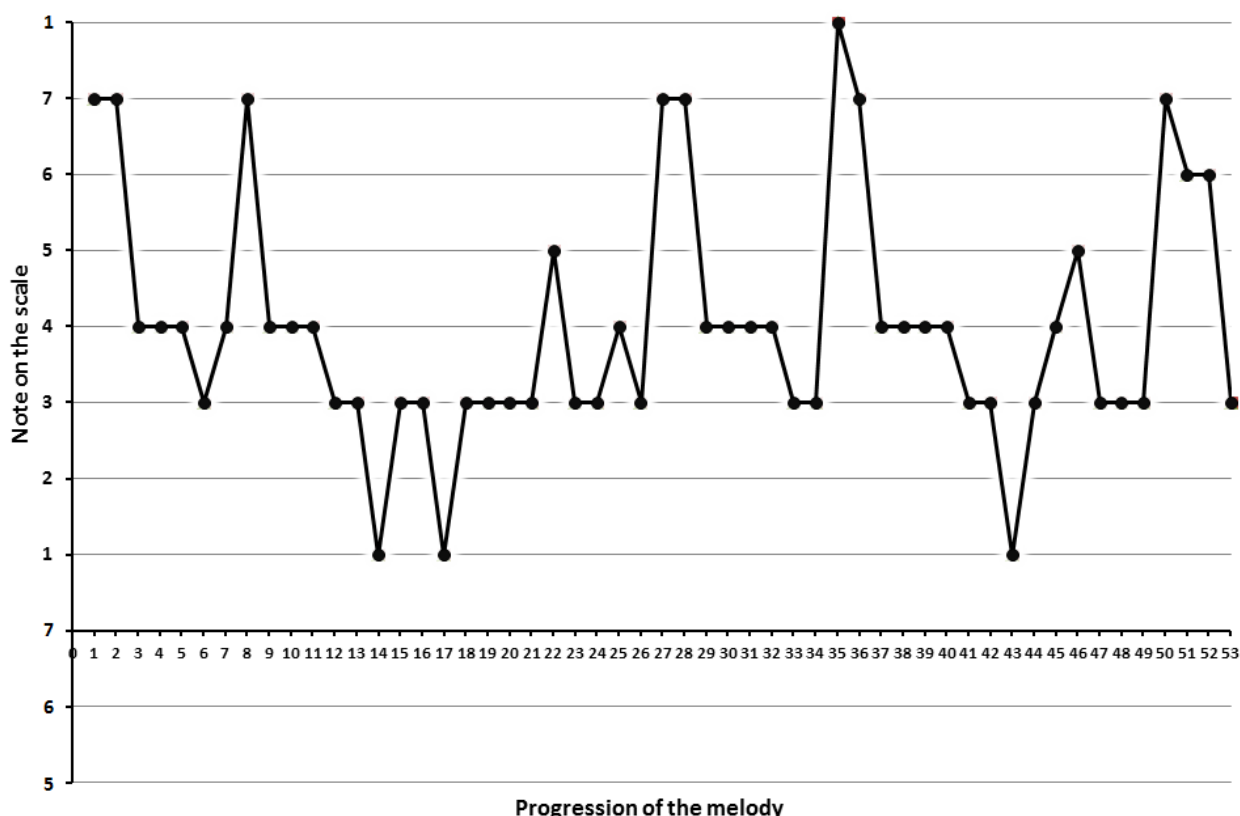


Figure 3.16 Melody of *Lonely* [0:20-0:38] – (Repeated at 1:43-2:01)

In terms of the pitch, this phrase, like the others that were analysed for this song, represents a fairly restricted degree of pitch movement, for although the range between the highest and lowest notes is an octave, the phrase itself consists of clusters of repeated notes that fall on the 3rd and 4th notes on the scale. Nevertheless, in amongst these clusters of notes, there are occasional rises and falls. This can be seen from notes 18-23, where the 3rd note on the scale is repeated four times before a rise to the 5th note on the

scale and then a return to the 3rd note on the scale occurs. This limitation in pitch movement combined with the repetition of the notes and the occasional rises and falls of pitch in between helps to create a restrained expression of sadness and resignation. However, unlike the cascades in pitch we saw in both *Gee* and *I Got a Boy*, no sense of “incoming emotion” is emphasised here (Cooke 1959; cited in Machin 2010). The traditional notion of passive and private femininity is hence not explicitly alluded to in this phrase or any of the other phrases that were analysed for this song, meaning that a male viewer, in this sense at least, is not made to feel empowered.

In this analysis of *Lonely* we have seen how through the use of *offer* acts, the oblique angle and the movement of both the camera and the girls themselves, the girls have been presented as being detached from the viewer and, moreover, resistant to any attachment with the viewer. With respects to a male viewer, it is, hence, difficult to see how this video could possibly satisfy a male gaze. With respects to a female viewer, however, the notion that solidarity with other women exists in the face of loneliness without a man is emphasised musically through the unison singing as well as visually in the final scene, where the girls come together and collectively look over their shoulders at the viewer.

3.2.2 Analysis of *I Am The Best* (2011)

3.2.2.1 Experiential analysis – Frequency count findings

A breakdown of the frequency for each of the *participants* in the video for *I Am The Best* can be seen in Table 3.17. A breakdown of the *Action Processes* and *Reaction Processes* can be seen in Tables 3.18 and 3.19 respectively. In addition to *Action* and *Reaction Processes*, there were also 20 *Filmic Events* and 27 *Verbal Processes*. The high number of *Filmic Events* can be attributed to the moving objects that the girls sit on in the video.

Participant Role	Frequency
Actor	147
Goal	65
Reacter	59
Phenomenon	6

Table 3.17 Frequency of *participants* in the video for I Am The Best (2011)

Action Processes	Frequency (Ordered from most to least)
NTA (Non-Transactional Action)	90
CTA (Connected Transactional Action)	47
UMDTA (Unmatched Disconnected Transactional Action)	10
EVENT	8
MDTA (Matched Disconnected Transactional Action)	0

Table 3.18 Frequency of *Action Processes* in the video for I Am The Best (2011)

Reaction Processes	Frequency (Ordered from most to least)
NTR (Non-transactional Reaction)	53
CTR (Connected Transactional Reaction) [Reacter in background]	5
CTR (Connected Transactional Reaction) [Reacter in foreground]	1
CTR (Connected Transactional Reaction) [Parallel Reacter]	0
MDTR (Matched Disconnected Transactional Reaction)	0
UMDTR (Unmatched Disconnected Transactional Reaction)	0

Table 3.19 Frequency of *Reaction Processes* in the video for I Am The Best (2011)

3.2.2.2 Circumstance: Means

There were a total of 80 *Circumstances of Means* occurring in the video, excluding 24 transactions involving dancing and 1 *Event* transaction. *Means* were broken down into those involving just body parts, and those involving either body parts and objects, or just objects. Table 3.20 shows a list of the objects used as *Circumstances of Means* in the video for *I Am The Best*.

Object used as Circumstance of Means	Frequency (Ordered from most to least)
Gun(s)	17
Baseball Bat(s)	9
Drumstick(s)	9
Mirror	3

Table 3.20 Objects used as *Circumstance of Means* and their frequency

The top three objects in this list are all associated with speed, action and movement. Furthermore, guns, sport and drumming are, traditionally at least, more often associated with men than with women. The top two objects in this list are also heavily linked to violent action, as they are commonly used as weapons in warfare or, in the case of the baseball bat, fights. Both guns and baseball bats, then, very clearly connote violence and threat, and in this video both are indeed used as part of violent actions. Whilst the *goal* of the action in which the gun is used as *means* is never shown clearly though, the *goal* of the action in which the baseball bat is used as *means* is shown clearly, where in 5 instances it is even shown in the same shot. Figure 3.17 shows one example of the violent use to which the baseball bat is put in this video.



Figure 3.17 Shot taken from the video for *I Am The Best* at 2:28-2:28

As we can see, both the guns and the baseball bats are being used as part of violent actions undertaken by the girls in this video. Whilst the girls in the video for *Lonely* avoid any action for the most part, the girls in this video, through their actions, pose a direct threat to any viewer who chooses to watch them. Such a reactionary stance naturally challenges the male gaze, which ultimately relies upon the submissiveness of the represented girls.

One object that does appear to sit uneasily amongst the other objects in this list, however, is the hand mirror. It could be said that, traditionally at least, hand mirrors are more often associated with women. In this sense, the inclusion of this object in the list stands out. Interestingly, all 3 instances of the mirror being used as a *Circumstance of Means* occur within 11 seconds of each other in this video, from 2:24-2:35. The reason for the inclusion of the mirror within this 11-second time frame will be discussed in greater detail in the next section on *Reaction Processes*.

3.2.2.3 Reaction Processes

In this video, the girl group members are *reacters* on 52 occasions. However, there are only 6 CTRs in the whole video, and they all occur within 15 seconds of each other. The mirror is *phenomenon* in 5 of these CTRs, and the large diamond is *phenomenon* in one of them. What is interesting to observe is that the section in which the mirror is used is arguably the most violent part of the video. This period runs from 2:24-2:40 and during this part there are 5 CTAs, 2 UMDTAs and 1 EVENT that are all associated with the glass cabinet getting smashed by the girls with baseball bats. Interspersed between these violent actions are, unusually, shots in which the girl (second from left at the back) preens herself in the mirror.

My contention is that the girl preening herself in the mirror during these scenes constitutes a parody of other K-pop girl groups such as *SNSD*, in whose videos such preening routinely takes place. My support for this argument comes from the data gathered as a result of the multimodal discourse analysis that was undertaken on this video. Had the data shown the mirror turning up in other transactions during other parts of the video, the use of the mirror during these violent scenes, whilst still unusual, would have been less pointed, but given that these are the only scenes in which the mirror occurs, and given that no other objects that could be said to be traditionally associated with femininity are used in this video, I would argue that the inclusion of the mirror in this video is marked and that the acts associated with it are intentionally parodic.

Interestingly, in relation to this point, at 2:00-2:01 there is an instance of a *demand* act in which one of the girls, using the *camera* as a mirror, preens herself. However, no sooner has this started, than she shoos off the viewer with her hand in the following shot. In both of these scenes, it would appear, attention is being drawn to hyper-femininity in order to subvert it, for not only is the act of preening re-contextualised in this video through juxtaposing it with violence, its very necessity in music videos at all is questioned. In the wider context of a highly gender divided country like South Korea, where female K-pop stars are expected to act in a hyper-feminine way in their videos, such an act becomes rather subversive, which is characteristic of the video as a whole. This video, then, naturally acts as a direct challenge to the male gaze, which comes to

accept such hyper-feminine behaviour patterns as the norm, in the context of female K-pop videos at least.

3.2.2.4 Interpersonal analysis – State transition machine findings

The interpersonal analysis for *I Am The Best* was guided by the results that were produced from the state transition machine. The state transition diagram for *I Am The Best*, which shows the combination of choices across the systems of STATIC CONTACT, STATIC ATTITUDE and STATIC SOCIAL DISTANCE, can be seen in figure 3.18. Table 3.21 shows the top six combinations of choices, which are also shown in the centre of the circle in figure 3.18.

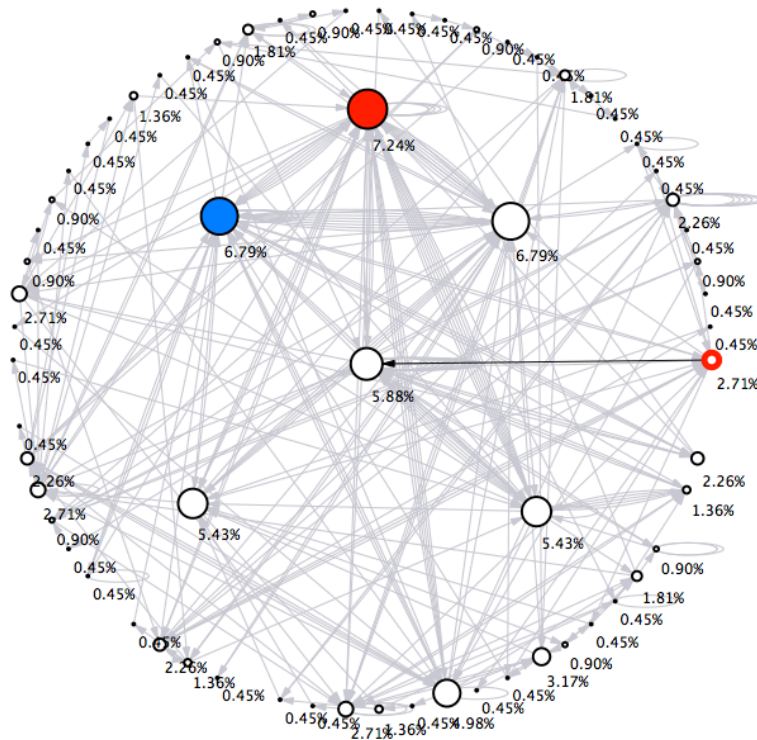


Figure 3.18 State transition diagram for *I Am The Best* using the systems of STATIC CONTACT, STATIC ATTITUDE and STATIC SOCIAL DISTANCE

Percentage of video time realised by combination (Time realised by combination is shown in brackets)	Combination of choices
7.24% (0:15)	Close, Frontal, Eye
6.79% (0:14) [Blue]	Demand, Close, Frontal, Eye
6.79% (0:14)	Demand, Medium, Frontal, Eye
5.88% (0:12)	Long, Frontal, Eye
5.43% (0:11) [Left]	Demand, Long, Frontal, Eye
5.43% (0:11) [Right]	Medium, Frontal, Eye

Table 3.21 Results from the state transition diagram for *I Am The Best* for the systems of STATIC CONTACT, STATIC ATTITUDE and STATIC SOCIAL DISTANCE

In table 3.21, as we can see, the 1st, 4th and 6th most common combinations involve no *demand* or *offer* acts. This is due to 24 instances of DYNAMIC CONTACT as well as 82 shots in which no choices from the system of CONTACT existed at all. The top two combinations of choices shown in table 3.21 were chosen for further analysis. The first combination, shown in red in figure 3.18, involves no CONTACT, and the second combination, shown in blue in figure 3.26, involves *demand* acts.

3.2.2.5 Shots involving *close, frontal, eye*

There were 7 instances overall of *close, frontal, eye* co-occurring with choices from the system of DYNAMIC CONTACT, and there were 19 instances of *close, frontal, eye* co-occurring with no CONTACT. In this section I will focus on the 19 instances of *close, frontal, eye* co-occurring with no CONTACT.

One of the primary functions of the *close, frontal, eye* shots with no CONTACT in this video seems to be to draw attention to the clothing that the girls wear. There are 13 instances where particular items of the clothing are foregrounded through the use of close

shots combined with frontal angle at eye level. This is not altogether surprising as clothing plays an important part in the sense of threat that is created in this video. Much of the sense of threat that is connoted in this video by the clothes is due to their direct association with violent acts. In the shots at 0:14-0:15, 0:15-0:15 and 0:20-0:21, for example, the girl touches the hood of a wrestling dressing gown she is wearing, whilst at 1:09-1:10, 1:12-1:13 and 1-13-1:14 the girl in these shots is shown wearing rings that are almost like knuckle-dusters. Furthermore, at 1:09-1:10, we can see her flicking one of her rings and making it spin around her finger whilst wearing a chainmail headpiece. Chainmail was traditionally worn in battles as a protective undergarment. All three of these clothing items, then, have direct associations with violent acts, helping to add to the sense of confrontation already established in the video.

However, the clothing and accessories the girls wear also connote a sense of threat due to the materials they are made from. In all 13 of these *close, frontal, eye* shots involving clothing, metal is being worn. Metal is often used to make guns and knives and other such weapons. Furthermore, unlike the delicacy of fabric, metal is a tough and heavy material that does not move or sway with the wearer. Wearing metal therefore involves a greater degree of physical exertion than ordinary clothes. Through wearing metal, then, the implication here is clear – these girls are strong.

Owyong states that, “the look created by a combination of various articles of clothing and accessories conveys a “message” to the viewer” (2009: 196). In these *close, frontal, eye* shots we have seen a few examples of how the clothing in this video plays an important part in the message of antagonism and confrontation that is conveyed to the viewer by the girls. For a male viewer, such a message comes across as a violent and reactionary challenge to his (assumed) authority.

3.2.2.6 Musical Analysis

So far we have seen how objects used as *Circumstance of Means*, clothing and actions have all helped to contribute to the sense of threat and aggression established in the video. This sense of threat is also transmitted through the music. An example of this can be seen in the phrase in figure 3.19.

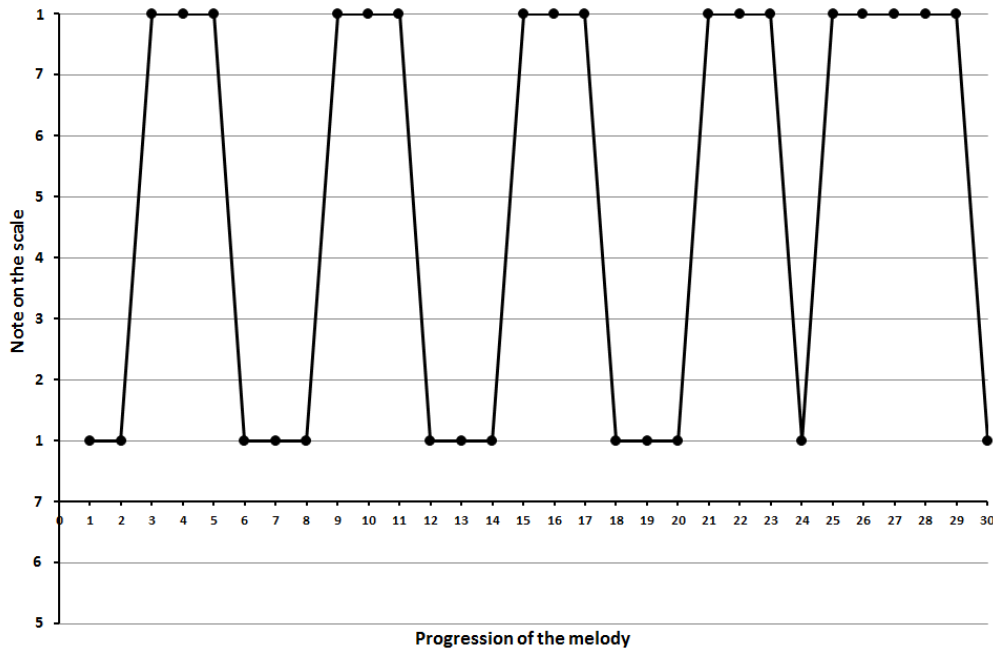


Figure 3.19 Melody of *I Am The Best* [0:09-0:23] – (Repeated at 1:24-1:38 and 2:24-2:38)

The pitch range of this phrase covers an octave. This is quite large, suggesting emotive expansion, which in the case of this phrase tends towards aggression. The ascents and descents in pitch are also not stepped, instead jumping and falling by an octave each time, which creates a jerky abrupt feel. This can be seen occurring on notes 2, 3, 5 and 6 on the horizontal axis. There is also repetition of the top 1st note on the scale on the vertical axis. This top 1st note on the scale is fairly high and is sung with a degree of loudness. This combination of high and loud in the phrase constitutes what van Leeuwen refers to as *dominance* (1999). Van Leeuwen states that this is the way of speaking or singing usually favoured by men (Ibid). Hence, we can see that the girls in this song are adopting a masculine stance when they sing like this. During this phrase, their voices also demonstrate a good deal of tension in terms of the voice quality. This tension combined with the loudness of their voices and the use of high pitch adds an aggressive edge to the phrase, which is emphasised by the sharp rises and falls between the two 1st notes on the scale. Hence, this rise in pitch and the range it covers combined with the fact that it falls on the two 1st notes on the scale ensures that a firm and confident sound is achieved. The tenseness of the voice when added to this results in an effect of assertiveness, which is further emphasised in the phrase through the mechanical repetition of the sung notes. This phrase forms the chorus of the song and is repeated three times. On the third

occurrence of the phrase, the aggression in the music is matched with the aggression on the screen, as the girls smash the glass disc cabinets.

In this section looking at the videos of *2NE1*, we have seen how neither of the videos explicitly serves to satisfy a male viewer. In the case of *Lonely*, a prevalence of oblique angles and *offer* acts are used to create a sense of detachment between the viewer and the represented girls, whilst choices from the DYNAMIC systems help to create a certain resistance towards any engagement with the viewer. However, in the video for *I Am The Best*, a mood of aggression is established primarily through the clothing and the music as well as through the girls' actions, which are frequently of a violent nature.

3.3 Summary of findings

In this analysis we have seen how the two *SNSD* videos and the two *2NE1* videos differ markedly in terms of the kind of relationships they form with the viewer. Whilst the two *SNSD* videos clearly appeal to a male viewer, the two *2NE1* videos clearly do not.

In the *SNSD* video for *Gee*, I argued that a discourse of infantilization is created in order to empower a male viewer. I argued that this is achieved through playful actions, which are observed obliquely ensuring we are not made to identify with the ludic world in which they occur, through the communication of basic childlike emotions such as happiness and sadness, which help to convey a world of juvenescence, and through the sense of make-believe and role play, which we saw emphasised through the use of Navy hats. In the music, meanwhile, I showed how large descents in pitch combined with images of the girls clenching their teeth, biting their nails or looking lost helped to multimodally convey to the viewer a sense of powerlessness as well as a sense of childish emotion, hence further empowering the male viewer.

In the video for *I Got A Boy*, I argued that the male viewer is gratified through placing him in the position of the male character on screen. I showed how this was achieved through using MDTAs to lessen the male's sense of agency and through reducing the representation of the male in all shots to a bare minimum. I also showed how the male gaze was gratified through bringing the focus on to the girls' faces in *Transactional*

Reaction Processes, where the girls are frequently looking at the male character from the background, thus facing the viewer. Furthermore, instances of *demand* acts in which the girls point at the viewer prior to and after certain scenes with the male were also pointed out as being examples of appealing to a male viewer. Finally, I showed that the music, like *Gee*, features large stepped descents in pitch, which is indicative of a sense of contemplation and private reflection.

The video for 2NE1's song *Lonely* differed markedly from all of the other videos. It was unique in its frequent use of oblique angles combined with *offer* acts as well as through the fact that, unlike the other three videos analysed, it lacked any dancing. A general avoidance of the viewer, or otherwise a resistance to engage with the viewer, was conveyed through the use of DYNAMIC ATTITUDE, which was higher for this video than any of the others. This lack of engagement with the viewer was also shown through the virtual absence of frontal angle shots. The only engagement with the viewer, in fact, occurs at the end of the video, where the girls unite and look over their shoulders at the viewer. I argued that this communicated a sense of solidarity with other women, which was further emphasised through the use of unison singing during the song.

Finally, in the video for *I Am The Best*, I showed how the semiotic choices made in both the visual and auditory modes help to convey a palpable sense of danger and threat. This is achieved through the clothes the girls are wearing as well as through the violent actions they engage in, where baseball bats and guns are involved as *Circumstances of Means*. The music, it was pointed out, adds to this mood of confrontation through jerky rises and falls of pitch and the tenseness of the girls' voices as they sing loudly at a high level of pitch.

CHAPTER 4

CONCLUSION

This dissertation has reported on a multimodal discourse analysis of four female K-pop music videos. The hypothesis, which stated that the semiotic choices made in the two *SNSD* videos would serve to satisfy a male viewer whilst the semiotic choices made in the *2NE1* videos would not serve to satisfy a male viewer, has, within the limitations set by this dissertation, been proven.

The framework chosen for the analysis of the visual elements of the four videos was KvL's grammar of visual design (2006). Elements of this framework were also adapted in order to suit the medium of film using van Leeuwen's article on film as a guide (1996). The *Multimodal Analysis Video* software program, meanwhile, was used to conduct the *Experiential* and *Interpersonal* analyses of the visual elements of the four videos, which meant that subsequent investigations of the four videos were based on data rather than *cherry-picked* examples. The framework chosen for the musical analysis of the four videos centred on van Leeuwen's system network for pitch in *Speech, Music, Sound* (1999). The meanings created by the visual and musical choices as well as the meanings arising from the combination of choices from both modes were considered in the analyses of all four videos.

In the analysis of the video for *Gee*, I argued that a discourse of infantilization is created in order to empower a male viewer. I suggested this was achieved through playful actions and gestures and the expression of basic childlike emotions such as joy and sadness.

In the analysis of *I Got A Boy*, I argued that the viewer is placed in the position of the male character on screen and is thus made to feel as if he is responsible for the girls' expressions of happiness, which are always shown directly to the viewer.

In the analysis of *Lonely* I showed how an avoidance of the viewer was achieved through the use of oblique angles and *offer* acts, whilst the use of DYNAMIC ATTITUDE also enabled a sense of resistance to viewer interaction to be communicated. The contact with

the viewer at the end, meanwhile, was suggested as being an expression of female solidarity in the face of a loneliness that results from not having a man.

Finally, in the analysis for *I Am the Best* I demonstrated how a sense of threat and confrontation was achieved in the video through the girls' actions and through the clothes they wore. I also suggested that a parody of hyper-feminine behaviour was enacted during the scene with the mirror.

Hence, whilst in the two *SNSD* videos the representation of the girls can be said to be designed to satisfy a male viewer, in the two *2NE1* videos the girls are represented as either refusing to acknowledge a male viewer, if we can indeed assume that it is one, or otherwise as directly challenging his assumed authority.

REFERENCE LIST

- Choi, Y.J. (2011) The Globalization of K-Pop: Is K-Pop Losing its Korean-ness? **Situations**, 5, 69-75.
- Cooke, D. (1959) **The Language of Music**. Oxford: Clarendon Paperbacks.
- Djonov, E. and Zhao, S. (2014) *Critical Multimodal Studies of Popular Discourse*. New York: Routledge. pp. 1-14.
- Halliday, M. A. K. (1978). *Language as social semiotic*. London: Arnold.
- Halliday, M.A.K. & Matthiessen, C. (2004) **An Introduction To Functional Grammar**. 3rd ed. London: Hodder Arnold.
- Jang, W. and Kim, Y.S. (2013) Envisaging the Sociocultural Dynamics of K-pop: Time/Space Hybridity, Red Queen's Race, and Cosmopolitan Striving. **Korea Journal**, 53/4, 83-106.
- Jung, S. (2011) **Korean Masculinities And Transcultural Consumption: Yonsama, Rain, Oldboy, K-Pop Idols**. Hong Kong: Hong Kong University Press.
- Jung, S. and Hirata, Y. (2012) Conflicting Desires: K-pop Idol Girl Group Flows in Japan in the Era of Web 2.0. **ejcjs** [online], 12/2. Available from: <http://www.japanesestudies.org.uk/ejcs/vol12/iss2/jung.html> [Accessed 8 March 2015].
- Kim, M (2012). 'gasujimangsaeng 'yanghyeonseog' hoesa-e julseoneun i-yu,' **Asia Economy**, [Online], 3. Available from: <http://www.asiae.co.kr/news/view.htm?idxno=2012020217444408857> [Accessed: 16 November 2014].
- Kress, G. and van Leeuwen, T. (2006) **Reading Images: The Grammar of Visual Design**. 2nd ed. London: Routledge.

Machin, D. (2010) **Analysing Popular Music: image, sound, text**. London: Sage.

Moon, S.S. (2002) "The Production and Subversion of Hegemonic Masculinity: Reconfiguring Gender Hierarchy in Contemporary South Korea". In Kendall, L. (ed.) **Under Construction**. Honolulu: University of Hawai'i Press.

Owyong, Y.S.M. (2009) Clothing semiotics and the social construction of power relations, **Social Semiotics**, 19/2, 191-211.

Seabrook, J. (2012) Factory Girls: Cultural Technology and the making of K-pop [online]. October 2012 Issue, *Annals of Music*, The New Yorker. Available from: <http://www.newyorker.com/magazine/2012/10/08/factory-girls-2> [Accessed: March 7 2015]

Shin, S.I. and Kim, L. (2013) Organizing K-Pop: Emergence and Market Making of Large Korean Entertainment Houses, 1980-2010. **East Asia**, 30, 255-272.

Smith, B.A., Tan, S., Podlasov, A. and O'Halloran, K.L. (2011) Analysing multimodality in an interactive digital environment: software as a meta-semiotic tool. **Social Semiotics**, 21/3, 359-380.

Van Leeuwen, T. (1996) Moving English: the visual language of film. In Goodman, S. and Graddol, T. (eds.) **Redesigning English: new texts, new identities**. London and New York: Routledge. pp. 80-103.

Van Leeuwen, T. (1999) **Speech, Music, Sound**. London: Macmillan.