The Report Committee for Kit Yan To
Certifies that this is the approved version of the following report:

The Voice of the Future: Seeking Freedom of Expression Through
VOCALOID Fandom

APPROVED BY

SUPERVISING COMMITTEE:

Supervisor:

Stephen Slawek

Veit Erlmann

The Voice of the Future: Seeking Freedom of Expression Through VOCALOID Fandom

by

Kit Yan To, B.A.

Report

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Master of Music

The University of Texas at Austin

May 2014

Dedication

To my mother

Acknowledgements

Abstract

The Voice of the Future: Seeking Freedom of Expression Through VOCALOID Fandom

Kit Yan To, M.Music

The University of Texas at Austin, 2014

Supervisor: Stephen Slawek

Hatsune Miku is a celebrity; she is also a virtual singer with no real entity. The phenomenon of her success in Japan and abroad provides the starting point for this report, which examines the different forms of collaborative creativity that grow out of social energy arising from a collective interest in the dazzling Vocaloid characters and VOCALOID singing synthesis technology. From an outside observer's perspective, the feverish reception of this anime character may be found to be uncanny. How can this "unreal" thing possess such affective power? By attaching themselves to a non-real object, are Japanese otaku (nerds) exhibiting pathological tendencies? Or has their frequent exposure to anime and manga predisposed them to be emotionally receptive to virtual characters in way that neophytes lack the experience to understand? Taking a cue from Bruno Latour, this report confronts these questions by opening the malfunctional social black box of otaku group formation. I try to understand how the meaning of otaku is made stable through a social explanation, and why the VOCALOID fandom is distinct from ordinary people. In contrast to the technological determinism and socially

determined use of technology, I apply Actor-Network-Theory's theoretical and ethnomethodological perspectives to the VOCALOID community constituted from sociotechnical networks. Based on a position of ANT that each actor interacts with other actors (human and nonhuman) that constitute the network, this report looks at the particular media platforms and their infrastructures that allow the distribution and circulation of songs and videos. I then provide my ethnographic account based on a local VOCALOID event in Hong Kong in an attempt to understand how fans are recruited to the network, and what their motivations to collaborate, create, and share are. Making the connections evident, I conclude that the stereotypical social theory is somewhat unessential to our understanding of social relations.

Table of Contents

List of Figures

Chapter 1: Introduction

"Sixteen-year-old Hatsune Miku filled stadiums of adoring fans on a recent concert tour. Her albums have topped Japan's music charts. She's attracted more than 55,000 Facebook followers. Oh, and one other thing worth noting: She's totally fake. Since Miku was introduced to the world in 2007, more than 30,000 songs and movies about her have been created, and you can see many of them on YouTube and the Japanese video-sharing site Nico Nico Douga. Though Miku-mania started in Japan, it's now catching on in the U.S. Last fall, Miku made her first American appearance at the New People center in San Francisco's Japantown."

---"The Hologram Rock Star", *New York Times Upfront* (Jan 10, 2011)

"We can see her but we can't touch her. We think she is a true idol, the purest kind of idol," gushes 21-year-old literature student Keisuke Umeda, decked out in a Miku T-shirt and a range of her merchandise. "She's so cute, and she dances so well," says his girlfriend Azusa Fushimi, 20, a design student dressed in a self-made Miku outfit for the occasion, complete with body-length aquamarine pig-tails.

An excited roar rippled through the crowd when Miku burst on to the stage as a cloud of pixels that morphed into the shape of a petite galactic manga vixen with thigh-high boots and an impossibly short miniskirt."

---"Japan's 3-D Avatar a star in the Real World", *The Australian* (June 15, 2011)

These vignettes depict the unusual craze in the live concert tour of the pop diva Hatsune Miku in Japan and overseas. Hatsune Miku[1] is the rarest kind of celebrity. She is cute, popular and has had number one chart-topping singles in her native country of Japan. This newest, stylish pop diva, moreover, differs from most of her peers in one crucial aspect: she is a 3D hyperreal hologram anime idol, who will never age or die, but will forever remain a sixteen-year-old,[2] and her voice will never falter. In fact, she is the

---

[1] Hatsune Miku (初音ミク) is the name attached to a popular synthesizer program VOCALOID designedto a blue pigtails anime female character.   Hatsune (初音) literally means "first sound." Miku (ミク) means "future." According to Crypton Future Media, her name is meant to signify the "first sound from the future."

[2] Hatsune Miku is marketed as "A Singer in A Box" with basic biographical information: 16 years old, 158cm tall, weighs 42kg, favorite genre (J-pop, Dance pop), favorite tempo (70-150BPM), best voice range

world's most famous crowd-sourced Vocaloid (literally, vocal android), a product developed and sold by Japanese firm Crypton Future Media, Inc. using VOCALOID[3] singing voice synthesis technology by Yamaha Corporation that uses concatenative synthesis to splice and process vocal fragments extracted from human voice samples. In other words, with this technology, singing can be produced without a singer. The expression and audio quality of Vocaloids are improved to the point where it is hard to tell the difference between Vocaloids and human beings.

Unlike the robotic and unnatural synthesized voice created by most computer synthesizers, the naturalness, intelligibility and accuracy of her voice is celebrated in which her "human" persona is very much indebted to the vocal samples from the voice actress Saki Fujita, who gives a spirit, life-force to the character. The songs performed on stage by the 3D hologram are mostly written by her fans, not professional producers in a studio. Ever since her debut in 2007, Hatsune Miku has become the world's most famous crowd-sourced musician, one with her own fan club, Facebook page, and record label. Taking advantage of the obsessive tendencies of otaku culture (or the geeks obsessed with Japanese anime and manga), her turquoise, ankle-length pigtails, slim figure, big eyes, and cute gestures, perfectly embody the beloved stereotypical young girl anime character. She has also appeared in a number of commercials in Japan for big names such as the Toyota Corolla car and Google Chrome. Her music and image have proved so popular that she has now gone on a sell-out live tour in Japan and overseas. Projected

(A3-E5). The design and biography are the only official information provided by Crypton, her other traits are left to the users and fans.
[3] VOCALOID ™ (ボーカロイド) is a registered trademark of Yamaha Corporation, whereas Vocaloid refers to vocal android (i.e. the characters) in this report.

larger than life on a screen, Miku sings and dances according to the direction of the

programmers who "choreographed" the concert accompanied by live musicians.

Thousands of fans frenetically scream, wave glow sticks, and pump their arms in the air

for the entire performance as if she is a real-life pop star. This is the same treatment you

would see given to any other pop diva, only Miku is not human.

The stark unreality of Hatsune Miku does not seem to bother her fans at all.

Thanks to the advanced holographic technology, nothing is more exciting than bringing

their idol to real life on stage from a fan's perspective. More profoundly, her fans argue

that she is "real," the most perfect idol because she is crowd-sourced, independent of the

mainstream music industry. Her fans are empowered to compose songs for her using

VOCALOID software, project personality traits onto her with their imaginations, and

circulate fan arts that enhance her persona and life story through illustrations and

animated videos. What do fans mean by "real"? One fan's commentary exemplifies the

love for Hatsune Miku as the ideal idol who can unite the world:

> Miku is not just a Hologram, she is a *alive* person in many people's heart. Her
> songs has bring hope in people's life. Even though many says she is not real but I
> would say that she is *real* in my heart. Everyone is a creator and Miku proved it to
> us as many people ha made many wonderful songs using her. Miku brings hope
> and future in Music, she is our *new hope*. Miku should really sing for the
> Olympics because she can bring us one step closer, to love, to peace. (Anonymous
> comment from "the Top Ten List," Italics by the author).[4]

Similarly, many commentaries not only from the same website but also other

video-sharing platform such as YouTube share a similar insight, arguing that she is the

---

[4]The Top Ten List. 2012. "Music Artist that You'd Like to Perform at the London Olympics
Opening Ceremonies." http://www.thetoptens.com/singers-perform-london-olympics-opening-
ceremonies/hatsune-miku-391821.asp (Accessed January 21, 2014).

embodiment of many aspiring composers, artists and musicians who have come together to fulfill their dreams. The VOCALOID fandom is a participatory culture of innovation and interaction in which it is not only the creators but also the fans who have made her what she is today.

Nevertheless, to claim an anime character to be as real as a human being is by no means a normal reaction, and, reasonably, is hard to understand. This unusual frenetic reaction of VOCALOID fans perplexes the public (whom I consider as "outsiders" in this essay). Not surprisingly, some might argue that paying monies to see a fake hologram concert is ridiculous. Despite the fact that her natural voice, charismatic performance on stage, and technological advances intrigue many people at first glance, many people criticize both Miku, because of her fakeness and computer-synthesized voice, and her fandom, because of their pathological qualities that relate to otaku[5] culture. Others express mild reactions to the hologram as some sort of hi-tech, fancy dancing puppet that sings to computer-generated music with little emotion, which is nothing more than a computer program trying to match a human. There is no doubt that her hologram is made by fabulous technology. Like many others, when I first watched her performance and the fan reaction on YouTube I regarded it as totally weird. Why would you feel the need to create a virtual performer when there are plenty of talented "real" people whom you could discover and make famous? Why would people find an anime singing hologram more desirable than a human singer? Under what kinds of conditions would an anime

---

[5] Otaku, which is a Japanese term synonymous with geeks, has now become a general term referring to an obsessive fandom of subculture strongly linked to manga, anime, video games, anime figurines, and so on.

singing hologram become favored? So how might we understand this new form of

disembodied yet physically present performer?

When I was first exposed to the virtual singers and VOCALOID fandom through

its online videos and media news in 2010, I was very reluctant to appreciate the frenetic

fan reaction toward a fake character in which I considered the overreaction of fans as

merely a kind of deviant obsession. Male fans tend to be sexually aroused by a cute girl,

an essentially fake anime character.[6] My initial encounter of Hatsune Miku and

VOCALOID was an online statistical poll in 2011 titled "Music Artists You'd Like to

Perform at the 2012 London Olympics Opening Ceremonies" via a Facebook link in

which Hatsune Miku had topped the list.[7] By that time, I was surprised by the fact that an

unknown Japanese singer managed to beat down all the popular K-pop bands and

English-speaking celebrities. Although there was no evidence regarding the credibility of

the poll, the commentaries were sophisticatedly well written in support of their choices.

Out of curiosity, I searched her name and found many YouTube videos with millions of

hits featuring her live performances and animated videos, as well as her fan-made arts.

On the one hand, I was more than intrigued to realize that she was an anime character,

having an amazing voice and spectacular stage performance that was not inferior to

humans. On the other hand, having seen how the fans fetishized an enormous anime

hologram in front of the screen, my impression of the VOCALOID fandom became

---

[6] There has been news reporting that Japanese men prefer to go out a date with virtual girlfriends
than real women. http://www.bbc.co.uk/news/magazine-24614830 (Accessed on January, 29).

[7] http://www.the-top-tens.com/lists/singers-perform-london-olympics-opening-ceremonies.asp
(Accessed on January 29, 2013).

immediately associated with deviant psychological attributes of otaku. For people who resist anime subculture, VOCALOID fandom is often associated with otaku, which is now a general term referring to an obsessive fandom of subculture strongly linked to anime, video games, anime figurines, pathology of social phobias, escapism or schizophrenic disorder in both Japan and overseas. Although I have been exposed to Japanese anime and manga culture since my childhood, I have never indulged myself in manga/anime consumption. Hence, I regard myself as an outsider, uncertain about why people have been persistently attracted to anime, video games, and now, virtual idols, which lead the trend.

My bias and resistance to the VOCALOID subculture, nonetheless, sparked my preliminary interest in studying the interaction between this fascinating subject and humans from an anthropological approach. The core purpose of the discipline of anthropology is the study of the human condition. The fan reaction to a hologram concert challenges many conventions in a way that stimulates interesting inquiries and causes us to rethink the boundaries regarding what constitutes human beings, what it is to be human. Clear distinctions between what is real and what is virtual, what is normal and what is abnormal, fracture. Given this technologically mediated voice and body, how can we understand what the self is, what a human is?

Anthropologists have posted inquiries contesting the shifting boundaries between human and nonhuman entities, which grounds much of the posthuman discourse (Haraway 1991; Hayles 1999). Yet, even the posthuman discourses have contributed

provocative accounts to our understanding of the human condition. I argue that such

binaries fall easily into the power relations that language has granted our epistemology.

Must the 2D character necessarily fall into the dual boundaries of human-nonhuman? Can

we get rid of the duality when we talk about it? In relation to that, I consider the anime

character as an image, an "encounter sign," which is, itself, without specific content or

meaning but which, according to Gilles Deleuze, through its affective sense leads us to

think (Deleuze 1972; Bennet 2005: 7). This approach takes us beyond the interpretive

closure of what images say, as well as what the viewers see in them, into the shock of

imagery affective engagement.

Framing this research from an ethnomusicological perspective is a difficult

task. Ethnomusicologists who study world music culture have often tended to ignore

electronic-based technology in their studies of music making, identity formation, and

community building. For instance, they are analytically inclined to favor traditional,

authentic musical instruments over those that originate with western electronic

hardware[8], even when such technologies are used with immense creativity.

The technology of altered sound has opened up new directions for musical

expression, inspiring new logics of music creation and engaging more people in the

process. People around the world nowadays are merging technological synthesizers with

traditional "musicking" (Small 1998) in unpredictable ways, and they are producing a

---

[8] The Hornbostel-Sachs system of organology organizes the instruments into four basic units:
membranophone, idiophone, areophone, and chordophone. It sometimes also extends to include
electronphone (i.e. synthesizers), though the system fails to distinguish among the electric
musical practices.

wide array of musics that are understudied, and thus deserve our attention. My research, first and foremost, tackles the affective power of images foregrounded by the virtual singer Hatsune Miku and contemplates how the new singing synthesis technology VOCALOID pushes and inspires the way people create, experience and appreciate music in a participatory culture regardless of regional boundaries. Ethnomusicology remains a discipline open to posing new questions, embracing diverse methodologies, and breaking with tradition when required by the empirical evidence. As an ethnographer interested in conditions of music technology and the techno-human encounter, I want to approach the problem and understand the fandom using a methodology and perspective different from the mainstream musicological/ethnomusicological literature in dealing with the global phenomenon of the virtual singer.

Apart from mainstream ethnomusicological studies, first, this report does not specifically place the phenomenon of virtual singers in a regional context. This is not to flatten the entire world into a unified global understanding. Moreover, it is clear that nation-state divides (to say nothing of area studies regional divides) fail to describe technological global ramifications. While the discussion of virtual singers revolves around Japanese popular culture, given their global popularity on peer-to-peer video sharing platforms through the Internet, virtual singers, nevertheless, impact the totality of human consciousness regardless of regional boundaries.

Second, this report does not look at a particular musical genre. Given that millions of VOCALOID songs and videos widely circulate on video-sharing platform for free,

VOCALOID accommodates a vast diversity of genres including J-pop, rock and experimental music (songs that beyond human's vocal range) that cannot be generalized into a subgenre called Japanese popular music. I am less interested in discussing specific musical genres than in discussing the crowd-sourced musician, for it is the human and technological factors involved in the various VOCALOID activities, collectively creating her persona and background story, that I would like to grasp.

Fieldwork has remained an important and central feature of anthropology and ethnomusicology. For anthropologists, technology now has a tangible presence not only at the research stage but also "in the field." Ethnomusicologists have come to reflect on the state of fieldwork. Despite the fact that my research subject is intimately connected with online media, it is by no means merely an Internet-mediated research project surveying people's online behavior indirectly. In his "Rethinking Digital Anthropology," Tom Boellstorff redefines the concept of digital anthropology as an approach to "researching the virtual that permits addressing that object of study in its own terms, while keeping in focus how those terms always involve the direct and indirect ways online sociality points at the physical world and vice versa" (2013: 40).[9] In other words, digital anthropology studies the relationship between the virtual (the online) and the actual (the physical or offline). Arguably, this relation determines what we take the virtual to be, what we take knowledge about the virtual to entail and what we understand

---

[9] Anthropologist Boellstorff expresses his discontents with the blurring notion of digital anthropology and virtual anthropology. He argues that digital should not be treated as an object of study but as a methodological approach (essentially founded in participant observation) for investigating the virtual and its relationship to the actual (40).

as the stakes of the virtual. His insights allow me not only to survey my informants'
online VOCALOID activities but also to learn about and take their life outside of the
Internet into account.

An inspirational and productive ethnographic approach to deal with complex
techno-human interactions is Bruno Latour's Actor-Network Theory (ANT), which
provides an analytical model to address the relationship between humans, science and
technology in a situation where the "social" is not easily defined. Latour's *Reassembling
the Social: an Introduction to Actor-Network Theory* (2005) rejects taking the word
"society" as a given; rather, he views society as something that is assembled by its actors.
Latour criticizes the word "social" for its implied confusing cause and effect facet that
forecloses the process of tracing the social connections and formations that it seeks to
explain. He highlights how human and nonhuman actors have to be considered equally
important in the constitution of networks. Society is not a force that determines the
actions of its actors; but, rather, it is a force that is constituted through and by the actors'
interactions that determine the form of the social.

Following Latour, I do not mean to provide a macro-scale "social explanation" of
how VOCALOID fulfills social needs, rather, I attempt to trace connections between the
human (e.g. user-creators, listeners) and nonhuman (e.g. Vocaloids characters, video-
sharing websites) actors. Since the success of Hatsune Miku is frequently attributed to
video-sharing websites, in particular to the Japanese-based YouTube competitor *Nico
Nico* Douga, Actor-Network Theory offers a vantage point for an analysis of participation

10

in virtual worlds and the enabling power of non-humans in contributing to collective life.

Before I delve into the network construction within the VOCALOID community, I would like to question the social explanation regarding the VOCALOID fandom as well as the anime and manga fandom as a whole. As noted earlier, due to the 2D anime characterization that comes with the VOCALOID synthesizer, otaku is a common term for people who resist anime subculture to label fans who are interested in the works and activities that link to the dazzling imagery of anime and manga characters. Contrastingly, I do not aim at criticizing the prior representations of a social group called otaku or adding another definition of what an otaku really is. Instead, I want to question the existence of such a homogenous, social group as a social phenomenon in the first place, especially if the term is so vague, yet powerful and imagined globally, particularly in East Asia.[10] Tellingly, media have today widely propagated the word otaku as a shorthand label to represent and categorize a group of young people that is identified with the social youth problem. But I have come to wonder: do people's personal attachments to 2D characters necessarily mean a rejection of society? One radical problem with this ubiquitous term is the implicit certainty that otaku represents the anime and manga fandom, and the public has taken its negative image as a given.

In this sense, this social discourse is reminiscent of Bruno Latour's concept of

---

[10] Since Japanese anime and manga is pervasive particularly in East Asia (e.g. China, Hong Kong, Taiwan), the term has been translated into Chinese character 御宅族. But it is widely used to identify with social youth problem in East Asia. However, contrastingly, English speakers do not identify otaku as pathological but conceive otaku as something "cool" under the rhetoric of "Cool Japan."

"black box." In Latour's account, cyberneticians use the term black box whenever something, such as a piece of machinery, is too complex to be represented or explained entirely. In its illustration as a diagram, it is drawn as a black box with only its input and output, but not its inner workings (1987: 3). Black box, according to its original usage, is a technical term in science and technology studies for a device, system or object when it is viewed in terms of its input, output and transfer characteristics without any prior knowledge required of its internal working. Along the same vein, the term "social" is frequently deployed as a shorthand from which something is made, and is often deployed as the explanation of transformations or forces between people (Latour 2005: 3-4). Social forces transform an input "X" to produce output "Y." I was intrigued with how Latour introduces the idea of black box with reference to Pandora's box from classical Greek Mythology. When Pandora opened the box, the evils therein escaped and spread over the world. The black box that Latour refers to connotes something mysterious and evil. In other words, believing a fact or claim without examination makes an idea more of a black box, whereas by opening it we can reexamine the bias and see what really happens inside. Reopening a black box implicates a destabilizing effect, because the fact is no longer taken as a given.

My inspiration for applying the black box in the VOCALOID study is the complex network and bizarre mixture it entails. As I have mentioned, since VOCALOID is closely perceived as an activity that is attractive to an otaku due to the cute anime girl characters, I am very interested in how the meaning of otaku is made stable, and why the

12

anime and manga fandom, as well as VOCALOID fandom, is distinct from "ordinary" people. The "malfunctional" social boxes of VOCALOID are urged to open up and make their connections bare. Therefore, my starting point in this report is to look at the actors that make up the "black box" of otaku's negative image and the following established practices that extend the governing logic (Foucault 1972). Similarly to a software program or pieces of lab equipment, I seek out the actors involved in transforming the input to the output "otaku"? Chapter 2 explores three actors that have contributed to otaku as a social phenomenon. Here I do not mean to delimit nor attribute the formation of the group to the three actors, but rather to consider the different interactions and actors within a network in establishing and populating a real, homogenous, solid social group called otaku.

In opposition to the social explanation of VOCALOID fandom and the group formation of otaku, Chapter 3 examines the roles and characteristics of nonhuman actors that emerge in time in relation to other human actors that constitute a sociotechnical network around Vocaloid characters. In this chapter I look at the affordances and constraints granted by two nonhuman actors, the VOCALOID singing synthesizer and Niconico video-sharing websites, in constituting the actor-network connections in VOCALOID song productions and video circulations. Chapter 4 draws on my ethnographic work in Hong Kong during the summer of 2013, and from ongoing ethnographic incursions into the Internet in order to examine the multiple reasons that VOCALOID music producers and fans identify with this subculture, and traces the

integral role and impact it plays in their lives. What new forms of collaboration are being enacted by people involved in distributing VOCALOID songs or fan arts? How are groups formed and sustained over time? How are new collaborators recruited to the group and non-contributing collaborators dropped from the group? What effects does this seem to have on the project itself? By tracing the connections in which the user-creators become motivated to share, collaborate and create, I argue that their motivations can be understood by looking at how they are recruited to the network.

Chapter 2: Opening the Black Box of the Otaku

VOCALOID is a popular sample-based vocal synthesizer developed by a Yamaha research team in 2000 and best known for its natural, realistic singing synthesis. It provides amateur user-creators an integrated condition in which the user can generate a singing voice easily and use it for music production. As often noted, it is the power of 2D anime characterization that has made this technology globally popular, integrating a technical software into a bundle of *moe*-inducing character elements. [11] An example of the *moe* elements of the Vocaloid characters is Hatsune Miku's distinctive turquoise, ankle-length pigtails, and matching black and turquoise outfits, which makes her character a remarkable success that people have a strong impression of her appearance without learning the fact that she is a "singer."

The idea of employing anime characters as commodities is not especially new in Japan. Yet, for this reason, VOCALOID fans are often labeled as otaku,[12] which is commonly perceived as a socially deviant "geek," having pathological problems that lead

---

[11] *Moe* (萌え "moh-ay"): A loose term for affectionate pseudo-desire for inanimate or virtual 2D characters, or more accurately, a reference to an internal response to something with no hope for a reciprocal emotional payback (Condry 2013: 187). In his analysis of otaku culture, Azuma (2009) argues that *moe* symbolizes a postmodern, "database" form of consumption. *Moe* is used to describe specific consumable items, examples of which would include images of desired body fragments originated from manga. Cat ears, cat tail, bells in Doraemon style, green hair, maid uniform, big and loose socks on high school girls, hair sticking up like antennae are examples of otaku's favorite *moe* elements (Azuma 2009: 42-43).

[12] As I have mentioned in the introduction, my observation of the perception of Vocaloid is that outsiders often associate Vocaloid with otaku culture because of the predominant assumption of the character in relation to anime and manga popular culture.

them to pursue meaning in pseudo love or sexual relationships with 2D anime and manga

characters.[13] Whereas the erotic, infantile, girl characters are often found in adult manga,

the frenetic reactions of male fans toward Hatsune Miku in the hologram concert is

controversial and makes one wonder why people express love for an anime character and

treat her like a human idol. Accordingly, the anime and manga fandom, as well as

VOCALOID, are associated with the concept of otaku. More profoundly, the idea of

otaku is exemplified through comics and anime consumption in a world whose presence

is equated with today's Japanese popular culture in general (Azuma 2009). Since Hong

Kong and Taiwan have close connections with Japan, what is at stake is that the concept

of otaku has been appropriated by mass media in those places and has become a

stereotype of the social youth problems in these areas for those fail to communicate with

others. Despite the fact that the pejorative connotation of otaku has decreased in Japan, it

is observed that the local reception of the term is a direct adaptation without much

modification of its initial, pathological connotation, but it is continually exploited with

deformation by local media. Using Hong Kong as the main reference, according to

outsiders, the association of Vocaloid characters with otaku culture brings about the

public's misunderstanding about VOCALOID. As noted by Hong Kong scholars Yiu and

Chan, the radical problem is that according to the local perception of otaku, otaku are

considered perverse and despised as lower-class individuals because their behaviors lead

---

[13] In Japan, the relationship between fans and their beloved anime characters took "an unusual turn" in
2008 (Condry: 186). A Japanese otaku by the name of Taichi Takashita has set up an online petition to call
for the enactment of Japanese law allowing citizens to marry a 2D anime character. He states, "I am no
longer interested in the three-dimensional world. I would like to become a resident of the two-dimensional
world." More than three thousands people had signed on the petition within two months, expressing their
love and support for this project.

to their social ostracism (2013: 853). Any anime and manga related activities are perceived as and considered as perverse.

Otaku is never easy to articulate; it is easy to know what otaku is in colloquial conversation but never easy to comprehend the meaning of otaku because of the multiple complex layers related to social youth problems. It is a vexing embodiment of many key problems not only in Japan but also that the world in general faces today. Its supposed hedonistic, immature, fantasies obsessed, antisocial, escapist youth-driven stereotype is one of the reasons for the public's attention. Moreover, there was a burgeoning industry of otaku studies by the late 1990s. Sociologists are keen on giving social explanations of this emerging otaku social dispute in an attempt to claim that the pre-existing understanding of otaku (i.e. dysfunctional, asocial, obsessive behaviors) is a false representation that has to be remedied (Grassmuck 1990; Gibson 2001; Azuma 2009; Ito 2005; Taneska 2009).[14] What is more, in Japan, recent studies of the perception of the term suggest that negative connotations have significantly decreased, in part because of the emergence of a new category called *hikikomori* (youth shut-ins) to which the socially pathological discourses of youth have shifted (Kikuchi 2008).

Considering the arbitrary, fragile, controversial and unstable states in the group

---

[14] Scholars and critics have given various definitions of otaku. William Gibson (2001) calls otaku "passionate obsessive." Volker Grassmuck (1990) describes otaku as "information fetishist." Lawrence Eng (2002) suggests the otaku spirit as one of "exploration, innovation, curiosity, dedication, and individualism." Hiroki Azuma (2009) situates otaku in a stabilized viewing position in front of the database. Mizuko Ito (2005) describes otaku as media connoisseurs. Taneska (2009) adds on the definition of otaku as the ultimate representative of information, consumption, resumption and re-creation.

formation of otaku, it is what Latour calls the first source of uncertainty in Actor-Network-Theory in which the formation of groups is meaningless in that it accommodates "neither the size nor the content" (Latour 2005: 28-29). And yet, sociologists are accustomed to define and stabilize groups based on arbitrary information in making up a society. Over the decades, the concept of a social group called otaku is set up and stabilized by various agencies in relation to manga and anime fandom. While no one has ever been able to define what an otaku exactly is, the word has been extensively accepted as a stereotype in daily colloquial conversation. It has become easy to use the term to label a group of people associated with manga and anime materials but difficult to talk about it.

Actors:
1. Media
2. Psychoanalysts
3. Sociologists
4. ...
5. ...

Input:
Otaku

Output:
Representations of
Otaku

Figure 2.1: the Black box of making up the otaku.

Latour's concept of black box is an inspirational approach to reveal some of the actors who have granted so much power to the social explanation of a group called otaku. The group formation of otaku is best understood by unpacking the various black boxes

18

and localities networked together, while also mapping the controversies about the otaku group formation. First, how did the word otaku emerge in the early 1980s, leading to a "moral panic"(Cohen 2002) during the controversy of a murder case in 1989 in Japan? Although the Miyazaki incident is not known outside Japan, I argue that the incident cemented the caustic link between otaku and its pathological association with 2D anime characters. The perception of otaku has undergone several changes over twenty years in Japan, but the arrest of Miyazaki marks the time when the word otaku was brought into discourse in reference to a social pathology and was popularized through the mass media. The stereotype of otaku was made stable and further consolidated in overseas mass media. The obsessive consumption of manga and anime was then considered abnormal behavior.

Second, I turn to a review of the psychoanalysis of the otaku male gaze found in English-language scholarships on manga and anime culture. The connection of the 2D anime characters with the genre of anime and manga pornography is another factor that contributes to the creation of the black box of the otaku negative image. Given the predominant stereotype of otaku as it evolved in Japan, the increased global anxieties over the link between anime and child pornography have put emphasize on the image of the otaku as a sexually promiscuous and potentially criminal group. The *moe* elements, especially those indecent images drawn from *ero* manga and anime, have been extensively discussed in English-language scholarships. In light of the psychoanalytic approach, English-language scholarship has often noted that the male gaze indicated in

the visual *moe* elements of female anime characters is akin to perverse sexual behavior

(Allison 2000: 30). Although the scholarly studies do not emphasize the deviant behavior

in viewing the erotic anime characters, I argue that, essentially, the psychoanalytic

approach conveys the deviant behavior of otaku. In fact, scholarship becomes the actor

that categorizes and stabilizes the anime and manga fandom as a social group whose

behavior is different from ordinary people.

Third, I then turn to look at a seminal work on otaku by the Japanese philosopher

and critic Hiroki Azuma, an actor who helped to legitimize and promote the otaku culture

within subcultural studies at the turn of twentieth-first century. First published in

Japanese in 2001, *Dobutsuka suru posutomodan: otaku kara mita Nihon shakai* (literally

as "Animalizing Postmodernity: Japanese Society Seen through Otaku") became a

bestseller and has gained wide public attention in Japan due to Azuma's ability to

commingle philosophical theory with the Japanese popular culture (i.e. otaku subculture)

in an accessible style and readable voice. Its mass reception in Japan brought about the

publication of Korean, French, and English translations of the book in 2009. While most

of the Japanese scholarship on otaku studies is written in Japanese and not available for

non-Japanese speakers, Azuma's book in English translation has become an important

source to analyze otaku culture from Japanese's perspective. Since this book is widely

received at local and global levels, I am interested in looking at his argument and his

underlying theoretical assumption in his analysis on otaku culture. Counter to the

psychoanalytic approach, Azuma makes information technologies central to the analysis

of otaku activities. He sees otaku as consumers of a new "database culture," which he links to a new mode of cultural reception dubbed "animalization" (2009: xv). However, I argue that his analysis of otaku culture in relation to media consumption appears to embrace technological determinism, a logic that implicates the cause and effect: media content dictates human behavior. Given the group formation of otaku that is a fundamental problem per se, I conclude that neither psychoanalysis nor database cultural analysis give an accurate account of what otaku do in daily life. No doubt that the otaku or VOCALOID black box is a complex phenomenon, the separation between technology and human elements is difficult and implausible. For this reason, I consider Bruno Latour's Actor-Network-Theory as an inspiring approach to examine the fact production in a complex network. In this Chapter, I will first open the black box, which makes up the narrative of otaku, and follow the key actors at the times and at the places where the bias and controversies were produced and circulated.

To start the enquiry, the first trajectory is how the negative image of otaku evolved and circulated in Japan in particular when a causative link between otaku and the murder case was drawn. The word "otaku" originally comes from Japanese, and its denotative meaning is "your home," but the primal use is as an honorific second person pronoun (Eng 2001). Moreover, it was in the 1980s that the word otaku started to describe and stereotype fans of anime, manga, video games and technology in its modern colloquial usage.

In an important sense, media discrimination against otaku began even as early as

since the word first circulated among the insiders. In 1983, critic Akio Nakamori first

referenced the fans of manga maniacs as "otaku" in his article titled "Otaku' Research" in

Manga Burikko examining the criticism of fans at comic conventions. The magazine

readers reacted fiercely against Nakamori's stereotypes of anime and manga fans. This

article stimulated a debate that lasted until 1984 (Okada 1996: 9). Consequently, the word

otaku started to circulate among manga artists and fans, and as a witty reference to

themselves. It also encouraged the readers of *Manga Burikko* to imagine a shared identity

in this fanzine community, though most people did not use otaku to refer to themselves

(Galbraith 2010: 216). Because the article was published in a minor fanzine catering to

manga fans, otaku, as a referent to a social group, still remained unknown beyond this

limited amount of people (Kinsella 2000:128). As the otaku narrative was present only in

a limited published fanzine running in a niche-market, it is necessary to ask how the term

emerged from the relatively insular world of otaku amateurs, eventually to permeate

public consciousness.

There are several key agencies that helped to establish the otaku image after 1989.

Mass media, not surprisingly, was the first key actor in disseminating the stereotype of

otaku. In 1989, the term otaku entered public discourse through the mass media reporting

of the notorious Tsutomu Miyazaki murderer case, in which the 26-year-old serial killer

Miyazaki was convicted of kidnapping and killing four little girls aged between four and

seven, mutilating their bodies and sexually molesting their corpses. After the arrest of

Miyazaki in 1989, video footage and pictures of his victims were found. His room was

reported to be stacked with collections of pornographic anime and slasher films, and then

the media quickly labeled him as the "Otaku Murderer"—instantly creating the

stereotype of the otaku as an asocial, deviant male youth who shuts himself up in his

room with anime and manga, thereby exposing the public to the term for the first time.

Due to the low crime rate, especially against children, in Japan, the incident of Miyazaki

shocked the public and fueled a moral panic against otaku (Toivonen and Imoto 2012:

13). In one of the early and influential sociological depictions of the otaku phenomenon,

Kinsella (2000) identified the otaku moral panic that ensued from the Miyazaki case. As

she notes, 'the sense that this unsociable otaku generation was multiplying and

threatening to take over the whole of society was strong" (1998: 311). "Otaku" was thus

recognized not only as a personality problem but also as a potential hazard to society.

No doubt that the media and journalists came to become the key actors in

spreading the public stereotypes and fears about otaku, accusing anime and manga of

making Miyazaki a serial killer. Considering that Japan was still in the thrall of its bubble

economy, a moment when money and pleasure seemed in endless supply, the mass media

went into frenzy over the arrest of Miyazaki, trying to explain what went wrong in

society. Starting July 1988 until the arrest of Miyazaki in June 1989, the mass media in

Japan continually reported the serial killer. Kamm (2013) opens the black box by

studying the connection between otaku and mass media. It was noted that *Asashi

Shimbun and Yomiuri Shimbun,* the two Japanese mainstream newspapers, had published

41 articles regarding the homicide in 1988. Moreover, the media coverage increased

23

rapidly from December 1988, a day after the body of the third missing girl was found.

Accounting for the number of articles that drew the attention to the murder, the remainder

of December observed 53 Asashi and 19 Yomiuri articles concerning the case. The public

was informed about the incident by journalists in details, including postcards, cardboard

boxes with body parts of the victim, the photos of the dead girls and the confession notes

sent by the killer to the victims' families over the past year. Five months later, the

public's interest culminated when Miyazaki was arrested. The mass media relentlessly

continued reporting on the victims' families and then shifted the attention to Miyazaki's

family, as media and the public tried to understand the motivation of Miyazaki for

committing a horrible crime. The individuality of Miyazaki, such as his hobbies and

interests, drew the media's interests. Since a large number of manga or anime related

materials and pornographic videotapes were found in Miyazaki's room, the pictures of

his room became a part of the evidence to explain and reflect Miyazaki's behavior and

mental state. Accordingly, the journalists soon relieved the public's suspicion by labeling

Miyazaki as an otaku: a lonely and psychologically disturbed young man who is a

frenetic collector of anime and manga pornography. Amateur fans became "the targets of

a nations-wide panic" about the murderous anti-social effects of involvement with violent

and erotic manga and animation (Kinsella 2000: 11).

However, there are disputes about the integrity of mass media regarding the case.

Eiji Otsuka, who is a manga critic, argues that Miyazaki's collection of pornography was

added by the photographer or police in order to highlight his perversity (Kamiya 2004).

Fumiya Ichihashi (2001), another critic, also suspected the released information was meant to culminate public stereotypes and fear against otaku as well as helping the police to cement a conviction.[15] Nonetheless, the riddle of the deviant crime had been resolved, at least according to the publicly known as the otaku narrative. Ever since then, the word otaku has become a dense node and a shorthand that stands for and conceals the networks that contributed to a group's existence and circulation. Otaku is meant to be pathological and dangerous.

Like the media commentators, academics are another key actor among those who are eager to explain what was wrong in the society. Both Japanese and overseas scholars from interdisciplinary backgrounds are interested in studying otaku. In the early 1990s, shortly after the Miyazaki incident, there was an increasing number of studies of otaku as a social phenomenon. Japanese social scientists quickly joined the media and worked with their statistics to define the group. In an empirical research project on university students, otaku were perceived as "unbalanced specialists" with low interpersonal skills, who tend to be avoided by others because of their unclean appearance (Miyadai 1994). In her book *Communication fuzen shokogun* (Dysfunctional Communication Syndrome), critic Azusa Nakajima stereotypes otaku as an anime/manga maniac who has asocial problems in which their pleasure can only be found in a relationship with objects other than with people (Nakajima 1991). Sharon Kinsella remarks that otaku came to symbolize the downfall of Japanese society itself in the recessionary 1990s. The early

---

[15] The news reporting the number of videotapes in his collection increased from 2.000 to 8.000 and later settled at 5.763 or 5.787. These discrepancies were reason to believe in evidence that there might have fabrication of data through the police or the photographers (Ichihashi 2001).

local research on otaku not only tried to explain otaku as a new youth problem but to draw a connection between otaku and males as the emergence and deficiencies of otaku signified the downfall of male status in the patriarchal Japanese society.

Indeed, anime and manga fans expressed their discontent toward the negative connotation of otaku promulgated by mass media and sociologists. By the mid-1990s, a number of individual agencies including artists and critics were involved in an attempt to change the perception of otaku. Intriguingly, the transformation and subtle shifts of otaku, previously perceived as deviant, is now perceived to be at the heart of Japanese cultural capital and identity. The anime producer Toshio Okada, who produced the influential TV anime series Neon Genesis Evangelion in the mid-1990s, was a key figure in the effort to change the perception of otaku. Referring to himself as the *Otaking* (King of otaku), Okada began to write prolifically and lecture at a university on otaku in the mid-19990s. The mission of Okada was to "eradicate discrimination against otaku." Accordingly, one of his strategies to elevate the otaku culture was to distribute Japanese anime abroad, where otaku culture was being consumed and identified as "cool Japan," thus elevating the otaku culture in Japan. It was so successful that 'Cool Japan' became a hot topic in Japan and grasped the attention of politicians after the circulation of McGray's influential article 'Japan's Gross National Cool' (2002), which noted the potential 'soft power' of Japanese popular culture under the rhetoric of 'Cool Japan.' Therefore, in Japan the concept of otaku gradually detached from its original stigmatized meanings to "innovators" and "heirs" of Japanese popular culture. Certain changes in

Japanese perceptions of otaku could be seen in the 2000s, when Akihabara became a

tourist spot for foreign anime fans and the TV drama featured an otaku as a hero.[16] The

meaning of otaku ramified and became increasingly appropriated and manipulated by

different parties.

The negative connotations of otaku have significantly decreased in Japan in recent

years due to different individual actors' efforts to change the public's fears against otaku,

whereas otaku has been stabilized as a social group that stands for anime and manga

fandom. While international audiences consume manga and anime in a context where

there is no historical association of manga and anime fans with the Miyazaki incident,

otaku is found to be a word roughly equivalent to another group called nerd or geek in

English with obsessive interests in anime, manga, and video games. The internationalized

concept of otaku remains open for translation and interpretation.

English-language scholarships noted the pervasiveness and soft power of

Japanese anime and manga culture. Beginning around 2000 and up until today, we see a

growing body of English-language scholarship centered on the study of anime and

manga.[17] In addition to the study of anime and manga as a diplomatic tool in "Cool

Japan," the internationalized concept of otaku was detached from its negative

connotations. Overseas scholarship is seemingly interested in the globalization of anime

---

[16] *Densha Otako* (Train man) is a 2005 Japanese TV drama, which tells a true story of 23-year-old otaku who intervened when a man started to harass a woman on train. He then begins to date the woman in the movie. The otaku protagonist is portrayed as a nice guy who is too shy to find a girlfriend.

[17] The pioneering work of scholars such as Anne Allison (2000, 2006), Ian Condry (2009), Sharon Kinsella (1998, 2000), Susan Napier (2000, 2007), and Joseph Tobin (2004) has helped established the study of manga and anime as a specialty.

and the affective power of the 2D anime characters. As Japanese anime and manga are often noted for their erotic and violent content that is considered perverse, scholars with Japanese literature and East Asian studies backgrounds are eager to study the sexual fetishism and the projection of sexual desire of cute girl characters from the perspective of psychoanalysis.

It is perceived that having a pseudo love or sex relationship with virtual characters and figurines is psychologically dysfunctional. The otaku fetishism of virtual characters is hard to articulate. In general, sex and homosexuality are the most noticeable topics in manga and anime where the images prone to fantasies of fragmented bodies identify female body parts, especially the fetishized breasts, which are imagined, scripted and represented by rage and violence in Japanese *ero* manga. *Hentai*[18] is the Japanese word that denotes a type of perverse or bizarre sexual desire or act of erotic anime characters. In particular, the sexual, *moe* appeal of anime characters and the otaku's sexual fetishism are widely recognized in popular discourse. In an important sense, Freudian and Lacanian psychoanalysis and the Marxian logic of commodity fetishism contribute to the popular analysis of the consumption of erotic desire (Allison 2000: 26-27). In light of Mulvey's psychoanalytic film theory (1975), the commodity fetish of the female body and figure situate otaku in the context of the so-called male gaze. Psychoanalysis has delved into the fundamental lack of male desire. For instance, female body fetishes in manga, anime,

---

[18] *Hentai* (変態 or へんたい) is a Japanese word short for perverse sexual desire. The term is used in theJapanese and English language sbut in a different context. In Japanese, the term refers to sexual material, but only of an extreme, abnormal or perverse kind. However, the English use of *hentai* is a generic term referring to erotic or sexual manga and anime in general (McLelland 2005).

video games, and figurines are the themes that circulate endlessly in the otaku's attention. Gender and sexuality are the predominant areas to be explored in explaining otaku's deviant behavior and male sexual desire. Similar to Nakajima's study, anime scholar Anne Allison defines the otaku male gaze in *ero* manga,[19] which is a "masturbatory substitute for 'real sex' when the alternative of actualization is not currently available" (2000: 30). The psychoanalysis of the male gaze situates otaku in a consumption of desire. Along the same vein, Susan Napier (2006) analyzes the male gaze in a number of anime cinematic productions. Through the consumption of virtual commodity one can displace drives or biological needs. Consequently, academics establish a narrative of otaku as a kind of endless extension of desires, a complex structure wherein desires are instilled with the desires of others.

The controversy is that the effect of the research is that of readily admitting that anime and manga fans have developed psychological problems, thus helping to construct the negative image of otaku as perverse and pathological; that is, one cannot help but desire into infinite. In other words, the involuntary desire of virtual characters discriminates manga and anime fans from ordinary people by labeling them as otaku: someone fails to develop normal human relations but now finds it easier to develop pseudo-relationships with virtual characters. Hence, the premise of the psychoanalytic approach to understanding otaku is problematic as it first situates the subject matter as sexual, abnormal, and deviant.

---

[19] *Ero* manga is a sexual manga where the imagery of female naked body is depicted and fetishized as the most dominant feature. It is meant to elicit the voyeuristic male gaze (ibid, 30-31).

In contrast to the sexuality analysis of psychoanalysis, Hiroki Azuma, the Japanese philosopher and media commentator, makes information technologies central to the analysis of otaku activities. Similar to the psychoanalytic approach, Azuma also shows interest in the analysis of the *moe* elements present in anime characters. However, he centers on consumer behaviors and media rather than on desire and sexuality. Like Toshido Okada, Azuma is an important figure who challenges the negative image of otaku and helps to legitimize the otaku culture in 2000. It is in his book *Otaku: Japan's Database Animal* that he persuasively argues that otaku culture presents the emergence of a new "database structure," which he links to a new mode of cultural reception dubbed "animalization." Otaku is the innovator, "posthistorical" man that is detached from the grand narrative (2009: 67).

Azuma further identifies two animalistic characteristics of the database consumption in otaku culture in terms of the postmodern condition: the propagation of simulacra and the decline of the grand narrative. First, he considers the high value placed on the omnipresence and proliferation of derivative works in comic markets and over the Internet (2009: 25). This is extremely close to the future of the culture industry as envisioned by French sociologist Jean Baudrillard. In the past, the original work was "an original" and the derivative work "a copy," degraded with no doubt. Moreover, due to technological advancements, Baudrillard (1994) predicted that the distinction between original products and commodities and their copies would weaken. On the other hand, an interim form called the simulacrum, which is neither original nor copy, becomes

dominant. He posits that in postmodernity the indefinite reproducible images of the media precede the real. The images no longer represent the reality but refer to each other and produce the hyperreality (1994: 2-3). In this regard, Azuma has described, "it is quite ambiguous what the original is or who the original author is, and the consumers rarely become aware of the author or the original. For them, the distinction between the original and the spin-off products does not exist" (2009: 39).

In otaku culture, simulacra are the amateurs' derivative works of original manga and anime largely circulated in the form of fanzines and peer-to-peer sharing platforms. The distinction between the original and the derivative works does not exist. The only valid distinction for them, in Azuma's double-layer analysis, is between the settings created anonymously (a database at a deep invisible inner layer) and the individual works that each artist has concretized from the information (a simulacrum on the surface visible outer layer), as a result of the so-called "database consumption" of moe-elements (2009: 39).

In sum, database consumption is a term Azuma coined for which otaku today construct databases toward the priority of their favorite *moe* elements of character over narrative story. An otaku no longer feels a need for large narrative worlds but prefers to construct small narratives by disassembling and reassembling the characters. As a result, Azuma sees *moe* as a behavioral response that makes otaku prey to technologies. He argues that otaku choose fiction over social reality not because they cannot distinguish between them but rather as a result of having considered which is more effective for their

human relations or virtual fiction (2009: 27). The controversy is that as if the consumption of *moe* entails a complete collapse of perceptual distance between the user and the computer. Although he emphasizes the fact that otaku is making rational choices between reality and virtual fiction, the underlying assumption is that otaku activities entailed a behavioral materialization of computer processes that digitalize everything, transforming the world into equalized bytes of information. The animalization of database (*moe*) consumption has affective power in that the simlation governs the user's animalized response. As a consequence, the computer controls the user rather than vice versa, thuswhich tending toward technological determinism.

This chapter provides a starting point by introducing the various actors that constitute the black box of the negative images of otaku. I described three actors that make up the narrative of otaku and otaku culture, namely, mass media in the Miyazaki incident, the psychoanalytic scholarships on the otaku male desire, and Hiroki Azuma's technological deterministic analysis on otaku culture. Given the newspapers, journal articles, and books produced by the actors, the documents serve as actant to close the black box of otaku by taking the fact of the dangerous group as a resource and moving on the disputes. This is the process by which the disputes become closed and ideas accepted, that is, with how the facts are made about what is known.

As demonstrated above, the medium of anime itself—with its dazzling imagery and lucid playfulness—is a complex subject that is made up of the controversies and assertions by outsiders. In dealing with this, the following chapter starts with the tenets of

Actor-Network-Theory. ANT distinguishes itself from other social theories by

considering both human and nonhuman elements equally as actors within a network in

order to overcome the unnecessary duality between humans and nonhumans. As

technologies are often overlooked in social theories, I turn to identify the technological

nonhuman actors and reveal their affordances and constraints that constitute the actor-

network connections in the VOCALOID community.

Chapter 3: Reassembling the VOCALOID Network Connections

Having opened the social black box of VOCALOID and questioned the social

construction of the homogenous social group otaku, this chapter centers on the Actor-

Network-Theory (ANT) and its application to VOCALOID activities within the online-

sharing and local networks. My VOCALOID study consists of a complex array of

techno-human interactions with the technologically mediated voice and anime characters,

digital singing synthesis technology and "Web 2.0"[20] media technology. Multiple

networks of human and technological components constitute and connect each other in

VOCALOID activities, leading to very complex relationships among me as an analyst,

creator-users, viewers, and multiple types of technology.

Considering the implication of technology's impact on human behaviors as well

as the socially determined use of technology, the long-held binary of social determinism

and technological determinism seems not appropriate to study this complex subject since

the separation between technology and human elements is difficult and implausible. For

this reason, I adopt the socio-technical account of Actor-Network-Theory as an inspiring

approach to treat people, technologies, Vocaloid characters, media, and music all on

equal terms in the constructions of VOCALOID actor-network activities. Hence, priority

is not given to either a technologically based or socially based explanation. ANT theorists

consciously use the term "sociotechnical network" or "heterogeneous network" in order

---

[20]A "Web 2.0" site refers to World Wide Web sites as a collective medium that allows users to interact and collaborate with each other in a virtual community in contrast to the earlier Web pages that people are passive viewers of contents. Social networking sites, video-sharing sites, Wikipedia, blogs are the examples of Web 2.0 sites.

to overcome the unnecessary duality between humans and nonhumans. The important fact here is not only to recognize the significance of both humans and nonhumans, but also that they are defined relationally in the network and function likewise. ANT insists that "the stability and form of artifacts should be seen as a function of the interaction of heterogeneous elements as these are shaped and assimilated into a network" (Law 1990: 113). This leads to a relational epistemology that rejects the positivist views of objects and actors as existing in themselves prior to any participation in the networks of associations. The position of ANT allows me to examine the effects and human interactions of heterogeneous elements such as musical and Internet technologies, shaping the inter-personal relationships and the manner in which the VOCALOID social connections are built.

In this chapter, I hope to bring the reader into the fans' world to examine the network connections in making VOCALOID music and other derivative works. According to ANT, culture and context emerge from complex networks and interactions rather than stabilizing groups and defining activities. Specifically focusing on VOCALOID music videos circulated on video-sharing websites, I first identify the actor-network in making VOCALOID original songs and videos, which in turns leads to a chain of covers and creative works. I use ANT to show the fans' "way of knowing" through online activities. The circulation of songs through the internet-based network enables fans to inspire and collaborate with each other. In particular, the examples I draw upon illustrate the several human and nonhuman actants, such as the song and the

Vocaloid characters, in contributing to the song production and circulation through shared online practice. An actor-network approach begins with opening the several black boxes of technologies, allowing the analyst to reveal network connections.

The Actor-Network Connections in VOCALOID Song Productions and Video Circulations

The circulation of VOCALOID videos on video-sharing sites has become a vital source for fans to acquire the knowledge of the voice synthesis technology and music. Thanks to the Web 2.0 media technology evolved in the late 1990s, videos are visible and easily accessible for users with a few clicks on computer. As a consequence, videos, especially those with high hit rates, attract specific actors to the technology and induce them to behave in novel ways. While VOCALOID technologies enlist specific types of users, the complex system of technologically mediated voice and adoption of anime characters also discourages others who have no interest in these things from entering the network. For this reason, it is difficult to tell how the interested parties are enrolled in or dropped from the networks. The VOCALOID community is evolving, yet unstable and shifting, as the networks are putting it together and pulling it apart at the same time.

The uncertainty of group formation moves my analysis away from a static listing of people and objects to a study of the dynamic associations built from activity. The concepts of actors and actants are two of the important concepts in ANT to settle the dispute of groupings. For ANT theorists, actor is any element that makes other elements to act upon itself so as to create an alignment of the other actors' interests with their own,

whereas an actant may be enrolled as allied to give strength to a position. An actor must be able to translate their will into a language of its own, and therefore translate others' will through its own material form and properties (Callon and Latour 1981: 287). In other words, an actor has an effect or impact on whatever actor. As a consequence, the creation of a heterogeneous network of aligned interests, namely, the actor-network, results. The strength and number of immediate connections that bind heterogeneous actors needs to be repeatedly "performed"; otherwise, the network will be dissolved. Accordingly, I attempt to identify the human and nonhuman actors involved in the activities in order to sustain the manipulation of networks. I start with disseminating the VOCALOID synthesizer as a black box that the network of activity begins to reassemble, followed by considering it as an actant that is granted to be the primary source of action. The perspective of ANT allows me to consider both human and nonhuman elements equally as actors in a network that takes shape through interactions with one another. Using an example of VOCALOID video as an actant, I then turn to analyze the circulation of videos that takes part in establishing the actor-network on video-sharing sites, which subsequently leads to a chain of covers and creative works that construct the VOCALOID culture and an amateur music-making community in turn.

1) YAMAHA Corporation as Actant and VOCALOID Synthesizer as Black Box and Actant

Since VOCALOID technology is a commercial singing synthesizer sold as "a singer in a box," most people are aware of the Vocaloid characters and their voices but

have no sense of how the synthesizer works. To start the enquiry, I treat the VOCALOID

singing synthesizer, its developer YAMAHA and third-party companies as the focal

actors that initiate the VOCALOID network by allowing users (or consumers) to make

original songs at home. In *Resembling the Social*, Latours defines network as "an

expression to check how much energy, movement, and specificity our own reports are

able to capture" (2005: 131). The energy and movements he refers to are the actions of

the agencies and their translations between actors. I unpack the synthesizer as the

technological "black box" resulting from translated agencies in the manufacture of the

product. The diagram in Figure 3.1 depicts an Actor-Network in the black box of

VOCALOID in which the developer YAMAHA may serve as "translator-spokesman."
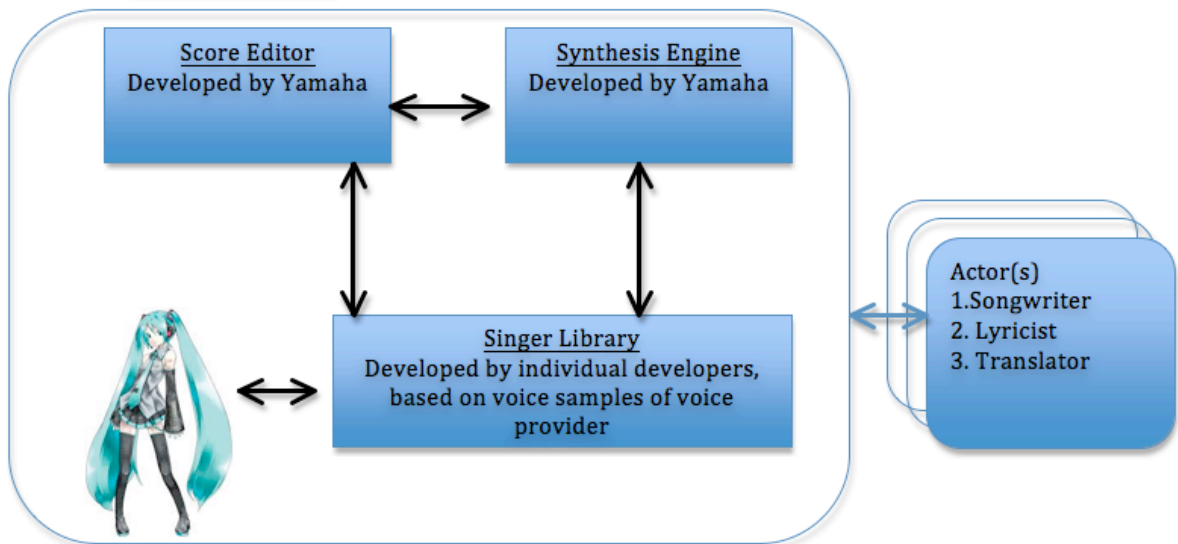

Figure 3.1: A diagram of Actor-Network based on YAMAHA as "translator spokesman"
in the "Black Box" of VOCALOID synthesizer.

Although artificial speech synthesis has existed since the 1960s, singing synthesis

is a much more recent endeavor. Researchers have been working on getting the computer to sing like a real human singer for a long time, but the complexity and flexibility of the human voice in imitating words and conveying emotions makes it the most difficult instrument to simulate computationally. The human ear is so attuned to the human voice that subtle nuances, errors, and tonal shifts are apparent. Nonetheless, recent VOCALOID singing synthesis technology offers new modes of agencies for professionals, amateurs, performers, retailers, and manufacturers. YAMAHA Corporation, the Japan-based multinational instrument manufacturer, has been developing instrumental synthesizer technology for many years, but the development of singing synthesis only began in 2000, and the first commercial singing synthesizer VOCALOID was launched in 2003, driving the technology out of the research laboratory into the reality of the consumer world.

VOCALOID products are designed and marketed by third-party companies under license from YAMAHA. On the technical side, the black box of VOCALOID technology operates in three layers: Score Editor, Synthesis Engine, and Singer Library. On the one hand, the Score Editor and Synthesis Engine, which are developed by YAMAHA Corporation, together function like a traditional sequencer (in composition processing) and a traditional synthesizer (sound data processing) respectively. On the other hand, the distinctiveness of Vocaloid's voice is generated at the level of a database called the Singer Library, in which its distinctive database is constructed from concatenated voice samples taken from real people. Individual companies have introduced products that

make use of the YAMAHA VOCALOID technology to develop their Singer Library in which the database must have all possible combinations of phonemes of the target language. When users utilize the Score Editor interface to enter notes and lyrics, and to edit and adjust the parameter as necessary to achieve the desired nuances, the synthesized singing voice is generated in the Synthesis Engine after the selection and concatenation of necessary sound fragments from the Singer Library. Additional expressions such as breathes, vibratos, brightness, and dynamics can be added to the voice in order to closely replicate a human singing voice.

Despite the innovative idea that makes singing synthesis a reality, the initial sales of the first version of VOCALOID were not good, as the software interface was reported to be difficult to manipulate and the concatenating technology still had room for improvement. The research team was reduced in size by 2007. However, they had a discussion with one of their licensed partners, Cryption Future Media, and chanced upon the incredible idea of having an anime character (i.e. Hatsune Miku) attached to the synthesized voice in the market release of VOCALOID 2. Hatsune Miku became a major hit in Japan, leading to an Internet-driven boom in the creation and dissemination of her products. Most notably, the anime character makes the synthesized voice more acceptable because audiences can place the character as the person singing the song. With the attachment of the anime character to the synthesized voice, a singing performance sounds more natural, and the characters become the virtual singers of the show. Producers are free to choose the desired vocalists to sing on their music (see figure 3.2). VOCALOID

software together with the anime characters becomes the actor as the voice and character become the factors for creator-users to choose their desired voice.

VOCALOID	Version	Developer	Voice Provider
Leon/ Lola	Vocaloid 1	Zero-G, 2004	unknown
Hatsune Miku	Vocaloid 2	Crypton Future Media, 2007	Fujita Saki
Kagamine Rin/ Len	Vocaloid 2	Crypton Future Media, 2007	Shimoda Asami
Megpoid 'Gumi'	Vocaloid 2	Internet Co Ltd, 2009	Nakajima Megumi
Yuzuki Yukari	Vocaloid 3	AH Software, 2011	Ishiguro Chihiro

Figure 3.2: Notable VOCALOIDs (Nozawa, http://www.semioticreview.com/pdf/open2013/nozawa_characterization.pdf. Accessed March 30, 2014)

To this end, YAMAHA is the focal network actor in the development of the commercial VOCALOID synthesizer and may function as "translator-spokesman." Translation is one of the central concepts in ANT as Latour frequently substitutes the terms "sociology of translations" in place of ANT (2005:9). Considering translations as the basis of all relations, Latour remarks, "There is no society, no social realm, and no social ties, but there exists translations between mediators that may generate traceable associations (ibid: 108). The definition of translation in ANT here refers to the conditions of interactions when agencies are brought into agreement with each other in such a way

that all actors agree that the network is worth establishing and building. In effect, network elements are defined and ascribed roles through translation. The spokesman's desire is ultimately translated to others' desires, gathering the actors (or interested parties) into a network of limited activities. The focal actor defines identities and interests and asserts an effect on other actors that is consistent with its own interests (Callon 1986). With this in mind, YAMAHA's intention to commercialize VOCALOID by licensing it to third-party companies, in effect, determined who would be involved (e.g. producers of vocal libraries, software marketing firms) and how they would establish an association. YAMAHA and the licensed company interact with each other to accomplish the technological improvements on the quality of simulated voice in the product, and to decide the marketing of the product.

The spokesman has an effect on the actors' desire, which involves a process of enrolling other actors to accept the interests defined by the focal actor. It is the technology that enables people to enjoy the result of making the computer to sing. According to the YAMAHA official site, their vision of building VOCALOID is to "create a product that would allow even an unsophisticated user to generate a singing voice at low cost and play it back in real time." With the emergence of electronic instrumentation in popular music, while a single synthesizer is capable to produce an enormous and rich orchestral sound using fewer instruments and musicians, humans are still required for all of the singing parts. YAMAHA observes that there is a potential to develop the simulated human voice if there were financial or logistic constraints in hiring

a human chorus. As the head of the VOCALOID research team at YAMAHA, Kenmochi

Hideki, revealed that the simulation technology of various musical instruments has

already become sophisticated, yet the singing voice is the last piece of the puzzle to be

placed.[21] The initial intention of developing the commercial software was not to

substitute for human singing but to build "a kind of new musical instrument."[22]

VOCALOID technology and the Score Editor interface make it easier for amateur

producers to create a song by entering notes and lyrics even if they could not read full

musical notation on a staff. YAMAHA's intention to develop singing synthesis is

translated to other actors' desires. The spokesman YAMAHA gathers actors (i.e.

individual manufacturers, creator-users) who have shared interests together, bridging

multiple desires of creating computer-singing songs from the onset of the VOCALOID

network.

      The role of "translator-spokesman" has an effect on the actors' desire through

technological associations in building a network. This aspect of translation also

implicates that the spokesman defines the network through exclusion. ANT theorist John

Law summarizes this outcome as the concept of interresement. Law writes, "The

elementary case of translation is thus triangular. One entity enrolls. Another enrolled.

And a third fails in its attempts to enroll" (Law 1986:71). While VOCALOID technology

enlists specific types of actors, it also denies others at the same time. Technology directly

---

[21] Hightrancesea. 2013. "Kenmochi on VOCALOID and Culture." *VOCALOIDISM Blog*. http://www.vocaloidism.com/kenmochi-on-vocaloid-and-culture/ (Accessed March 9, 2014).

[22] Kaufman, Rachel. 2011. "Speech Synthesizer Could 'Resurrect' Dead Singers." *Wired Magazine*. http://www.wired.com/underwire/2011/12/ueki-loid-speech-synthesizer/ (Accessed January 27, 2014).

affects the spokesman and interresement processes that enroll and deny others from participating in a network. The technical constraints in manipulating the software are an example of interresement, as technical constraints block others, especially those who have no prior knowledge of computer music to use the software. Once producers encounter some problem but have no one to help them to solve it, the process of interresement provides a focal point for us to consider the dissolution of connections and the failure of networks.

The singing synthesis software, especially the first version VOCALOID released in 2003, is intended for professional musicians as well as DTM (Desktop Music) users and was sold based on the idea that the limits of the software depend on the users' skills. Since the VOCALOID synthesis engine is designed for singing more than reading the texts aloud, the Score Editor offers various parameters to edit and add expressions to the singing voice. The user is supposed to optimize those parameters that best accommodate the singing voice with respect to the specific source. The complex system of the synthesis engine enlists professionals, but denies amateur users. The major drawback of the first version VOCALOID was reported as the daunting amount of detailed editing of the various expression parameters that was needed to craft a natural-sounding vocal line.[23] According to my informants, the core members of the Project H.K.V.I.P. (short for Project Hong Kong VOCALOID Independent Producers) DJ Ninja and Y.I.N. express that the complex interface and the parameter editing skills have been barriers for

---

[23] Walden, John. 2004. "Vocaloid Miriam: Singing Synthesis Software (Windows)." *Sound on Sound*. http://www.soundonsound.com/sos/dec04/articles/miriam.htm (Accessed March 10, 2014).

beginners who want to use the software. The intention to form this group was to gather VOCALOID producers in Hong Kong, publish a VOCALOID CD album annually, and provide technical support for members and those who are interested in writing VOCALOID songs. For that, H.K.V.I.P. had organized several tutorials for beginners to learn the editing techniques. However, DJ Ninja notes that it is not too daunting for beginners to edit short phrases, whereas trying to craft a natural vocal line for a full-length track would be arduous. For these reasons, the tutorials were not successful, as many of the beginners did not adapt well to the program and even quit the tutorials. Despite the fact that the latest version of the VOCALOID engine has been improved in the intelligibility of the singing, the technical aspect of the singing synthesis software not only expends associations through recruiting actors in the network, but also blocks other potential users in the meantime.

The language of the software is one of the common examples of constraints encountered by non-Japanese speaking users, as it may discourage the actors from not speaking the language to use the software, cutting off other potential creator-users. Even though the VOCALOID software is primarily available in Japanese and English and other five languages in the later versions including Spanish, Korean, Chinese Mandarin, Catalan, and French, the concatenating technology is optimized best in Japanese.[24] DJ

---

[24] Generally speaking, due to linguistic factors, VOCALOID in the Japanese language generates more natural singing than in other languages. In Japanese, most syllabic sounds are open syllables ending in a vowel. When compared to English, English has many closed syllables ending in a consonant, and consonant-consonant and consonant-voiceless diphones. In comparison, Japanese requires 500 diphones in the database, whereas English requires 2,500. Thus the synthesized singing generated from a Japanese VOCALOID is considered to be more natural due to its relatively less complex pronunciation.

Ninja and Y.I.N. remarked that the synthesized singing sounds better and more natural in Japanese when compared to the English and Chinese Mandarin. Within voice-tuning techniques, for instance, the voice of Japanese singing Vocaloids such as Hatsune Miku, Lin & Len, and Gumi are easier to adjust the parameters to get the desired singing effect, whereas the English Vocaloid Luka requires more editing techniques to make it sound natural. For this reason, they prioritize the use of Japanese Vocaloids in creating songs or making fun. Nonetheless, Japanese lyrics become a part of the concern in songwriting. While DJ Ninja and Y.I.N. are fans of Japanese anime and video games and have been exposed to Japanese texts for many years, their Japanese is not sufficiently proficient to write the lyrics for the songs. To deal with this problem, DJ Ninja first drafted the lyrics in Chinese and then translated them using the Google Translate online translation engine, to see if the lyrics fit into the melody. He also asked his friends for those good at Japanese to translate and proofread the lyrics, or posted the lyrics on forum if anyone replie to him. Accordingly, DJ Ninja admits that the process of editing lyrics and songwriting goes back and forth because it takes time to communicate with the translator. Another scenario is to give up the lyrics but allow the voice to sing a single sound throughout the song. For instance, their fellow member Sky Flight created a song "The Legend of Revival" using Vocaloid IA by inputting a single vowel "a" to sing the entire song with accompaniment, though the overall effect is less fulfilling.

As a result, a Vocaloid song is not produced by a sole person; rather, the process could involve any interested parties, thus making the song an outcome of a collaborative

effort. The act of music making is shaped through the technological affordances of the software. By this, affordance is the capability of technologies to enable interaction in particular ways. When considering language as a mode of affordance, language has become a factor that affects interresement processes that block others from participating in the network. The connections of the network may be interrupted and dissolved if the producer fails to find any actors to solve the problem. The process of interresement also indicates that the success of the network ultimately depends on the interpersonal relationships between the actors in the VOCALOID community.

In the above examples of interresement, interactions between the software and users disrupt and limit associations as much as they build associations. Computer knowledge and artifacts play an important part to define the network through exclusion. In the following section, I will consider video-sharing sites as actors that shape and affect the networks and whatever flows through them through their technological affordances and constraints. As another example of interresement, more importantly, the next section explores a set of questions: how do the various human and nonhuman actors in this network affect each other, and what is the effect on VOCALOID community as a whole?

2) **Video-Sharing websites as Black Boxes and Actors; Songs/Videos as Actors**

Turning my analysis from the synthesis software to the media itself, I now treat the video medium and the video-sharing websites as actors that contribute to the massively collaborative creations building on the VOCALOID network. In an important

sense, many of the VOCALOID original songs were made into animated music videos with either moving or still images of the Vocaloid character, and then were uploaded and widely circulated on video-sharing sites. The songs are created in the manner of interesting videos in hopes to draw as many hits as possible, which serves as an indicator that the works are valuable to others. Therefore, video is the medium that has inspired viewers (i.e. new actors) to participate in the network and to create new contents such as songwriting, illustration, video editing, and so on. With this in mind, the video-sharing website, which is an example of the Web 2.0 media technology, allows users to perform these actions. In the meantime, the video hosting sites also impose additional constraints on video-sharing. Since VOCALOID videos are mostly uploaded and circulated on YouTube and the Japanese video-sharing site Niconico (literally, "smile" in Japanese), I first consider different affordances and constraints of the two video-sharing websites imposed on video production and style of participation. Next, my informants will give an account of their considerations when they promote their works on video-sharing platforms. The affordances of media technologies, including their features and constraints, offer focal points determining the distributions of videos and the participations of users through the Internet network.

I first start with an examination of the features of Niconico, not only because it is the award winning, most popular video-sharing website in Japan, but also since it is the most well-known primary site for VOCALOID producers to share their original works. The head of the VOCALOID research team Kenmochi once reiterated the importance of

Niconico in promoting VOCALOID fandom because it helps amateur creators to spread their works in Japan as well as overseas.[25] As a video-sharing website, the basic service of Niconico closely resembles that of YouTube: users can upload, view, and share video clips, but it has some unique functions that differ from YouTube and other video-sharing sites, and thus has attracted users rapidly in Japan. Significantly, the most unique function is the synchronized and direct overlaying of public commentaries as the video plays. Viewers enjoy not only posting comments but also seeing the comments on the live streaming video, virtually creating an interesting online message board and a sense of a shared watching experience. In addition, viewers can add their responses to videos at a specific position in which the video maker can know which moment is appreciated by the viewer. One may note that the commentary may inspire or encourage the video maker or others to produce new videos related to the video content.

[25] Hightrancesea. 2013. "Kenmochi on VOCALOID and Culture." *VOCALOIDISM Blog*. http://www.vocaloidism.com/kenmochi-on-vocaloid-and-culture/ (Accessed March 9, 2014).

Figure 3.3: Screenshot of a Hatsune Miku video "Miku Miku ni shite ageru" (literally, I'll make you Miku Miku-ed"). The overlaid comments are the lyrics of the chorus.

Another unique feature of Niconico is that both uploaders and viewers can add or edit the tags of the videos in order to categorize the clips, whereas up to five tags may be locked by the uploader. Viewed in total, Niconico invites viewers to add additional effects onto the videos. Although YouTube shares similar functions such as channel subscriptions, video views, sharing, "likes" and comments, YouTube does not allow users to post comments to a specific time. Niconico is more entertaining to watch as it offers an alternative way for viewers to engage in the video-watching experience, which encourages a kind of social interaction and collaboration.

But while the commenting system of Niconico may seem to liberate the passive

viewership on other video-hosting sites, it has other features that serve as constraints, a type of agency that affects the VOCALOID producers to upload videos. I suggest that Niconico constrains video creation and distribution through membership. The video maker must understand the video editing technologies and the video hosting sites constraints.

As a video hosting site, Niconico make use of the live streaming technology that distributes multimedia contents to multiple users at the same time. In general, the transmission speed of the streaming video is constrained by the available outgoing bandwidth from the streaming server. Due to limited server capacity, registration is needed in order to watch videos at Niconico. Niconico allocates the limited bandwidths by setting up two membership accounts for users, free membership and premium membership. Niconico limits the amount of free users accessible to the website during peak hours (7 pm to 2 am). The playing speed and the quality of video is affected at the time of heavy user traffic. Nonetheless, by paying a small amount of monthly fee (525 yen, about US$6), premium members enjoy faster upload speeds, high quality videos, priority upload, no restrictions on the number of videos uploaded, and special access to the most popular shows. The table in Figure 3.4 shows the different upload services provided by Niconico for the two accounts.

Free Membership	Premium Membership
Uploading	**Uploading**
Uploading queued during peak hours	Upload anytime
Video storage space	**Video storage space**
Always 2GB	Starts at 8GB and additional 1GB each month
Upload Format	**Upload Format**
File size: limited to 40MB Bitrate: 600kbps, Resolution: 1280x720	File size: up to 100MB for high-quality videos Unlimited bitrate and resolution
Time Release	**Time Release**
Uploads release immediately	Specify when to release uploads

Figure 3.4: Upload services for free membership and premium membership. Source: Niconico official site http://www.upload.nicovideo.jp/upload (Accessed March 15, 2014).

As seen from above, the upload activity of free users is restricted at Niconico. The duration of 40 MB video is about six minutes, while all YouTube users can upload videos up to 12 hours in length and the file size is limited to 2 GB, or up to 20 GB if requested. In addition, YouTube allows a resolution of up to 3840x2160 pixels, which is thrice as pixels than Niconico.[26] Given the different affordances and constraints imposed by Niconico and YouTube, Y.I.N. explains his preference regarding video-sharing sites when he share his works:

> Watching Niconico video is a spectacular experience, as the overlaying comments on the video are definitely a unique feature of Niconico. Commentators can choose the font size and color at specific position and time. On top of Japanese,

---

[26] YouTube. https://support.google.com/youtube/answer/4523193?rd=2 (Accessed March 15, 2014).

comments in different languages such as Korean, English, and Chinese could be seen too. But sometime the "cloud" of comments is quite annoying because there're many meaningless words and icons. But for myself, I never leave any comment on Niconico in Chinese, because I feel it's impolite and inappropriate to leave comments in other languages on the video made by a Japanese uploader as I assume they can't read it anyway […] I enjoy browsing VOCALOID videos on Niconico because it is the place where the latest, original videos were uploaded. Thanks to the membership system and tags, it's easy to trace the original video and its songwriter […] however, there're some drawbacks of the website. When I come to share my works on Niconico, the limited file size of the uploaded video is a constraint for free member. I can't upload videos with too many animations as the file size might exceed the range. Also, the storage space on Niconico is limited too. I can't upload as many videos as on YouTube. But I'd post the videos on both websites, one with compression or no animation on Niconico and one with more animation on YouTube. Because I'd like my works to be appreciated by Hong Kong people as well as Japanese audiences (Y.I.N. 2013).

ANT theorists suggest that the network is the result of actors being tied together and constrained through multiple connections. In turn, the songs and the videos may be remade in order to fulfill the requirements of the hosting sites. The following diagram in Figure 3.5 represents the multiple translations that occur as we consider video hosting sites and the video as actors in the VOCALOID video-sharing network.

"Black box" of the original video produced and uploaded by producer

Black box of
video maker

Song made by
VOCALOID

Illustration/
Animation of the
Vocaloid character

Video Editing
Technology

Other Uploaded Cover
Videos
e.g. singing, dance, fan-edit
videos, cosplay, etc.

Niconico or YouTube as
hosting site
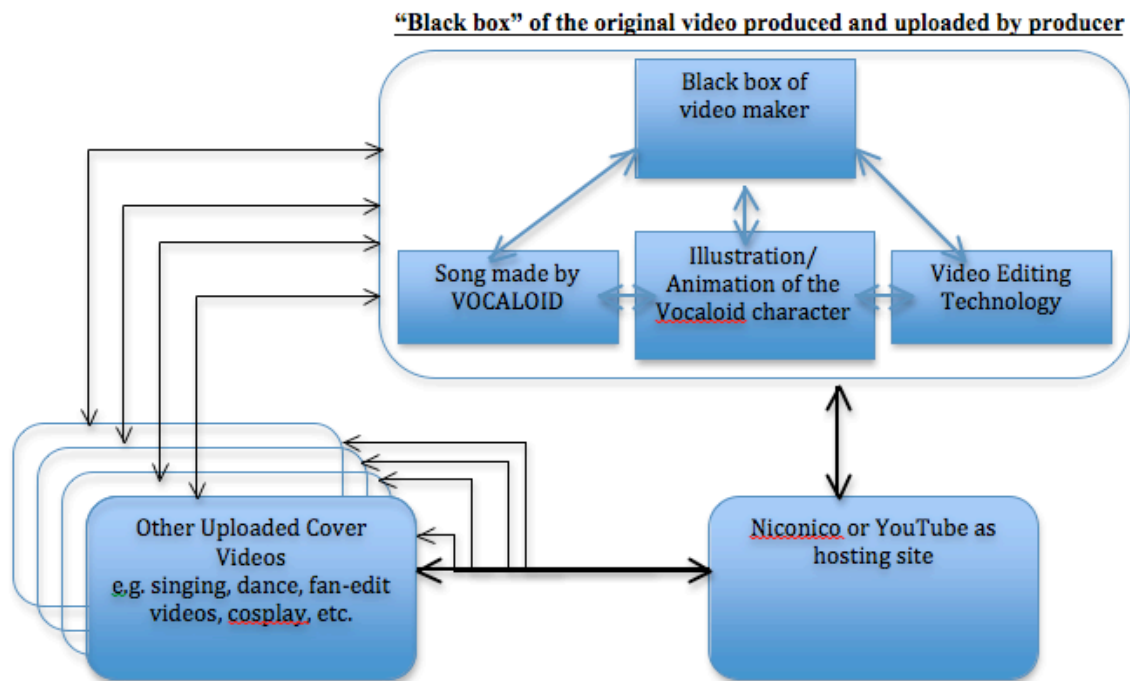
Figure 3.5: Network diagram of VOCALOID video on video-sharing sites

Niconico constrains the file size of video uploaded by free users, whereas the

components of the video black box, 3D animation program and video editing

technologies, determines the amount of editing on the video clip. To reach different

audiences, Y.I.N. considers making up compressed versions of videos in order to upload

them to different websites. This constraint reveals itself through the visual quality of the

video. The lower resolution of video owes to the compressed file and slow streaming

speed. The Figures 3.5 and 3.6 compares the video quality of the specific playback time

(00:19) viewed on Niconico and YouTube respectively.

Figure 3.6: Screenshot of "ライバルと書いて☆友達と読む" video on Niconico. Music and animation by Y.I.N.

Figure 3.7: Screenshot of "ライバルと書いて☆友達と読む" video on YouTube.

From the above figures, the resolution of the YouTube video in Figure 3.7 is better than that of Niconico video in Figure 3.6, whereas the facial expression of the characters, drawing, color, and lyrics are blurred when viewed on Niconico. Illustration and animation are the components not only to deliver the story and the message of the lyrics, but also to attract viewers and fans. The low resolution of Niconico video may not be well received by viewers. Given that the quality of video could be improved, Y.I.N. never thinks of upgrading to premium membership. The monthly fee is one of his concerns for sure. Arguably, Y.I.N. notes that VOCALOID fans value music more than the video itself. He would be happy if fans appreciate his tuning techniques in the natural sounding of the synthesized singing voice more than the video and Vocaloid characters.

As for video views, taking the VOCALOID song "ライバルと書いて☆友達と読む" (literally, studying with friends and competing with each other in an exam) as an example, the numbers of views on Niconico and YouTube have accumulated to 773 and 951 respectively. Since the majority of the viewers at Niconico are from Japan and the VOCALOID culture in Japan leads the trend, Y.I.N. is more thrilled to receive video views at Niconico in hopes of having cultural exchange with Japanese VOCALOID producers. Niconico serves as a localized VOCALOID network catering to Japanese audience. Yet, Y.I.N. and his fellow members cannot deny the fact that the membership system of Niconico becomes a barrier for those who do not have account at Niconico to access their works.

As demonstrated above, Niconico allows VOCALOID producers to upload their

works; on the other hand, it also carefully limits actors to perform activities. The affordances and constraints instigated by Niconico may assume a position of power as an effect. Nonetheless, it does not guarantee the stability of the network. Rather, the concept of networks in ANT is premised on their instability. Actors, translations, interresements often cause communities and networks develop in unpredictable ways. In fact, Niconico, YouTube, and other social networking websites, cannot control the translations and stability of the network, as every node in a network is a mediator, contributing and shaping the networks as networks flow through each node (Latour 2005: 31). Any actor involved in the network has the affordance to transform the network by the act of translation.

Having identified how the technological nonhuman actors play their roles in establishing the actor-network, the next chapter will step fully into my ethnography of VOCAOID fan activities in Hong Kong. In the summer of 2013, I follow the actors, Project H.K.V.I.P. members, to visit a VOCALOID event and interview VOCALOID fans that provide me invaluable insights in the network and the community as a whole. I would like to offer a few profiles of VOCALOID fans in an attempt to reveal the connections and different roles of actors in a localized VOCALOID community. Fieldwork can help us to understand some of the dynamics of the before, during, and after aspects of the community.

Chapter 4: An Ethnography of VOCALOID Fandom in Hong Kong

A snapshot of current VOCALOID fandom requires a methodology that integrates interpersonal interactions, media and musical technology into a united framework. ANT can be used as a tool to reveal the complexities of the sociotechnical VOCALOID fandom. Latour states that ANT "was never meant to be a theory of what the social is made of…[but] simply another way to be of being faithful to the insights of ethnomethodology" (Latour 1999: 19). As a methodology, ANT studies calls for direct observations and focuses on local interactions. In a similar vein, the ethnomethodology applied to localities produces thick description as conceived by anthropologist Clifford Geertz (1973). Like ANT scholars, Geertz believes anthropologists' role was "not to generalize across cases but to generalize within them" (Geertz 1973: 20). Examinations of local activities set out an account for how informants interpret their activities, interactions, and what happens in a network. Like Geertz's thick description, ANT theorists focus on the interpretation of localized activities, but argue for a "panorama of localities," which consists of localities connected by traceable associations and actions that take both people and technologies into account (Latour 2005: 183).

Within the past decade, different forms of media constitute the ties between localities. Consequently, global and regional connections are loosely comprised of localities of varying sizes. By the same token, although VOCALOID is a technology that originated in Japan, the circulation of songs and videos in localities is made possible by Web 2.0 Internet technology, assembling the global and regional connections across

nations. With this in mind, this report does not mean to provide an overarching account of the global effect of the VOCALOID technology and suggest how the new musical technology involves more amateurs in creating songs with ease. Instead, I describe the connections between the people and define how people are assembling within the unstable and shifting networks through localized activity.

ANT offers a unique methodology in the treatment of VOCALOID activities. Methodologically, ANT sets out to "follow the actors" via interviews and ethnographic research in order to trace through whose eyes to look at the process of network construction (ibid: 28). As each actor is bound to leave traces of relations, following the actors is the only way that these relations can be observed. The implications of this perspective directly address the notion of community by following the actors' own activities, particularly the ones that challenge the predominant social hegemonic power structure. Local activities can no longer be summarized into the homogenous term such as regional, social or global; any anime and manga related activities should not be categorized as otaku culture. The theoretical ANT network replaces all that is commonly referred to as "social," yet requires the analyst to define the network through localized activity. Every locale offers the possibility of its own aggregates of constituents, relationships and behaviors that make up the heterogeneous networks.

In this study, the perspective offered by ANT has directed my analytic focus towards localities instead of global forces. I use ANT to focus on the VOCALOID activities in Hong Kong in accordance with the music, lyrics, illustrations, and many

derivative works produced by fans. The ethnography mainly draws on my interviews

with an active group called Hong Kong VOCALOID Independent Producer (H.K.V.I.P.)

in Hong Kong during the summer of 2013. Following my informants, I attended the

VOCALOID events organized by a local subculture association. By tracing the

connections in which the fans become motivated to share, collaborate and create, I argue

that their motivations can be understood by looking at how they are attracted and

recruited to the networks, and behave in novel ways. But before I delve into my

ethnography in Hong Kong, I first identify the various technological nonhuman actors

and their connections with human actors in the VOCALOID network.

Figure 4.1 & 4.2: Offical osters of the two events: "Miracle Vox 3" and "Netstage Fiesta

2013" concert.

In the summer of 2013, during my short trip back to my hometown Hong Kong I came to conduct an ethnographic fieldwork in a one-day "VOCALOID-only" free admission event called "Miracle Vox 3" at Kowloon International Trade & Exhibition Centre. Thanks to my informant DJ Ninja, I came across this event and met with his fellow members. Gradually, I was able to observe the community and talk to a couple of VOCALOID fans, and gained a better understanding of their love of the VOCALOID subculture.

"Miracle Vox 3" (literally, Miracle Voice 3) was the third annual event organized by the Hong Kong based subculture promotion company Midgard Development Limited. The one-day event comprised two sections in the afternoon and evening. The first section was a marketplace from 11am to 5pm in which exhibitors sold self-produced wares at their booths, then followed by a live concert at 7pm performed by Japanese "Niconico singers," who were singers who made VOCALOID song covers on Niconico and became famous. All in all, it was a brilliant idea to hold the two related events in a day as fans could attend the events at the same venue at ease.

During my ongoing ethnographic incursions into the Internet in Spring 2013, DJ Ninja invited me to a forthcoming VOCALOID event in late June called "Miracle Vox 3" where they would have a booth selling their second CD album "The Legend of Revival." He suggested that I could have a detailed interview with their members at their booth. I was more than thrilled because not only this was one of the few "VOCALOID-only"

events in Hong Kong, but also I was able to meet my informants for the first time since I

contacted Project H.K.V.I.P. via their Facebook fan page in April 2013.

Project H.K.V.I.P.: DJ Ninja and Y.I.N.

Let me start with my first informant DJ Ninja who suggested that I check out the

event. DJ Ninja is one of the core members and the spokesman of the fan group Project

H.K.V.I.P.. He is also the administrator of the Facebook fan page, and has been very

helpful in answering my inquiries since I first contacted the group.

Figure 4.3: Project H.K.V.I.P. booth at "Miracle Vox 3." Y.I.N. was promoting their two
albums "Disaster and Prophecy" and "The Legend of Revival," June 30, 2013.

Photograph by author.

During our initial conversation on Facebook regarding his interests in VOCALOID, DJ Ninja first introduced himself as a longtime anime fan and an "ACG DJ," an amateur disc jockey who mixes recorded anime songs for audiences in ACG DJ party.[27] Ninja, which is his adopted pseudonym, denotes his professional skills to cut up and remix the soundtracks, like a ninja's ability to perform covert assassination with his sword in feudal Japan. His affinity for anime music began in the days of Nintendo game music and Japanese animation since his childhood. He explains that he is drawn by the energetic, expressive singing and catchy melody of Japanese anime title songs in general. He began to collect anime musical soundtracks and started to remix the music in order to create interesting effects. The art of turntable techniques such as cutting, bridging, scratching, cueing, and harmonic mixing lead him to enjoy the process of music creation. In relation to how he started to write VOCALOID music, he explains,

> "As an anime fan, we all know that the Vocaloid characters are very popular in Japan and Hong Kong in recent years, and the VOCALOID software grants the grass-root amateur songwriters and Nico cover singers a lot of opportunities to express their talents in the competitive music industry. Since I'm an ACG DJ, I have to closely follow the track of the popular trend of Japanese anime culture. I have listened to VOCALOID songs on Niconico and played VOCALOID songs in the ACG party, I have a strong feeling that I have to get to learn how to write VOCALOID songs in long run. So I started to do some experiments on the Hatsune Miku and Luka VOCALOID software. FYI, unlike my fellow members

---

[27] "ACG" is a short hand that stands for anime, comic, and game. According to DJ Ninja, ACG DJ party is a type of party in Japan that features anime music and performances in bars and nightclubs. ACG DJ selects and plays anime songs using multiple turntables to mix and blend the recorded anime music, creating different grooves to instigate the audiences. ACG DJ party has become a popular phenomenon in Japan. There are ACG DJ parties almost every night in Japan but it is not common in Hong Kong. For this reason, DJ Ninja would like to be the pioneer to bring about this culture in Hong Kong.

who have prior experiences in playing musical instruments, I actually don't read musical notations nor play any musical instruments. I don't have an instrument with me to try out the note. Though I have a good sense of musical styles and remixing of musical elements/styles owing to my DJ experience. The VOCALOID software enables me to create a song with a singing voice and musical accompaniment. Although it takes me a lot of time to figure out how to use the software in order to create the desired effect, the overall outcome is pretty cool as I couldn't imagine I could create a song by myself; with all the necessary components of a song such as singing voice, lyrics, accompaniment, all are done with my Mac!"

DJ Ninja's story in many ways encapsulates the Hong Kong anime fans' entrance into VOCALOID fandom. They stay tuned to every popular trend in Japanese anime and manga culture. Most of the fans get to know the Vocaloid anime characters before listen to the VOCALOID songs on Niconico. DJ Ninja's explanation as to what appeals to him about VOCALOID is also not familiar:

"Well, I have to confess that one of the first things that attracted me to VOCALOID is the cute girls. My favorite Vocaloid character is Luka. But after trying out to create songs using the software, what VOCALOID really attracts me now is not the Vocaloid characters. The creativity built on the character and software is intriguing. Niconico is a good place to learn about people's creativity on music, drawings, and animation. As such, the technology ignites my capability of creating music, which is something that I never imagined before. I find myself connect to the community that share the same interests, which always makes me feel good."

To a certain extent, the creativity grows out of social energy arising from a collective interest in the Vocaloid characters. In general, moreover, the VOCALOID fandom in Hong Kong is not particularly enthusiastic in VOCALOID song creations. As Y.I.N. points out that, "Within the Chinese-speaking VOCALOID communities, Chinese and Taiwanese VOCALOID fans are doing more VOCALOID song creations when compared to Hong Kong fans, as there are many Internet discussion sites where Chinese

64

and Taiwanese VOCOCALOID producers share their works and knowledge. I also learn a lot of voice-tuning techniques from these forums." For this reason, Y.I.N. and DJ Ninja explain that this is their motivation to gather as many as VOCALOID producers in Hong Kong in hopes to promote their original VOCALOID works as well as bring about more collaborative creations and derivative works in Hong Kong VOCALOID fandom.

Founded in February 2012, Project Hong Kong VOCALOID Independent Producers (Project H.K.V.I.P.) is the only VOCALOID music producer group in Hong Kong with nine active members. Having similar interests in games, anime and VOCALOID music, DJ Ninja and Y.I.N. came to know each other through their YouTube channels, and then decided to set up the first VOCALOID producer fan group in Hong Kong. As suggested by the group name, every producer in the group is a VIP; everyone is independent and treated as important as one another. Their goal is to release an album as a group every year. Not only this act can lower the cost of the production, Y.I.N. adds another point to reveal the arduous process of VOCALOID songwriting,

> "Creating a VOCAOID song is not easy… We are all amateurs. VOCALOID is one of our hobbies, thus the time we spend on it is limited. It may take a couple of sleepless nights or several months to write a song as the software is easy to comprehend, yet the parameters are difficult to adjust the desired voice. Just like humans, VOCALOIDs have different characteristics because of different sources of the voice library. It really depends on the producer's acquaintance with the specific VOCALOID and 'luck.' As an amateur songwriter, my dream is to release an album with all my compositions. Yet, it might be impossible and stressful to do so. However, it would become more feasible if everyone just contributes one song to the album. To take a step further, just like the way we get inspired by the songs created using VOCALOID, we hope we could be the one who inspires other."

To achieve the goal, the members first compromise with a theme to tie all the

songs together. For instance, the theme of their first album was "disaster," while the

theme of the second album centers on "legend." Accordingly, they propose a deadline,

and then work individually or collaboratively to create the song using their desired

Vocaloids. Despite the fact that members work independently to write their songs, the

group facilitates interpersonal communications among them, and thus they are more

likely to complete their individual work and motivated to work toward the common goal,

that is, to release an album. As Y.I.N. notes that,

> "It's easy to start writing out a melody because it's always fun to hear the
> synthesized voice singing out the melody in your head. Yet, it's difficult to
> complete a full-track song though because it requires a large amount of editing.
> Frankly, the long process of editing often put me off. I have a couple of
> incomplete works with some musical ideas I really like. I've also posted some of
> the incomplete work to ask for ideas. The best part of VOCALOID is to allow me
> to express myself and write anything I like, and interact with each other. There are
> many possibilities of applying the software in the doushiji culture.With a deadline
> and target, it's more likely to work it out with the support of fellow members."

The idea of working together toward a common goal is more productive. Having

different talents and songwriting experiences of using the software, members benefit each

other with their knowledge and talent whenever they encounter technical problems of the

software. DJ Ninja remarks that a member who speaks excellent Japanese helps the

translations and proofreading of lyrics. The finished work results with better quality,

hence the creator attains to higher degree of satisfaction.

Members have brought not only their talents and skills but also their social

networks to the group. Contacts in a network provide information, opportunities, and

perspectives that can be beneficial to the members in the network. As noted earlier, a VOCALOID song is often a result of collaborative effort. The social networks of members bring different talents to the group and help the preparation of the album in many ways. To lower the cost of production, DJ Ninja notes that the process of preparation is more than putting the songs together. Some of their friends have offered help in providing free use of recording studio, and contacts of inexpensive printing agencies. Given that Project H.K.V.I.P. is the only VOCALOID producers fan group to sell their music in the convention, when asked if their album is a lucrative business, Y.I.N. points out that in any event it is unprofitable to produce an album in such a small community. Moreover, he emphasizes the fact that their intention of producing the album is to publish their work instead of making a profit. Nothing is more exciting than the product comes into fruition.

To sum up our conversation, I asked about their opinions on the correlation between otaku, and VOCALOID and anime fandom as a whole. Not surprisingly, given that longtime fans of anime, manga, and video games (which are otaku's all-time favourite), they do not regard themselves as otaku. They see otaku associated to a group of young people who are lack of social skills. In contrast to the asocial behavior of otaku, DJ Ninja and Y.I.N.consider themselves as sociable and outgoing persons because of what they produce and contribute to the fandom. More importantly, they recognize that VOCALOID has expanded the demographic group of fandom from anime fans to music lovers and professionals. Many people, who are not necessarily anime fans, listen to

VOCALOID music, and hence get a better understanding of the Japanese anime and manga subculture. This act considerably rescues the pejorative connotation of otaku.

Syaoran

I met Syaoran at the "Miracle Vox 3" convention. As far as I observed in the convention, the convention was hosted in an exceptionally small conference room where there were only thirteen booths selling their self-produced wares—mostly books of Vocaloid character drawings, stickers, posters, etc. Project H.K.V.I.P. and Syaoran were the only booths selling albums. While Project H.K.V.I.P. were selling an album with their VOCALOID original songs, Syaoran's album compiles a set of VOCALOID cover songs (sung in Cantonese by amateur singers) with his new Chinese lyrics. Syaoran's booth was next to Project H.K.V.I.P., yet they had not known each other. I talked to Syaoran after my interview with DJ Ninja and Y.I.N., and I am impressed to learn that Syaoran's creative work sheds lights on a different kind of network in the VOCALOID fan community.

Figure 4.4: Syaoran's VOCALOID album. Photograph provided by Syaoran.

As with most anime fans, people are friendly and happy to help me in my interviews. Syaoran is clearly thrilled and enthusiastic to talk about his hobby to anyone who has an interest in VOCALOID. Syaoran is in his early forties but looks younger, slim with a neat white shirt. I can tell his interests in anime from his pseudonym, as Syaoran is the fictional character and protagonist in the popular anime *Tsubasa: Reservoir Chronicle*. Surprisingly, Syaoran's "real" career lies in anime, over the years he has been the producer who specializes in animation dialogue dubbing at the largest commercial television station in Hong Kong, Television Broadcasts Limited (commonly known as TVB). Syaoran explains that his profession is to translate and edit the Japanese

scripts into Chinese, and help the voice actors to record the Cantonese dubs in anime and drama series. Having graduated with a university degree in Chinese language and literature, Syaoran has a profound interest in Chinese rhetoric and writing in the tradition of classical Chinese poetry. During his leisure time, he is an active fan involved in various anime activities in his high school and university in which he regards himself as an anime critic and amateur screenwriter. He especially enjoys the collective spirit in the anime fandom where he can put to use his talents.

Figure 4.5: Naru is one of the singers featured in Syaoran's VOCALOID album. She dressed herself in a self-produced outfit, resembling the image of the VOCALOID character Gumi (Figure 4.6 on the right). Photograph by author.

Coming from a literature background, seemingly, it is hard to imagine how a person with no musical background happened to produce and sell a VOCALOID album

in a fan convention. When asked what inspired him to create a VOCALOID album, he

explains,

> "Needless to say, the concept of VOCALOID is beautiful and awesome. It invites many people to express their talents in many aspects. For me, I really like listening to VOCALOID songs and watching those fascinating animation on Niconico. A couple of songs with its stories and catchy melodies stuck in my mind, which inspire me to write new lyrics to express my feelings of the music. Having been an anime fan who has studied Japanese language for several years, I'm confident to write the lyrics in Japanese. Indeed I feel the best to articulate the feelings in mother language. But the reason why I insisted upon writing the lyrics in Chinese due to my profound love in Chinese poem and rhetoric since childhood. As a tonal language, the tone of Cantonese is beautiful, yet challenging to match the melody. It's been a pleasure to write the lyrics that perfectly match the melody in addition to the message or story I want to convey in the lyrics. After I accomplished some works and posted the lyrics on forum, and then I received some fans' request to make a cover song using my lyrics. Indeed everyone is welcome to use my lyrics with acknowledgement. The commentaries of the lyrics were very good after she posted her cover on YouTube, and I'm impressed with the fact that my work is performed by a human voice. The overall effect is fulfilling. Consequently, I came up with the idea to invite someone who has an interest to record the songs and compile them into an album."

Syaoran's album is a compilation of eight songs performed by seven singers. The

production involves several actors, which reveals the high production cost and complex

network. Syaoran explains that although he is the person who starts the project, it is

impossible to complete the project on his own. The project would not have been possible

without the generous help and support of many wonderful people and their talents. In

particular, he expresses his gratitude to the record producer who managed the recording

sessions with all the singers and supervised the entire process through mixing and

mastering. In the meantime, a friend helped him to contact the copyright holders in Japan

to deal with the copyrights of the songs. When asked about the copyright cost, Syaoran is

grateful that the copyrights holders agree the use of their music with no cost as long as Syaoran cites their names in the album, although it took awhile to acquire the permissions from the VOCALOID producers in Japan. "The copyright holders did not ask for a penny. They didn't mean to create the work for money," he said, "in fact, they are more happy to share their work and inspire others. This is also the most beautiful thing in VOCALOID or anime/manga fandom. People are more willing to share their work than any creative medium." With the efforts by different parties, the album had taken for nearly half a year to come into fruition. Yet, Syaoran especially enjoys the process because the music project is new to him. From this perspective, his VOCALOID album is no difference from his prior experiences in other anime *doujishi* activities. What appeals to him about VOCALOID is the creativity and collaboration inherent in the anime fandom. It is always fun to meet people who are passionate to collaboratively create derivative works.

In conversation with Syaoran, in fact, his booth was very popular as many fans talked to him and purchased his album. Syaoran was very busy to keep our conversation ongoing. However, Syaoran was very kind and patience to answer my inquiry even when he was having his lunch. As I observed that not only the fans but also almost everyone in the convention seemed to know Syaoran, I was wondering if he was a big name in the fandom. My puzzle was solved after I talked to Lynn, a VOCALOID fan who came to show her support to Syaoran's album at the convention.

Lynn is a university freshman majoring in Chinese Language and Literature. Lynn is a fan of Syaoran, as she confesses that she came to the convention to show her support to Syaoran after she learnt about Syaoran's effort in making his first album. She has known Syaoran for a long time when she participated in an amateur radio drama in which Syaoran was the screenwriter and dialogue-dubbing director. According to Lynn, since Syaoran has extensive interests in languages and Japanese anime and manga subculture, Syaoran has become a key organizer in various *doujishi* projects. As a renowned anime critic and writer, Syaoran has been a frequent speaker in the events hosted by university anime fan clubs. In particular, she compliments him on his beautiful rhetoric writing and enthusiasm devoted to anime and manga subculture in Hong Kong. Lynn especially enjoys reading his literary works and reviews on his blog where he shares his insider's insights on the development of anime and manga subculture. Using different pseudonyms, his essays are widely read by fans on various forums, though he goes by Syaoran for most of the time. Thanks to Lynn, I can tell why so many fans came to Syaoran's booth to purchase his work and showed their admirations to his multi-talents.

Syaoran was shy when I complimented him on his remarkable contribution to the fandom in Hong Kong. For Syaoran, the best part of the anime and VOCALOID fandom is the freedom and collaborative creativity. Interest in anime has put to use his creativity and talents and has brought him friendships and many people's support and love. My impression of Syaoran by far was a humble, gentle, and talented man who enjoyed very much in what he had been doing. However, he became intimidating and hostile when I

73

asked him about the outsiders' criticism. His furious and rigorous reaction was out of my expectation. He says he recognized that people always antagonized the anime and manga fandom for no reason:

> "Many people are contentious to make irrational and false claims about any anime or manga related activities based on their own value judgment. Regarding this, I understand very well why people called us otaku because of our attraction to virtual characters seemingly. I also never deny the fact that the virtual characters look attractive to me, but there are also many things I learnt from the fandom, which has become the more important factor I love being involved in the community for so many years. The beautiful artwork and the creative storytelling of anime and manga are simply fascinating and inspiring for me. The freedom of expression is something I value the most beyond all things. I believe anime is no different from other hobbies. Agree with me or not, everyone has different tastes and hobbies. I don't see their point to defame our effort. If anime or the virtual characters is not your cup of tea, go away and feel free to ignore what we're doing. You outsiders are not welcome to comment on us with no understanding of our hard works. Anime is my hobby, while *doujishi* is the medium where I find the intimate bonding with people who share the same interests. I've never meant to establish myself as a renowned artist or anything. I'm not asking anyone to appreciate my work neither. My work has never meant to please anyone. I create something merely because I want to do so. We put many efforts in producing a work together in which outsiders would never understand. I'm not asking for their understanding neither. But our effort does have a right to be respected in any event."

I was in awe of his agitated reaction and discontent toward outsiders' criticism. Being an active leader in the fandom, Syaoran tends to put many pressure on himself. This is bittersweet moment when it comes to bear different people's value judgment on VOCALOID or the anime/manga fandom as a whole. Sometimes, however, a slight note of wistfulness creeps into his voice as he talks about the days when anime fandom was a more intimate and personal kind of experience. Yet Syaoran is delighted that so many young people are becoming interested in anime and manga. In some ways, VOCALOID

is now taking the trend to bring more fans into the fandom in which teens and young adults are clearly the top demographic group. Lynn is my next portrait that reveals how VOCALOID draws attention to a teenage girl.

Lynn

As noted earlier, Lynn is a VOCALOID fan who came to "Miracle Vox 3" for two reasons: to purchase Syaoran's VOCALOID album and to attend the evening concert "Netstage Fiesta 2013" performed by Japanese Niconico Singers. Deeply impressed by the music as well as the cute characters, she has been a VOCALOID fan since the Hatsune Miku boom in 2007. Prior to the conversation with Lynn, I had no interests in going to the concert because I had no idea about the performer "Niconico singers" in the first place. At the very least, I knew it was not a Hatsune Miku hologram concert for sure. However, after learning from Lynn that she booked ahead the most expensive concert ticket (HK$880, about US$110), her fan reaction made me wondered if it's something fabulous that I shouldn't miss. I decided to follow her to go to the concert out of curiosity. The concert turned out to be a good opportunity to learn about another type of VOCALOID live performance and what Lynn likes about VOCALOID.

Like many others, Lynn became interested in VOCALOID because of the dazzling imagery of Vocaloid characters. The blue short hair, twenty-year-old Kaito is her favorite character. In fact, I was surprised to learn that Kaito was her choice. Kaito was not the most popular Vocaloid characters I had known by far as the songs created

75

using Kaito's voice is not as popular as other characters. Kaito is the early Vocaloid developed and released by Crypton Future Media in 2006, which is a year before the Hatsune Miku boom in 2007. Lynn explained that she particularly likes his gentle, kind-hearted, and oblivious characters. Kaito is like a kind, elder brother to her. Interestingly, Lynn notes that she hasn't fallen into Kaito at the first glimpse, not until a song about Kaito having ice cream that captures her heart. I found her reason intriguing in a way that ice cream becomes a *moe* element that she really likes about Kaito.

For outsiders who knew nothing about the Nico singers and the popular VOCALOID songs, it was unlikely for them to pay for the ticket and go to the concert. Lynn explained about the Nico singers in more details when we were lining up to get into the venue. On Niconico, there is a craze about covering or dubbing over an original song that someone has created using VOCALOIDs. According to Lynn, the Japanese term "*utaite*" (歌い手) is used to describe amateur singers who cover pre-existing songs and post them on Niconico, whereas *utaite* are known as "Nico singers" to the VOCALOID fandom outside Japan. Moreover, as the recognition and popularity of their song covers grow in Japan, *utaite* begin to make their public appearance and perform in live concerts hosted by Niconico and tours outside Japan.

Lynn told me that the "Netstage Fiesta 2013" concert was the second *utaite* concert she attended in Hong Kong. Although the vocalists that performed in the concert were a few of the popular ones, she still looked forward to the concert as she enjoyed the *utaite* concert last year very much. "In Japan, several VOCALOID concerts and live

76

performances have taken place ever since Hatsune Miku performed her first live concert in 2009. Niconico also helps to promote VOCALOID by hosting a live concert annually called "Nico Nico Douga Natsu Daikaigi." It is a showcase for the popular *utaite*. I watched a few of these concerts and deeply impressed by their amazing voices and energetic performance. However, the VOCALOID fandom in Hong Kong is much smaller, we had the very first VOCALOID concert not until two years ago (i.e. 2011). As such, I never miss any VOCALOID live performances in Hong Kong, such as the Hatsune Miku concert in 2011 and 2012, and the Nico singers concert last year," she said, "I really like going to the live concerts. Indeed the acoustics is much better. Moreover, the best moment in the concert is to scream and wave glow sticks with other fans. While most of the time I sit in front of the computer watching the Nico videos and enjoying the music alone, I feel it's always more fun to get into the crowds and enjoy the music with other people who share the same interests."

Seemingly, her reason to go to a VOCALOID concert makes no difference to other live music performances. However, she notes that the *utaite* performance is somewhat different from other mainstream concert. First, *utaite,* like most voice actors, are usually effaced or obscured from the public gaze at large. It reveals the fact that many people have an amazing voice but their appearances are not attractive enough to become pop star. For Lynn, *utaite* that performed on stage are truly talented singers and music lovers. Lynn enumerates some of her favorite Nico singers who have released their albums under official labels. 96Neko was one of her favorite who performed in the

77

"Netstage Fiesta 2013." Second, she observes that VOCALOID allows *utaite* to interpret the song with greater freedom. As demonstrated earlier, the success of VOCALOID is attributed to the collaborative creativity and freedom that constitute a network. Occasionally, *utaite* collaborate with each other, forming groups or pairs that they collaborate in. She told me that Niconico video-sharing site hosts a weekly ranking of the top 100 VOCALOID songs, which has been her favorite website that allows her to constantly explore the popular VOCALOID original songs or song covers. The most enjoyable moment for her is to explore different beautiful renditions of the same song recorded by different *utaite*.

As a side note, although *utaite* can do song covers for any type of music such as anime and game title song, and J-pop music, the issue of copyright infringement has often led to video-sharing sites deleting cover videos on pop songs and cancelling the uploader's account. With regards to songs created using VOCALOID, everyone is welcome to access the song and make derivative videos or illustration, triggering a chain of creative activity followed by other creative communities. In fact, from the standpoint of the VOCALOID producers, the increase of derivative works raises the value of the originals, with the result that more producers are motivated to share their works. With millions of VOCALOID songs available on Niconico, VOCALOID songs become great resources for amateur singers to express their talents at ease without being put off by the copyright infringement.

Stepping into the venue, Lynn expressed her excitement and started to fan-girling, once the concert was about to kick off. She then took out two yellow glow stick and gave one to me. "Waving glow sticks is a must when one attends any live concerts!" she blinked her eyes. Overall, the concert was very entertaining, featuring twelve *utaite* and two kids. Some of them appeared with their Vocaloid cosplay costume, while most of them dressed in causal outfits according to their personal styles. Occasionally, a translator, who was also the representative of the organizer Midgard Development Limited, appeared during the break and helped the Japanese performers to interact with the audiences. For me, some of the performances were impressive, though I was bored for most of the time because I didn't know the songs and singers. On the other hand, Lynn enjoyed herself very much. Apparently, she knew most of the songs, as she was able to sing along with other fans in the chorus parts. Watching her screaming and waving her glow stick frenetically, I gained a better understanding of the many reasons that she is drawn to VOCALOID fandom.

Conclusion: Thinking Outside the Box

Throughout this study, I start with thinking outside the box of the criticisms and social constructions on the 2D anime characterization and VOCALOID fandom. For that, I have attempted to understand the group formation of otaku strongly linked to anime and manga fandom as a socially determined pathological behavior. As I have demonstrated in chapter 2, the term otaku is a complex black box encompassing social, pathological, dysfunctional, sexual, technological encounters, and so on. I am discontented with the fact that the group otaku is imagined homogenously and globally, yet appropriated and misrepresented as dangerous by scholarships and outsiders in several ways. For this reason, I reject the postulated social explanation of VOCALOID fandom. I have attempted to unpack some of the taken-as-given networks built by the actors that governs the public's misunderstanding of any anime and manga related activities as well as the subject in this report, VOCALOID fandom. The fact is transformed and simplified into a black box by which the network of many actors has helped to stabilize. Once the black box is considered only in terms of its input and output, the narratives are taken for granted or counted as resource. By opening the social black box of otaku formation, it becomes clearer that the group has undergone many interpretations and translations and has been treated as a unified entity by different agencies at play in the "social context." This reveals the constraints of groupings, which reduces explanations to social, natural categories and disregards the heterogeneous elements in the community.

To overcome the constraints of the reductive sociological approach in studying the VOCALOID fandom, the use of "network" in Actor-Network-Theory allows me to shift the analytical focus from the sociocultural theories to observable agencies, which consists of a heterogeneous array of human and nonhuman actors. "The social is visible only by the traces it leaves (under trails) when a new association is being produced between elements which themselves are in no way 'social,'" according to Latour (2005: 8). By this, I try to grasp as much of the visible traces left by the human and nonhuman actors and analyze the connections in these in order to delineate the elements that make up the VOCALOID networks. ANT has provided a useful means of analyzing collaborated practice while foregrounding VOCALOID technologies, music production and consumption, and online media.

I started with looking at the technological agencies, the VOCALOID singing synthesizer and Niconico video hosting site, which are the two major nonhuman actors that establish the actor-network in VOCALOID song productions and distributions in the first place. In chapter 3, I first unpack the black box of the VOCALOID synthesizer in an attempt to identify the various actors involved in the development of new technologies. On the one hand, I identify how the focal actor YAMAHA defines interests and then asserts an effect on other actors or interested parties, gathering them into a network of activities. Yet while VOCALOID technology enlists specific types of actors, at the same time it also denies others from participating in the network. The connections of the network may be interrupted and dissolved if one fails to find the way to solve the

81

problem. On the other hand, the distribution and circulation of songs/videos depend on particular media platforms and their infrastructures, which regulate and restrict the direction and transmission of materials. I suggest that the restrictions of Japanese video hosting site Niconico imposed on the free users result in modifications of the video contents. Viewed in total, in the following ethnographic account of Project H.K.V.I.P., I have shown that the success of a network must be sustained by reinforcing the inter-personal associations of actors. While the constraints and affordances of technologies link the actors into a network, activities within a network must expand the energy arising from a collective interest in order to stabilize social and technological associations.

As a theoretical means as well as an ethnographic methodology, ANT allows me to locate the actors and source of actions in the VOCALOID network. Certainly, the creativity that grows out of the collective interests in Vocaloid characters and/or the music software is the key factor that holds the actors together in the network. To understand the motivations for actors to create, share, and collaborate, I follow the actors' path and let them represent and speak for themselves. In chapter 4, I have introduced the reader to four fans whom I feel are in some ways representative of their different roles in the VOCALOID community. As must be obvious, DJ Ninja, Y.I.N., Syaoran, and Lynn are all very different individuals and have obtained somewhat different satisfactions from their interests in VOCALOID. DJ Ninja, Y.I.N., and Syaoran are clearly creative types, while Lynn shows her appreciation of the creativity and communal energy inherent in the fandom. In the case of Syaoran, my findings reveal that VOCALOID inspires an array of

creativity that does not delimit itself to musical expression. Instead, the open source

culture provides a platform for creative people to express their love in VOCALOID. As I

have mentioned, it is impossible to sum up the VOCALOID fandom in a few portraits. At

the very least, however, these portraits should suggest that VOCALOID fans as a group

are far from the outsider's stereotype of an anime/VOCALOID fan as an otaku who shuts

oneself in his or her room communing only with an anime character or fan object. As I

observed, these are people who are passionate, creative, talented, involved, affectionate,

and enjoy communicating their enthusiasms.

To this end, my analysis identifies human and nonhuman actors, but it does not

mean to indicate the nature of the network in the future. Based on just a few months of

fieldwork and roughly a year of research, the communities and networks I identified may

now disappear or transform within a year's time. According to my most updated e-

ethnography by now (April 2014), my key informant, Project H.K.V.I.P., is no longer

active on their Facebook page since last year's September. When searching the members'

video sharing channels, the last upload activity in Y.I.N.'s YouTube channel was eight

months ago. DJ Ninja is still active on his Facebook fan page updating the recent news of

ACG activities, and has continually uploaded his remixed works on soundcloud.com in

the past year; however, none of the works are VOCALOID songs. Seemingly, the group

is no longer motivated to release a new album this year. As for the VOCALOID-only

annual event "Miracle Vox 3" I attended last summer, the organizer Midgard

Development Limited made an announcement on their official website regarding their

disbandment in November 2013. On my recent visit to their website in March 2014, however, the organization resumed its operation in the preparation of the VOCALOID events this summer. It is impossible to forecast the VOCALOID scenarios in Hong Kong in the future by the date of the completion of this report; nonetheless, the fluidity of group formations in VOCALOID fandom reaffirms the fact that all social relations are, according to ANT, arbitrary, unstable and shifting (Latour 2005: 24).

I, as an outsider of anime and manga fandom, have once taken VOCALOID fandom as a branch of pathological otaku activities. Reflecting on my experiences learning about what the VOCALOID fans are doing while analyzing the socio-musical networks through ANT theory, I have discarded the earlier negative image of otaku and have noted the misunderstandings of otaku and VOCALOID fandom. Things aren't always what they seem to be.

References Cited

Allison, Anne. 2000. *Permitted and Prohibited Desire: Mothers, Comics, and Censorship in Japan.* Berkeley and Los Angeles: University of California Press.

_____. 2006. *Millennial Monsters: Japanese Toys and the Global Imagination.* Berkeley: University of California Press.

Azuma, Hiroki. 2009. *Otaku: Japan's Database Animals*. Translated by Johnathan E. Abel and Shion Kono. Minneapolis: University of Minnesota Press.

Baudrillard, Jean. 1994. *Simulacra and simulation*. Michigan: University of Michigan Press.

Bennett, Jill. 2005. *Empathic Vision: Affect, Trauma, and Contemporary Art*. California: Stanford University Press.

Boellstorff, Tom. 2013. "Rethinking Digital Anthropology." *Digital Anthropology*. Edited by Heather A. Horst and Daniel Miller. London: Bloomsbury.

Callon, Michel, and Bruno Latour. 1981. "Unscrewing the big Leviathan: how actors macro-structure reality and how sociologists help them to do so." In *Advances in Social Theory and Methodology: Toward an Integration of Micro and Macro-Sociologies*. Edited by Knorr-Cetina and Cicourel. Boston: Routledge & Kegan Paul. Pp. 277-303.

———. 1986b. "The Sociology of an Actor-Network: The Case of the Electric Vehicle." In *Mapping the Dynamics of Science and Technology*. Edited by Michel Callon, John Law and Arie Rip. London: The Macmillan Press LTD. Pp. 19-34.

Cohen, Stanley. 2002. *Folk Devils and Moral Panics: The Creation of the Mods and Rockers*. New York: Routledge.

Condry, Ian. 2013. *The Soul of Anime: Collaborative Creativity and Japan's Media Success Story*. Durham and London: Duke University Press.

Deleuze, Gilles. 1972. *Proust and Signs*. Translated by Richard Howard. New York: G. Braziller.

Eng, Lawrence. 2001. "The politics of otaku." *Cornell Japanese Animation Society. Cornell Japanese Animation Society*.

Eidsheim, Nina. 2008. *Voice as a Technology of Selfhood: Towards an Analysis of Racialized Timbre*. PhD dissertation. University of California, San Diego.

Foucault, Michel. 1972. *The Archaeology of Knowledge*. New York: Pantheon Books.

Fumiya, Ichihashi. 2001. *Scenario Filled: Tsutomu Miyazaki Incident*. Tokyo: Shinchosha.

Galbraith, Patrick W. 2010. "Akihabara: Conditioning a Public "Otaku" Image." *Mechademia*, Vol. 5, Fanthropologies, pp. 209-230.

Geertz, Clifford. 1973. *The Interpretation of Cultures: Selected Essays*. New York: Basic books.

Gibson, Willam. 2001. "Modern Boys and Mobile Girls." *The Observer* 1. http://www.guardian.co.uk/books/2001/apr/01/sciencefictionfantasyandhorror.features. (Accessed February 4th, 2014)

Grassmuck, Volker. 1990. "I'm Alone, But Not Lonely": Japanese Otaku-Kids Colonize the Realm of Information and Media, a Tale of Sex and Crime from a Faraway Place." *Lawmune's Netspace. Cornell Japanese Animation Society*.

Haraway, Donna. 1991. "A Manifesto for Cyborgs: Science, Technology, and Socialist Feminism in the 1980s," *Simians, Cyborgs and Woman: The Reinvention of Nature.* New York and London: Routledge. Pp. 149-181.

Hayles, Katherine. 1999. *How We Became Posthuman: Virtual Bodies in Cybernetics, Literature and Informatics*. Chicago: University of Chicago Press.

Hightrancesea. 2013. "Kenmochi on VOCALOID and Culture." *VOCALOIDISM Blog*. http://www.vocaloidism.com/kenmochi-on-vocaloid-and-culture/ (Accessed March 9, 2014).

"The Hologram Rock Star." 2011. *New York Times Upfront*, 143(8), 5. January 10. Retrieved from http://go.galegroup.com/ps/i.do?id=GALE%7CA246178303&v=2.1&u=txshracd2598&it=r&p=AONE&sw=w&asid=019d44597811156dff59babfd3b2d600 (Accessed March 27, 2014)

Ito, Mizuko. Okabe, Daisuke and Tsuji Izumi. Ed. 2005. *Fandom Unbound: Otaku Culture in a Connected World*. New Haven & London: Yale University Press.

Ito, Mizuko. 2005. "Otaku media literacy." http://www.itofisher.com/mito/publications/otaku_media_lit.html. (Accessed February 4th, 2014)

Kamm, Björn-Ole. 2013. "Making-up the Otaku." https://www.gcoe-intimacy.jp/images/library/File/event/20130107Oleさんセミナー.pdf. (Accessed March 30, 2014).

Kamiya, Kousetsu. 2004. "大塚英志「おたく」の精神史" (Eiji Otsuka's Psychoanalytic History of Otaku). http://www1.odn.ne.jp/kamiya-ta/gisi.html. (Accessed March 30, 2014).

Kikuchi, S. 2008. "Otaku stereotype no hensen to Akihabara burando" (The Transformation and Diffusion of Otaku Stereotypes and the Establishment of Akihabara as a Place-brand) *Chiiki Burando Kenkyu,* 4: 47-78.

Kinsella, Sharon. 1998. "Japanese Subculture in the 1990s: Otaku and the Amateur Movement." *Journal of Japanese Studies* 24(2): 289-316.
———. 2000. "Amateur Manga Subculture and the Otaku Incident." *The Subcultures Reader*. Pp.542-543.

Latour, Bruno. 1987. *Science in Action: How to Follow Scientists and Engineers Through Society*. Bristol: Open University Press.

———. 1999. *Pandora's Hope: Essays on the Reality of Science Studies*. Massachusetts: Harvard University Press.

———. 2008. *Reassembling the Social: An Introduction to Actor-Network-Theory*. New York: Oxford University Press.

Law, John. 1986. "On the Methods of Long Distance Control: Vessels, Navigation and the Portuguese Route to India." In *Power, Action and Belief: A New Sociology of Knowledge?* Edited by John Law, 234-263. London: Routledge.

_____. 1990. Technology and heterogeneous engineering: The case of Portuguese expansion. In The social construction of technological systems: New directions in the sociology and history of technology, Eds Wiebe E. Bijker, Thomas P. Hughes, and Trevor J. Pinch, 111-34. Cambridge, MA and London: MIT Press.

McGray, Douglas. 2002. "Japan's Gross National Cool." *Foreign Policy* 130 (May/June): 44-54.

Miyadai, S. 1994. "Shinjinrui to otaku to wa nan datta no ka." (What was the meaning of the shinjinrui (new breed) and otaku?), *Seifuku Shojotachi no sentaku.* (*The Choice of the School Uniform Girls*). Tokyo: Kodansha.

Mulvey, Laura. 1975. "Visual Pleasure and Narrative Cinema." *Screen* 16 (3): pp. 6-18.

Napier, Susan. 2001. *Anime: From Akira to Princess Mononoke*. New York: Palgrave Macmillan.

———. 2007. *From Impressionism to Anime: Japan as Fantasy and Fan Cult in the Mind of the West*. New York: Palgrave Macmillan.

Nakajima, A. 1991. *Communication Deficiency Syndrome*. Tokyo: Chikuma Shobo.

Nozawa, Shunseke. 2013. "Characterization." *Semiotic Review* (3). http://www.semioticreview.com/pdf/open2013/nozawa_characterization.pdf (Accessed March 30, 2014).

Okada, Toshio. 1996. *Otakugaku nyummon*. (*Introduction to Otakuology*). Tokyo: Ota Shuppan.

Small, Christopher. 1998. *Musiking: the Meaning of Performing and Listening*. Hanover: University Press of New England.

Taneska, Biljana Kochoska. 2009. "Otaku–the Living Force of the Social Media Network." In *Third International Communication Conference--Glocal: Inside Social Media. New York University Skopje, Macedonia*, vol. 15.

Tobin, Joseph. Ed. 2004. *Pikachu's Global Adventure: The Rise and Fall of Pokémon*. Durham: Duke University Press.

Toivonen, Tuuka and Imoto, Yuki. 2012. "Making Sense of Youth Problem." *A Sociology of Japanese Youth: From Returnees to NEETs*. Edited by Roger Goodman, Yuki Imoto and Tuuka Toivonen. New York: Routledge.

The Top Ten List. 2012. "Music Artist that You'd Like to Perform at the London Olympics Opening Ceremonies." http://www.thetoptens.com/singers-perform-london-olympics-opening-ceremonies/hatsune-miku-391821.asp (Accessed January 21, 2014).

Y.I.N. 2011. "ライバルと書いて☆友達と読む." YouTube video. 3:01. Uploaded December 8, 2011. https://www.youtube.com/watch?v=hUkNc8uWHBQ. (Accessed March 30, 2014).

———. 2011. "ライバルと書いて☆友達と読む." Niconico video. 3:00. Uploaded December 7, 2011. http://www.nicovideo.jp/watch/sm16373535. (Accessed March 30, 2014).

Yiu, Wai-Hung, & Chan, Alex Ching-Shing. 2013. ""Kawaii" and "Moe"—Gazes, Geeks (Otaku), and Glocalization of Beautiful Girls (Bishōjo) in Hong Kong Youth Culture." *Positions 21*(4): 853-884.

Zeller, Frank. 2011. "Japan's 3-D Avatar a star in the Real World." *The Australian.* June 15. Retrieved from http://www.lexisnexis.com/hottopics/lnacademic/?verb=sr&csi=244777&sr=HLEAD(Japan%20s%203-D%20avatar%20a%20star%20in%20the%20real%20world)%20and%20date%20is%202011 (Accessed December 14th, 2013)