CHINESE GOVERNMENTAL CENSORSHIP AND ITS INFLUENCE ON UYGHUR
PEOPLE'S LANGUAGE USE IN AN ONLINE SHOW

By

Xiayimaierdan Abudushalamu

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Linguistics – Master of Arts

2023

ABSTRACT

In 2017 there was a massive crackdown on the Muslim population in Xinjiang, China, by the Chinese government through putting 2 million Muslims (mostly Uyghurs) into concentration camps. This thesis investigates how intensified government censorship associated with this sudden political change in Xinjiang might have influenced the language use of characters in a popular Uyghur online show, *Anar Pishti*. It operationalizes the proportion of Chinese Mandarin and the average length of Chinese Mandarin utterances spoken in the show as indicators of Chinese influence on Uyghur language use. Additionally, the project also describes the changing semiotic representations of Chinese Han culture as the index for Chinese government authority in the show. Furthermore, Uyghur-Chinese code-switching is analyzed to see whether the use Chinese Mandarin in the show is the result of long-term language shift in Xinjiang instead of government censorship. Lastly, by seeing Chinese government as an overhearer of the show based on Bell's audience design model (1984), present project also investigates whether the Chinse government overhearer could have stronger effect on the speakers than the addresses and auditors in the show.

In summary, results of quantitative analysis show the proportion of Chinese and mean length of Chinese utterances spoken in the show increase drastically starting from Season 4 released in 2017 to Season 6. The results of qualitative analysis show that there is a clear increase of semiotic representations of Chinese Han culture from Season 1 to Season 6. And the results of code-switching analysis indicate that most of code-switches are "superficial" (e.g., simple lexical insertion) and Uyghur is still the dominant language in Uyghurs mental grammar, thereby ruling out language shift as a potential reason for language use changes. These results also indicate that Chinese government overhearers have stronger influence on speakers' language use than the immediate addressees and auditors in the show.

ACKNOWLEDGEMENTS

I would like to thank Suzanne Wagner, my advisor, for supporting me throughout this whole project, this project wouldn't be completed without her help, her care and her continuous guidance and feedback. I also want to thank Gabriela, Deo, Betsy, and Professor Lin for their valuable input and suggestions for my project, I learned a lot from them, for which I'm very grateful!

I also want to thank the PhD students in the linguistic program, especially my friend Komeil for his long-term support as a friend and as a colleague. We spent countless hours together in East Lansing, talking about linguistics, food and soccer, all of which are precious memories for me.

Last but not least, I would like to thank my beautiful wife Ray, I wouldn't be who I am today as a researcher and as a writer without her love and strong support. We went through all the ups and downs together as a family in the past few years, which were the toughest but also the happiest time in my life, for which I'm forever grateful!

TABLE OF CONTENTS

1. Introduction

1.1. Overview

This thesis aims to show how government censorship influences people's language use in online environments, especially when there is sudden change that drastically increases political pressure on the language speakers. More specifically, this thesis presents an investigation of the effect of Chinese government censorship on Uyghur characters' language use over a period of five years in an online Uyghur show called *Anar Pishti.* The show began airing in 2016, a year before a sudden government repression of the Uyghur people in Xinjiang in 2017. As a casual viewer of the show myself[1], I had noticed that characters seemed to speak Chinese increasingly frequently over time, especially after the 2017 government crackdown on Uyghurs in Xinjiang (BBC News, 2020). The thesis traces the decrease from 2016 to 2020 in linguistic and other symbolic representation of Uyghur culture in *Anar Pishti*. It will show that creators of *Anar Pishti* increasingly design their characters' language for Han Chinese government *overhearers* (Bell, 1984), and that alternative explanations, including language shift (since there is an ongoing shift in Xinjiang toward Chinese) and creative choices made by the creators, are not as convincing.

1.2. Language data from fictional media

Language used in media environment has often been viewed as 'non-natural' and distinguished from conversational language in the community, but at the same time, people have come to realize that language use in TV or cinematic fictions are not as "controlled/standardized" as people

---

[1] I'm a Uyghur and a native speaker of Uyghur. I grew up in Xinjiang, China, and I had watched *Anar Pishti* before I started this project.

previously think since there's much more room for free creation in a TV/fiction production setting (Androutsopoulos, 2014). Although being deemed as "non-authentic" based on the traditional sociolinguistics view (Dynel, 2011), fictional/scripted discourse has come to be viewed as equally important as everyday speech since it can shape a sociolinguistic reality on its own instead of merely reflecting it (Stamou, 2014). It's capable of constructing a particular version of language and the world, in which viewers can understand the language and the world based on their own perspectives, which in return can reflect the viewers' opinions about a certain language or speech community that's depicted by the fictional world (e.g., Xinjiang). This on its own has a deep value in understanding the sociolinguistic environment of a bigger speech community (e.g., China). Besides, the increasing numbers of sociolinguistic analyses of language variation and change in fictional media can attest to the importance of it (e.g., Androutsopoulos, 2012; Bleichenbacher, 2012; Bucholtz, 2011; Dimova, 2012; Planchenault, 2012; Si, 2011; Stamou, 2011; Trotta & Blyahher, 2011; Tsiplakou & Ioannidou, 2012; etc.). Therefore, I view the data collected from *Anar Pishti* as equally important and as valuable as the everyday speech spoken in Xinjiang.

In addition, building on Bell's early work (Bell, 1982,1984), the audience design model has also been more and more been used in sociolinguistic studies of media language, from reality TV shows (Eberhardt & Downs, 2015) to social media (Androutsopolous, 2014). And in terms of the audience design model of linguistic style (Bell, 1984), the Chinese government is an *overhearer:* an audience that is unaddressed, unratified (namely, the speaker does not acknowledge the listener's presence in the speech context), but known to be present in many environments[2]. Apart from overhearer, there are three other types of audience: addressees, who are directly addressed,

---

[2] One could also imagine the Chinese government as an eavesdropper in Bell's terms: unaddressed, unratified, and unknown. However, an eavesdropper is a potential audience, whereas the creators of Anar Pishti will have reasonably assumed that the Chinese government is always or at least very likely monitoring the show. In this sense they are known, and I classify them as overhearers.

known, and ratified; auditors, who are not addressed but known and ratified; eavesdroppers, who are unaddressed, unknown, and unratified. Collectively, studies employing an audience design approach have shown that the audience's effect on a speaker's language use increases when the distance between the audience and the speaker decreases. In other words, an addressee typically has a stronger effect on a speaker than an auditor, and auditor effect is typically stronger than overhearer effect, also similarly, an overhearer typically has stronger effect than an eavesdropper. Therefore, under the audience design model, overhearers are predicted to have very little influence on a speaker's language choices, compared with e.g., the influence of a direct addressee. However, it's precisely because the writers of *Anar Pishti* may be designing the show's script for the Chinese government censors, it's a good type of data for examining Bell's theory with regards to overhearers. Namely, to see whether in an authoritarian society like contemporary China, the overhearing government will have *more* influence on Uyghur speakers than other audience types in the environments like film sets or show sets where they speak directly to co-actors who can be either addressees or auditors.

1.3. Censorship in China and the oppression of minorities

Discrimination against the Uyghur people within China is part of a national racist ideology that is facilitated by censorship. In 2018, the Chinese state television organization CCTV broadcast its New Year's gala. The gala show included a racist skit featuring a Chinese actress as an African woman with a blackface make-up and huge fake buttocks, accompanied by an African man dressed as a monkey. People around the world were horrified and outraged but most Han Chinese—the majority ethnic group in China—were confused and annoyed, some even being angry, saying that it was a gesture of friendship and kindness (BBC, 2018).

This utmost ignorance of racism and race issues in general by the Han Chinese was the product of long-term government censorship by the Chinese Communist Party (CCP) on every aspect of people's life. In the People's Republic of China, citizens are subject to widespread government surveillance and control (Leibold, 2020; Kam & Clarke, 2021) that often manifests itself as censorship, but this is especially true for the people living in ethnic minority areas and provinces (Zeng, 2020; Weller, 2012), like Xinjiang, a minority language community in north-western China.

In Xinjiang, censorship blocked the news about racial conflicts, so that Han Chinese around China believed there was no racial oppression; it distorted Xinjiang history, so that Han Chinese believed they were welcomed by Uyghurs and other ethnic minorities when they entered and colonized Xinjiang in 1949; it controlled education, so that racism was never mentioned in schools, and Han Chinese had no concept of it growing up. However, most importantly, censorship strongly affected ethnic minorities (mostly Uyghurs)' life. Not only did it make them endure the result of ignorance of racism from Han Chinese, but it also took away their rights to practice their religion, to express their culture, and ultimately, to learn and speak their own language. And in recent years, the situation worsened because of a massive crackdown on the Muslim population in Xinjiang by the CCP, who put around 2 million Muslims (mostly Uyghurs) into so called "Re-education" Camps (BBC News, 2020).

Apart from news and education, one of the major targets of Chinese censorship is online media, and although studies of Chinese censorship on online language exist (e.g., King, Pan & Roberts, 2013; Wozniak, 2015; among others), they have not focused on Xinjiang, and they have primarily analyzed written language. Further, there has been little sociolinguistic examination of the influence of such non-ratified *overhearers* (Bell, 1984) on language use in a media context.

4

1.4. *Anar Pishti*

Data for this study are drawn from an online show named *Anar Pishti*, an online mini comedy drama (2016-2020) that takes place primarily in the capital of Xinjiang, Urumqi. The show's name translates as "Pomegranate is ripe", which is related to a frequently used government slogan you can see everywhere right now in Xinjiang: "Let's unite tightly like pomegranate seeds." The implicit message from the government is that every ethnic group in Xinjiang should unite tightly together like pomegranate seeds and create so called "racial harmony" in Xinjiang. The first season was released in 2016 on the Chinese online video platform, YouKu. The subsequent 5 seasons were released over the next 4 years, concluding with Season 6 in 2020[1]. *Anar Pishti* features a fixed cast of 5 actors (2 females, 3 males, Appendix 1), and scenes that are not narratively interconnected. The sketches reflect the broader characteristics of ethnic minorities (mainly Uyghurs) and online humor culture. Each season consists of 15 to 20 episodes, and each episode is composed of at least five scenes. Each scene is 1-3 minutes long, and each scene has at least one humorous moment. Each actor has his/her fixed identity (e.g., same fictional name) throughout all seasons but plays different characters in different scenes of each episode. The main language used in the show is Uyghur and some very little Kazak (mainly used by one single character), and they are all supplemented with Chinese subtitles.

Because of its youth elements and incorporation of up-to-date social topics in the Xinjiang region, *Anar Pishti* quickly became popular among Uyghur and some Chinese audiences; nevertheless, at the same time, because of its use of strong Uyghur elements and Uyghur language, it has become a likely target of the Chinese government's censorship since 2017 (release year of Season 4) when the government's crackdown on the Uyghur Muslim population began (BBC

News, 2020). This thesis will look at linguistic evidence for *Anar Pishti*'s creators' reactions to the censorship over time between 2016 and 2020.

1.5. Outline of the thesis

Chapter 2 will serve as an introduction to Xinjiang's recent social and linguistic history. This provides necessary context for a summary of the media censorship situation in China and Xinjiang.

In Chapter 3 I lay out my background assumptions, my research questions, and the hypotheses that I will test in the next two chapters.

Chapter 4 will show the results of two broad-stroke analyses of how the Uyghur and Chinese languages are represented on *Anar Pishti* over time. I begin with a qualitative analysis of the proportion of Chinese non-linguistic symbolic content on the show over time. Next, I take a quantitative look at the relative proportional use of these two languages, season by season.

In Chapter 5, I examine Chinese language use on *Anar Pishti* in more detail. The frequency and type of code-switching is reported, and I provide an audience design analysis of language use to (direct) addressees and (co-present, but not directly addressed) auditors by characters on the show.

Finally, Chapter 6 discusses the results of the prior chapters, draws some conclusions, and points to future directions.

2. Xinjiang and censorship

2.1. Xinjiang demographics and history

Since *Anar Pishti* is a Uyghur show that primarily takes place in Xinjiang, it's important to know the history and demographics of Xinjiang so that one can understand the causes behind the intensified Chinese government in Xinjiang that led to the change of language use in *Anar Pishti*. In this section, I draw principally from the history of Xinjiang provided in Kamberi (2015). The Xinjiang region is home to the Uyghur people, and accounts for about one sixth of China's territory, as the largest Autonomous Region of and province of China. As shown in Figure 1, The Uyghur region includes a great portion of Central Asia, from the northeast to the southwest; it borders Mongolia, Russia, Kazakhstan, Kyrgyzstan, Tajikistan, Afghanistan, Pakistan, Tibet, and India.

Figure 1. Uyghur Autonomous Region of Xinjiang (Encyclopedia Britannica, 2023)

The Xinjiang region was an independent state in different historical periods and was not fully controlled by the rest of China until the Qing invasion of 1877. The last Uyghur republic, established in 1944, was strongly supported by the Soviet Union. From 1946 to 1949, Russia and China attempted many governmental structural reforms in Xinjiang. During these reforms, both Russian and Chinese government representatives promised the Uyghurs again and again that the presence of the Chinese army in Xinjiang was intended to promote democratization, free elections, and greater autonomy, to help build the new Xinjiang, even to provide for the eventual independence of Xinjiang. None of these promises were fulfilled. After 1949, it was fully controlled by the Chinese Communist Party and has been part of the People's Republic of China since then.

Moreover, historically, because of Xinjiang's unique geographic location that connects many countries from west to east, Xinjiang has been a home for many ethnic groups, including Uyghur, Han, Hui, Kazakh, Kirghiz, Mongols, Tajiks, Uzbeks, Russians, etc. Table 1 below shows the population and percentage information of each ethnic group in Xinjiang based on Chinese government's 2018 demographic data ("Xinjiang," 2018).

Table 1 Number and proportion of ethnic groups in Xinjiang in 2018

Nationality	Population	Percentage
Uyghur	11,678,646	51.145%
Han	7,857,370	34.410%
Kazakh	1,574,930	6.897%
Hui	1,015,700	4.448%
Kirghiz	208,346	0.912%

Table 1 (cont'd)

Mongols	178,993	0.784%
Tajiks	51,355	0.225%
Xibe	42,772	0.187%
Manchu	27,372	0.120%
Tujia	N/A	N/A
Uzbek	19,652	0.086%
Russian	11,604	0.051%
Miao	N/A	N/A
Daur	6,793	0.030%
Tibetan	N/A	N/A
Zhuang	N/A	N/A
Tatar	5,019	0.022%
Salar	N/A	N/A
Other	156,024	0.683%

The ethnic demographics of Xinjiang have changed radically since the mid-twentieth century. In 1949, Uyghurs accounted for more than 90 percent of the region's population, while Han Chinese accounted for only 5 percent. But in the past 20 years alone, Han Chinese in the Xinjiang region increased from 30% to about 40% of the total population. This huge amount of increase led to a series of conflicts between Han Chinese and Uyghurs in the region since 1949, especially in the 1990s-2010s, and among these conflicts, major conflicts include the Baren Township Riot in 1990 (Davis, 2008), Urumqi Bus Bombings in 1997 (Dillon, 2003), Ghulja Incident in 1997 (Rodríguez, 2013), July 2009 Urumqi Riots (BBC News, 2009), Hotan Attack (The Economist, 2011), April 2014 Ürümqi Attack (BBC News, 2014), May 2014 Ürümqi Attack (The Guardian,

2014). Finally, in 2017, it culminated in a massive crackdown on the Muslim population in Xinjiang by CCP through putting around 2 million Muslims (mostly Uyghurs) into so called "Re-education" Camps, which are basically concentration camps, as a way to eliminate the culture and language of Uyghur people so that they can "solve" the conflict (BBC News, 2020).

2.2. Uyghur and Chinese Mandarin in Xinjiang

There are 49 languages and dialects spoken in Xinjiang as of 2020, including Uyghur, Chinese Mandarin, Kazakh, Uzbek, Russian, Kyrgyz, Oirat, etc. ("Xinjiang," 2020). However, it has to be noted that, based on my own experience of living in Xinjiang, it's required that all ethnic minority languages speakers have to learn and speak Chinese Mandarin, but there is no such a rule for Han Chinese in Xinjiang, therefore, almost no Han Chinese in Xinjiang can speak Uyghur or other minority languages, but almost all Uyghurs and other ethnic minorities can speak Chinese Mandarin to a certain degree (from somewhat fluent to native-like) based on their education background. Since the present study focuses on the two most widely-spoken languages, Uyghur and Chinese Mandarin, as they are used on the show *Anar Pishti*, I briefly introduce these two languages and their social contexts here.

The Uyghur and Mandarin languages are different enough that one person can notice the switch between these two languages (see Chapter 5, where this is relevant to the analysis of code-switching). Uyghur is a Turkic language with 8 to 11 million speakers, spoken primarily by Uyghur people in the Xinjiang Uyghur Autonomous Region of Western China and remaining Uyghur people residing overseas. Uyghur belongs to the Karluk branch of the Turkic language family. Like many other Turkic languages, Uyghur displays agglutination and vowel harmony, lacks noun

classes or grammatical gender, and it has subject–object–verb word order ("Uyghur language," 2023).

Similar to Uyghur, the majority of Chinese Mandarin native speakers reside in China, with the population of around 920 million. Chinese Mandarin is the official language of China, and it belongs to the Sino-Tibetan language family. It also lacks noun classes and gender, but unlike Uyghur, it has subject-verb-object word order. Also, it is considered as isolating language, so unlike Uyghur, it has much simpler morphological system in terms of morphological inflection or affixes. Chinese Mandarin is also a tonal language, with four primary tones (Ethnologue, 2020).

Uyghur is an official language of the Xinjiang Uyghur Autonomous Region (Xinjiang) and is widely used in both social and official spheres, as well as in print, radio, and television in Xinjiang. However, even though both Uyghur and Chinese Mandarin are considered as official languages, since 1949, Xinjiang has been undergoing language shift from Uyghur to Chinese Mandarin. And language policies implemented by Chinese government in education system in Xinjiang over the years since 1949 is one of the leading causes of this language shift, and the next section will introduce these policy changes in detail.

2.2.1. Language shift in Xinjiang education

The first wave of encounters between Uyghurs and Hans happened since very early age in the history, it can go back to as early as Western Han Dynasty (202 BC–9 AD) (Dai & Dong, 2001), but it was not until Qing Dynasty there were Han schools being built in the region; however, those schools were mainly teaching Chinese Han culture (e.g., Confucianism) (Chen, 2008). This situation lasted until CCP took over the control of China in 1949. Before 1949, in almost all schools in Xinjiang, Uyghur was used as the medium of instruction (MoI) in classrooms, which also lasted

after 1949, until 1980s; however, Chinese Mandarin was already widely promoted in all schools after 1949. And during 1980s, mastery of both Uyghurs and Chinese Mandarin, was emphasized in government documents (Wang, 2012), which led to the emergence of bilingual schools.

Starting from 1990s, there were many Han schools built in Xinjiang, in which Chinese Mandarin is used as the only MoI. Those schools not only accept local Chinese Han students, they also accept Uyghur students, separating from ethnic minority bilingual schools, which only accept Uyghur students. This separate schooling lasted until 2000s; however, it is shown that 80% of Uyghurs in Xinjiang work in the most poorly paid working positions, while Han Chinese usually occupies most of the best paid jobs in the region, mainly because Chinese Mandarin is used as the only working language in most of those working environments (Cote, 2015). In response to this situation, Xinjiang government started to enforce Chinese Mandarin as MoI in most of the bilingual schools in Xinjiang. In 2004, it was proposed by the local government that all courses in Urumqi senior high schools should use Mandarin as MoI, except for the Uyghur language course, and this decision further extended to pre-school education in 2005 (Ma, 2009). In 2010, this mode of bilingual education was further planned by the government to cover 90% of the schools in Xinjiang by 2020 and 85% of the pre-schools in Xinjiang by 2015 according to "Outline of Mid- and Long-term Educational Reforms and Developmental Plan for Xinjiang Uyghur Autonomous Region (2010–2020)".

As can be seen the history of Xinjiang language policy above, since 1980s the status of Uyghur language in the bilingual education system had changed from being the MoI for all subjects in schools to being gradually marginalized and becoming a subject course (like P.E or music class) in almost all schools, and finally giving way to Chinese Mandarin as MoI. This kind of rapid change clearly reflects culture assimilation, rather than the Xinjiang government's original goal of

so called 'Minhan jiantong', the mastery of both the ethnic minority and Han languages, written in the 1982 document "Suggestions for Strengthening Han-language Education in Ethnic Schools". It therefore lays the groundwork for long-term language shift in Xinjiang from Uyghur to Chinese Mandarin.

These changes of language policy also had influences in Xinjiang online media environment, including online TV shows, and these influences are normally manifested through governmental censorship, which will be the topic for next section.

2.3. Censorship in China

Censorship is generally defined as "the suppression of speech, public communication, or other information, on the basis that such material is considered objectionable, harmful, sensitive, or 'inconvenient'" by governments, private institutions, or any other authorities ("Censorship," 2020). In government hands it is a powerful tool that can influence many aspects of people's lives, including their language use (Wozniak, 2015; Emery, 2002; Cowlishaw. 2000; King, Pan & Roberts, 2013). In this particular project, for obvious reasons, I will be focusing on Chinese Internet censorship. There are three major ways implemented to censor people's expression on the Chinese Internet (King, Pan & Roberts, 2013):

1) "the Great Firewall of China": a computer system that blocks almost all foreign big websites all over the country, especially those big Internet firms like Facebook, Google, Twitter, etc.

2) "keyword blocking": also a computer system required to be implemented in all Chinese websites, which automatically censors any posted sensitive words or texts that include such words. For instance, The Tiananmen Incident ("天安门事件"), Falun Gong ("法轮功"), freedom ("自由"), etc.

3) a large number of Internet police ("网警") hired by central and local governments and other types of censors hired by Internet firms: they manually detect and remove those texts and posts that are hard to be detected by those two methods above. It's been proven to be very effective and systematic (Chen & Ang, 2011).

Because *Anar Pishti* is an online show, it is subject both to the internet censorship practices just described, and to Chinese censorship of media more generally. According to the CCP's National Radio and Television Administration (NRTA)'s "Regulations on the Administration of Internet Audio-Visual Program Services" (NRTA, 2007), NRTA is responsible for "supervising" and "managing" internet audio-visual programs in China. They also stated that local Radio and Television Administrations are responsible for the local programs. In 2014, they also post a notice called "Notice of Strengthening Regulations on Internet Audio-Visual Programs like Online Shows and Micro-Movies" (NRTA, 2014), in which they clearly state that any companies/organizations that produce online shows or micro-movies required to hire "compliance audits who have passed the national or provincial Internet Audio-Visual Program Association's training." Moreover, each production company/organization has to obtain the "License of Dissemination of Internet Information and Audio-Visual Programs" issued by NRTA. All of these regulations clearly indicate that *Anar Pishti* is under CCP's censorship.

2.3.1. Censorship in Xinjiang

Plus, Xinjiang's unique political situation only increases the intensity of the censorship. For instance, in 2017, March 29, Chinese government passed the Xinjiang Uyghur Autonomous Region Regulation on De-extremification, in which they clearly state that those media/online websites will be forced to close or shut down if they contain anything that is considered by the government as "extremified negative content" ("Xinjiang Uyghur Autonomous Region Regulation on De-extremification," 2017). And according to the "Learning and Identifying 75 Religious Extreme Activities in Parts of Xinjiang" released by Xinjiang government in 2017 (Department of Xinjiang United Front, 2017), "resisting modernist clothing", "preventing children from learning Chinese Mandarin" and "defaming bilingual education" are among those "Religious Extreme Activities".

Although the regulation doesn't specify how much percent of the dialogues in the show have to be in Chinese Mandarin or what kind of Chinese cultures should be incorporated in the show, I believe it is exactly this type of "vagueness" that makes it possible for the government to shut down any online media that, they think, are "extremified" without any clear explanation. "Vagueness" is one salient trait of CCP's censorship, which promotes people working in creative businesses to actively self-sensor by constantly "guessing" where the "redline" is.

2.3.2. Censors as overhearers

Anar Pishti's creators' self-censorship is the direct reflection of their awareness of the existence of Chinese censors as overhearers. Even though previous studies employing an audience design approach have shown that an addressee typically has stronger effect on a speaker than an auditor, and auditor effect is typically stronger than overhearer effect, Bell (1984) points out that

this is not always the case: Sometimes the effect of a further audience can override the effect of a closer one. For instance, Gal (1979,p.124) found out that in Oberwart, Austria, Hungarian-German bilingual speakers always switched from Hungarian to German as long as there was a monolingual German speaker as an auditor. Nevertheless, to my knowledge, no study has yet observed that the effect of an overhearer is greater than the effect of an addressee or auditor. Bell shows that ignoring overhears or eavesdroppers has consequences by mentioning Jimmy Carter's interview with Playboy magazine in 1976 (Solomon, 1978, discussed in Bell, 1984, p.177). The interview, in which Carter used some risqué language, became a blot on Carter's presidential campaign, because he failed to anticipate the effect of overhearers and eavesdroppers, who were also his supporters but disliked his behavior during the interview. However, when the consequences are severe enough, e.g., ignoring the Chinese government censors can lead to canceling of TV shows or movies, there's a reason to believe that the effect of overhears or eavesdroppers could override the effect of addressees and auditors. And because of the awareness of those sever consequences, the prediction follows that even the possibility of censorship will lead *Anar Pishti*'s creators to change the show as the potential threat of adverse consequences even increases after 2017. They can't create an impossible to penetrate code (as in Wozniak's 2015 paper where he describes how Chinese netizens avoid Chinese online censorship by exploiting Chinese phonology, morphology, and orthography so that they could create an impenetrable-to-outsiders code) but we will see evidence of them responding linguistically to censorship.

2.4. Summary

In summary, it can be seen that Chinese government's deliberate changings to Xinjiang's demography and language policies after 1949 has led to many conflicts between Uyghurs and Han

Chinese in the region , which culminated in a massive crackdown on the Muslims, mostly Uyghurs, in 2017; meanwhile, this crackdown made the Chinese government censorship that already existed for many years and that were already intense especially in ethnic minority regions like Xinjiang even more intense. And based on Bell's audience design model and how intensified the censorship has become in Xinjiang after 2017, it is reasonable to believe that *Anar Pishti*'s creators has gone through self-censorship over the years linguistically because of the effect of Chinese censors as the overhearers of the dialogues in the show. To test this prediction, the next chapter will lay out research questions and the relevant hypothesis.

3. Research question and hypothesis

3.1. Background assumptions

As mentioned in the previous chapter, government censorship can influence people's language use, including their use of languages in movies and TV shows. One relevant example is Korean-Japanese code-switching in Korean films during World War Two when Korea was under the control of the Japanese Empire (Kwon, 2013). During that period, Korean film producers had to use Korean-Japanese code-switching in their films to "linguistically, culturally, and politically align itself with the wartime empire." (Kwon, 2013, p.10). Therefore, similarly, I assume that *Anar Pishti's* creators make use of Chinese Mandarin and Chinese cultural elements in their show to align themselves linguistically, culturally, and politically with Chinese government. I also assume that *Anar Pishti's* creators must have been aware that Internet police could be surveilling their show at any time, and that surveillance likely increased after 2017. In fact, *Anar Pishti*'s popularity gives it 'collective action potential': A potential that the Chinese government would not want to see realized. According to King, Pan and Roberts (2013), Chinese censorship of the internet is in line with "collective action potential" theory (King, Pan & Roberts, 2013), which posits that the target of censorship is those who express their thought and ideas online collectively, which has the potential to create collective action outside the internet, like protests, regardless of whether the posters have expressed being opposed to or supportive of the government. Given the size of the audience that potentially could be reached by the *Anar Pishti s*how, and since the Uyghur language is the most important aspect of a collective identity recently targeted by the CCP, I don't see any reason to reject the possibility that *Anar Pishti* could be a potential target of Chinese internet censorship, according to the "collective action potential" theory. For this reason, I assume that

Anar Pishti is subject to a fairly high degree of Chinese government censorship, and that this government pressure has increased since 2017.

Further, I take it for granted that the creators of *Anar Pishti* made the same assumption about the degree of surveillance to which the show was likely subjected. Following the predictions of Bell's audience design model (1984), I anticipate that the creators will design the language of the show – in some empirically observable way – for the censors as possible overhearers.

The thesis will address the following three research questions:

1.   Will the amount of Chinese cultural and linguistic representation increase in *Anar Pishti* after 2017?

2.   Will the integration of Chinese language into *Anar Pishti* become more complex over time?

3.   Will *Anar Pishti* characters increasingly design their speech for government overhearers rather than other characters in the show?

3.2. Hypotheses

3.2.1. H1: The amount of Chinese cultural and linguistic representation will increase
       in *Anar Pishti* after 2017.

Xinjiang is a bilingual community, and *Anar Pishti* includes Chinese language throughout its run. But I expect Chinese language and Han Chinese ethnic representation to be more evident in Seasons 4-6 than in Seasons 1-3. This is because I anticipate that the show's creators will feel much more pressure (including implicit pressure, as discussed in Chapter 2 above) to dilute *Anar Pishti's* strongly Uyghur ethnic identity. The timing of an increase in Chinese Han linguistic representation after 2017 would be suggestive of a government censorship by the show's creators,

rather than merely a reflection of ongoing community language shift in Xinjiang. In Chapter 4, I will test this hypothesis by calculating the proportion of Chinese language used on the show season by season. I will also present qualitative evidence from observations of semiotic non-linguistic representation of Han culture, in the form of clothing and other symbolic markers.

3.2.2. H2: Integration of Chinese language into *Anar Pishti* will not become more complex over time.

Even if Hypothesis 1 is confirmed, the correlation of a big rise after 2017 in the amount of Chinese culture and language on *Anar Pishti* with the government crackdown could just be spurious. After all, Xinjiang is undergoing language shift away from Uyghur and towards Chinese, which could potentially explain any increase in Chinese language use on *Anar Pishti.* In addition, political regime change is an initiator of language shift (Romaine, 1994). For example, Willemyns (1997, citing De Jonghe, 1967) describes how a deliberate Frenchification policy (e.g., obligatory use of French in the school system, in local administration and in all official documents) was imposed on towns near the border with the Netherlands after the French Revolution in 1789. The Dutch language faced almost complete extinction in that region as a result. As I explained in Chapter 2, the situation in Xinjiang is very similar, because language shift due to long-term pressure on Uyghur from Chinese may recently have accelerated because of sudden political change in 2017. Therefore, peaks in the trend line might be random. However, we would expect language shift to be accompanied by the production of increasingly complex linguistic structures by L2 speakers. David (1996), for example, argues that language shift cannot be very advanced if, when speakers code-switch from their L1 (minority, 'ethnic') language to their L2, their switches are made up of only "minimal linguistic items". Independently of Hypothesis 1, therefore, I

anticipate that Chinese language use on *Anar Pishti* will be linguistically superficial and will remain so over the seasons. In Chapter 5, I will operationalize "superficial" and "complex" in terms of types of code-switching and Myers-Scotton's Matrix Language Frame (MLF) model (1997). MLF was created to explain intrasentential code-switching, and within a typical intrasentential code-switching sentence, the dominant language is considered as the Matrix Language (ML) and the other one is called Embedded Language (EL)[3]. In an intrasentential code-switching sentence, ML consists of System Morphemes (e.g., function words, inflectional morphemes, etc.) and EL consists of Content Morphemes (e.g., nouns, verbs, adjectives, and some prepositions). Building upon this model, I define the integration of Chinese Mandarin in a Uyghur-Chinese code-switching sentence as "complex" when System Morphemes in the sentence are from Chinese Mandarin. And for my hypothesis to be confirmed, within intrasentential code-switches, System Morphemes will still come from Uyghur, not from Chinese Mandarin, over the seasons of *Anar Pishti.*

3.2.3. H3: *Anar Pishti* characters will increasingly design their speech for government overhearers rather than other characters in the show.

Even if the amount (H1) and type (H2) of Chinese language on *Anar Pishti* does not implicate a post-2017 reaction to government censorship, I anticipate that government influence will nonetheless be detectable. One way to uncover it is to see if Uyghur characters on *Anar Pishti* speak Chinese with each other (i.e. notwithstanding their speech to Han Chinese addressees) more frequently over time? And do they do so even when there are no Han Chinese characters co-present

---

[3] Here "dominant language" refers to the language that is dominant in the mental grammar of the multilingual individual, not socially dominant.

in the scene? If yes, then we would be forced to conclude that *Anar Pishti's* creators had inserted Chinese dialogue in non-natural ways, most likely with the potentially watching government in mind. Put another way, I predict that on *Anar Pishti,* government overhearers will have more linguistic influence on Uyghur characters than addresses and auditors – in contravention of the predictions made by Bell's audience design model.

4. Analysis: Representation of Chinese language and culture

In this chapter, I analyze the amount of semiotic representation of Chinese culture and the amount of Chinese Mandarin used by the characters in different seasons of the show to address the Hypothesis 1.

4.1. Semiotic representation of Chinese culture

In this particular section, I'm looking at the amount of semiotic representation of Chinese culture in different seasons of the show to see whether the amount of Chinese cultural and linguistic representation will increase in Anar Pishti after 2017, as claimed in Hypothesis 1.

4.1.1. Data and method

I watched every episode of all 6 seasons of the show. As I did so, I recorded my observations of characters' language content, behavior, and dress, as well as other aspects of the show that struck me as symbolic of Chinese – and not Uyghur – culture. I noted the season and episode for every observation. I include just a few representative observations here. In what follows, I refer to Seasons 1, 2 and 3 as "early" seasons, i.e., pre-2017 government crackdown, and Seasons 4, 5 and 6 as "late".

4.1.2. Observations

4.1.2.1. Clothing

Female characters start to wear short pants starting from Season 4, as shown in the comparison between Season 2 and 6 in Figure 2 below. I interpret this as the penetration of Han modern culture

since Uyghur culture tends to be more conservative when it comes to women's clothes because of the influence from Muslim religion. Also, since Han culture has been increasingly in contact with Uyghur culture over the generations, one wouldn't necessarily expect to see this abrupt shift in Season 4 in terms of clothing, so this could again suggest that post-2017 political pressure on the show is responsible for this semiotic change.

Figure 2 Clothing change from earlier seasons to later seasons

Season 2   Season 6

4.1.2.2. Education

Classroom scenes changed from Uyghur school classrooms to Chinese school classrooms in later seasons, as an indicator of the change of the education system, namely, Chinese-Uyghur bilingual schools being replaced by Chinese schools in Xinjiang. However, the shift from Uyghur school classrooms to Chinese ones happened abruptly in the show, not gradually as it occurred in the history of Xinjiang, which indicates that there is government pressure on the show after 2017 and the show is not simply reflecting language shift in Xinjiang.

4.1.2.3. Use of religious terms

In Season 1, there are religious terms used in characters' speech, for instance, in Season 1 Episode 2 at 10:05, Ximu said "hudayim boyrisa", which means "if god permits." This type of religious term never appears in later seasons.

In summary, the observations support Hypothesis 1, that the amount of Chinese cultural representation increases in *Anar Pishti* after 2017. And since language/culture shift in Xinjiang already started way before 2017, the observations also support Hypothesis 1 that the timing of the increase in Chinese Han cultural semiotic representation after 2017 would be suggestive of a government censorship by the show's creators, not to ongoing language/culture shift.

4.2. Proportion of Chinese spoken per season

In this section, I'm looking at the proportion of Chinese spoken per season to see whether the amount of Chinese linguistic representation will increase in *Anar Pishti* after 2017, as claimed in Hypothesis 1.

4.2.1. Data and method

The analysis of employs audio from the first six episodes of *Anar Pishti* for each of seasons 1 to 6. The mp4 files were ported to ELAN 6.4 (Max Planck Institute for Psycholinguistics, 2022), an application that allows for multi-tier annotation of time-aligned audio. I manually identified and segmented all utterances in Chinese on a *Chinese_words* tier. Chinese loan words were not included. One additional tier was used for the manual identification and segmentation of *Silence and adds*. *Silence* was defined as a time length longer than 0.5s where there is no human speech that are relevant to the content of the show being produced. Therefore, this type of speech also includes all the commercial advertisements.

25

The segmented and annotated files were exported to Excel as .csv files, where the length of each *Chinese, Silence,* and *Non-speech/Adds* segment was automatically reported and could easily be summed. To simplify the analysis, *Silence* and *Non-speech* were treated as a single non-speech category. I first calculated the minutes of speech per season by subtracting the sum of all non-speech segments from the total run time across the season's episodes. For example, Season 4 has a run time of 5887.068 seconds; of this, 3186.918 seconds were Silence. Therefore, speech accounted for 2700.15 seconds of the season. I then calculated the total amount in seconds constituted by all Chinese utterances in each season as a proportion of the total speech in that season. For Season 4, Chinese constituted 10.2 % (276.58 seconds) of the total speech in the season.

4.2.2. Results

The amount and proportion of Chinese spoken by all characters in the dataset is shown in Table 2 and Figure 3. One can see in the right-hand column of Table 2 and in Figure 3 that there is no notable difference between Season 1, 2 and 3 in the proportion of Chinese used in the show, which is very low at around 1-2% of all the speech in these seasons. However, there's a big increase in the proportion of Chinese spoken starting from Season 4, when the proportion leaps to approximately 10%. This finding is in line with Hypothesis 1, namely, the amount of Chinese Mandarin produced in the show will increase starting from Season 4 (which is released in 2017).

Table 2 Chinese Spoken Per Season of *Anar Pishti*, as a Proportion of all Speech in Seconds

(6 episodes per season only)

Season	Total Airtime	Silence	Airtime Without Silence	Chinese	Chinese/Without Silence (%)

Table 2 (cont'd)

1	4500.03	2690.07	1809.96	13.47	0.7
2	5860.089	3719.989	2140.1	36.86	1.7
3	5875.342	3218.422	2656.92	36.04	1.4
4	5887.068	3186.918	2700.15	276.58	10.2
5	6000.778	2984.798	3015.98	482.07	16.0
6	6831.204	2451.004	4380.2	1083.4	24.7

Figure 3 Proportion of Chinese Produced in the First Six Episodes of Six Seasons
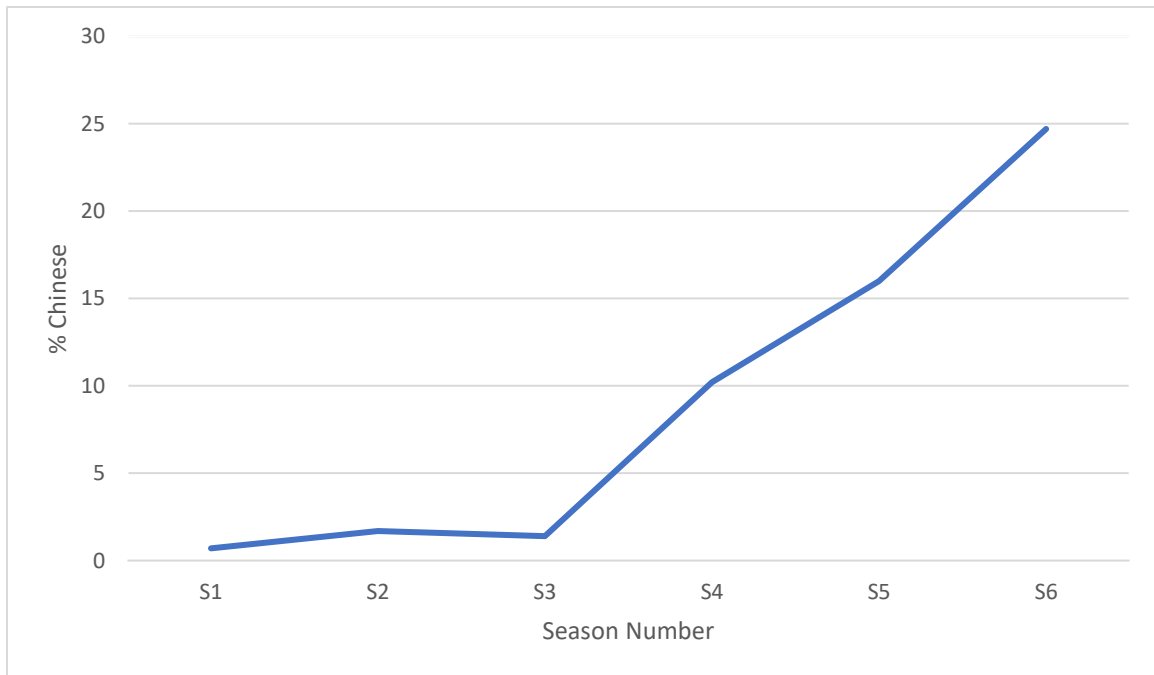
4.3. Length of Chinese utterances per season

One possible alternative explanation for the dramatic increase in the proportion of Chinese in Season 4 onwards is that the characters just happen to use certain Chinese words more frequently when speaking Uyghur.

4.3.1. Data and method

In order to rule out this possibility, I tested whether the mean length of utterances in Chinese increased from Season 4 onwards. I employed the same dataset as described above for the analysis of the proportion of Chinese. I operationalized "utterance" based on the time length of silence. If there's a silence longer than 0.5s[4], it marks the beginning or ending of an utterance in a speech. I then calculated the average length of Chinese utterances produced in each season by dividing the overall time length in seconds of Chinese utterances by the number of utterances in the whole season.

4.3.2. Results

Table 3 and Figure 4 below show the result of the calculation.

Table 3 Average Length of Chinese Utterances in Seconds, by Each Season

Season	Number of Utterances	Chinese (average length)
S1	14	0.39
S2	48	0.48
S3	72	0.54
S4	158	1.22
S5	219	1.60
S6	346	3.42

---

[4] There seems to be a unified time length for the boundaries of utterances in language acquisition, which is 2s (e.g., Language Development Project 2022), but since adult speech is faster than child speech, I use 0.5s as the boundary.

Figure 4 Average Time Length of Chinese Utterances in Each of Six Seasons
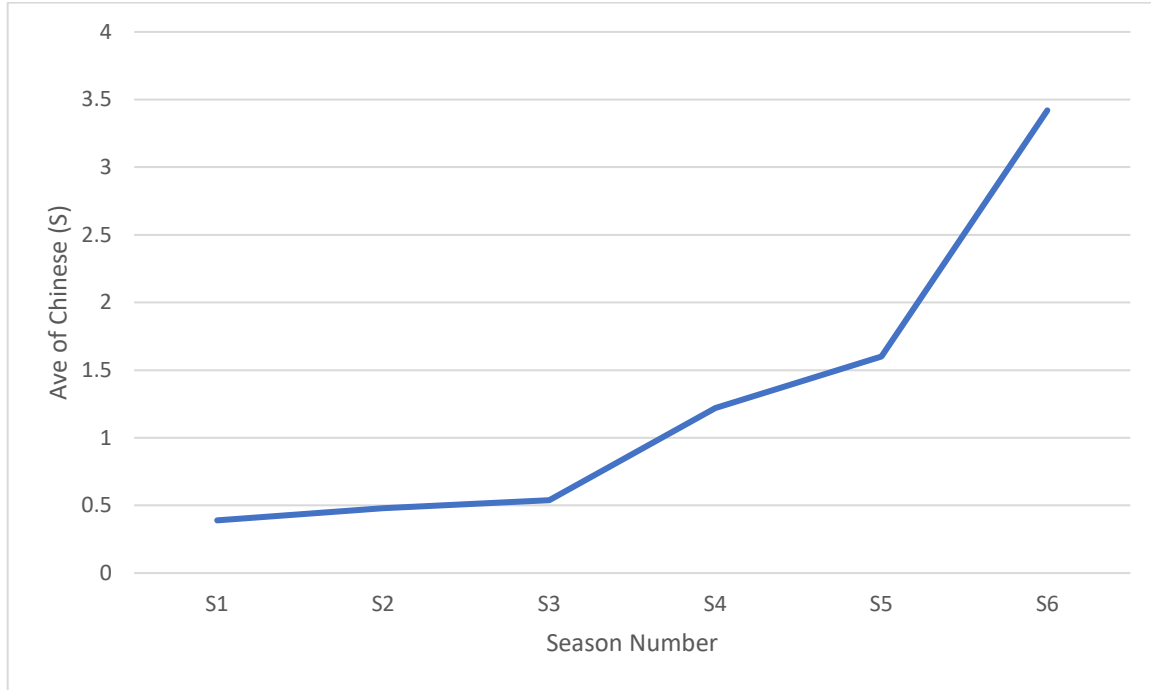
From Figure 4 one can see that, similar to Figure 3, there is no notable difference among first three seasons in terms of the average time length of Chinese spoken. However, there is a notable increase from Season 3 to Season 4, and from Season 4 to 6. This refutes the possibility, floated above, that the characters just happen to use certain Chinese words more frequently when speaking Uyghur. Rather, utterances in Chinese become longer over time, so they are not merely single word insertions.

4.4. Summary

In summary, based on the results of the qualitative and quantitative analysis of the amount of Chinese cultural and linguistic representation in six seasons of the show, it can be seen both of these measurements increase in *Anar Pishti* after 2017, which supports Hypothesis 1. Furthermore,

the results of the analysis of the mean length of Chinese utterances in each season suggests that the Chinese utterances produced by the characters in the show become longer over time and they are not merely single word insertions. In the next chapter, in order to see what's really happening with the increasing length of Chinese utterances, I'm going to move to an utterance-level analysis to include code-switching, and audience design.

5. Analysis: Audience effects on Uyghur and Chinese language use

Another semiotic representation that can reflect the influence of Chinese government censorship on the language use in the show because of the sudden political change is Code-Switching (CS). Therefore, in this chapter, I analyze the amount of CS used in each season, the type and direction CS used in each season to test Hypothesis 2, namely, to see whether integration of Chinese language into *Anar Pishti* become more complex over time.

In addition, it has to be noted that *Anar Pishti*'s creators must write dialogues with multiple audiences in mind. First, characters speak to other characters (addressees). Second, characters design speech for co-present, non-directly addressed characters in the same scene (auditors). Third, the online viewing audience constitutes a kind of auditor for whom the *Anar Pishti* creators are designing the language of the show. Fourth, the Chinese government is a potentially unratified audience (overhearers) for whom the language of the show must also be designed. Therefore, I am also looking at the Audience effects on the Chinese language use of the characters, particularly, Uyghur characters in the show in each season to test my Hypothesis 3, namely, to see whether *Anar Pishti* (Uyghur) characters will increasingly design their speech for government overhearers rather than other characters in the show.

In this chapter, section 5.1 begins with an analysis of CS, where we can see how characters use language with direct addressees and whether CS utterances become more structurally complex when the mean length of Chinese utterances increases as shown in the previous chapter. Section 5.2 looks at the proportion of Chinese used by Uyghur characters in each season and whether that is influenced by the presence of Han Chinese addressees and auditors or not.

5.1. Code-switching

Code-Switching (or Code-Mixing (CM))[5] is generally defined as using two or more codes, i.e., languages or language varieties within an exchange by one single speaker across turns, within a turn or within a single utterance (e.g., Auer, 2013; Milroy & Muysken, 1995; Bullock & Toribio, 2009). Where the code-switching occurs in an unchanged setting, this has often been defined as 'conversational code-switching', in contrast to 'situational code-switching' which occurs when situational factors change (e.g., change of physical setting/location: office vs. home) (Blom & Gumperz, 1972). I will include both types of code-switching in my analysis.

CS is often related to speaker's social background. For instance, Myers-Scotton (1993b) states that bi/multilingual speakers normally assess their social environment based on various social factors (e.g., ethnic identity, education background, gender, etc.) before they choose whether or not to switch to another language. Eventually, they will make the most unmarked linguistic choice based on their assessment. Relevant to the present thesis, studies of language use and language attitudes in Xinjiang (Cao & Wang, 2009; Sun, 2012; Elterish, 2016) have shown that most Uyghurs, regardless of their age, gender, or education background, speak Uyghur with their family members, with their friends, or peers, in public, at home, or at work, and most of the time. Language choice varies slightly depending on the age of the addressee, namely, younger Uyghurs use slightly more Chinese overall than older people. They also have mostly positive attitude towards people who speak Uyghur. However, the situation changes drastically once the authoritarian element comes into play, namely, the probability of speaking Uyghur significantly

---

[5] Muysken& Muysken (2000) separate CS from CM, stating that switching happens inter-sententially (at the sentence or clause boundary), while mixing occurs intra-sententially (within the sentence or clause). However, as Bullock and Toribio (2009) indicate, the difference can be just the degree of mixing, where CM has greater degree of mixing than CS. Therefore, in this study, I will consider both of them as one single CS, instead of two different phenomena.

decreases, and the probability of speaking Chinese significantly increases when people's occupation is related to the government or when they are dealing with the government (Sun, 2012). A similar pattern seems to exist among Uyghur students when a Chinese teacher (regardless of their ethnicity) is present. Therefore, this ultimately shows the correlation between speaking Chinese Mandarin and the governmental authority as well as the correlation between speaking Uyghur and authentic Uyghur identity in Xinjiang.

In the remaining sections of this chapter, I focus on the analysis of CS data in each season, to see whether the results of the analysis support Hypothesis 2, that integration of Chinese language into *Anar Pishti* become more complex over time.

5.1.1. Data and method

5.1.1.1.   Data and dependent variable

I operationalized CS as instances of both conversational and situational code-switching since both happen in the show. Since, as stated in an earlier chapter, Uyghur and Chinese Mandarin are linguistically highly dissimilar, it was usually quite straightforward to determine when a code-switch had occurred. However, lexical borrowing, particularly in a speech community like Xinjiang where there is a high degree of contact between Uyghur and Chinese Mandarin, presented a more difficult challenge. As Bullock & Toribio (2009, p.5) point out, the term 'borrowing' is normally associated with two different types of forms. One is assimilated lexical borrowing, which involves the morphological and phonological integration of the borrowed words, like Japanese word *basubaru* from English "baseball," which is fully established in the monolingual Japanese lexicon. The second one is unassimilated loan words, also called *nonce borrowings* (Poplack et al., 1988), which can preserve their original morphological and phonological features when they occur

in a CS speech. From this perspective, nonce borrowings should be considered as CS when it occurs (Bullock & Toribio, 2009). Therefore, I excluded all the assimilated borrowings and counted nonce borrowings as a type of intra-sentential CS.

All tokens come from the speech of two main male characters (Baimao and Daodao), because they produced most of the CS in all six seasons of *Anar Pishti*. Tokens were exhaustively extracted from these characters for the first three episodes in each of the six seasons (cf. the semiotic representation analysis, for which I reviewed every episode in every season). Only CS that happened within the main content (i.e. the main script) of each episode were included. I excluded 13 CS tokens because they were in dialogues that occurred in blooper reels, as in (1) below, where the utterance is addressed to the director of the show, and not part of the script.

(1) Ximu: Di-gi-næ hæ,　导演　　boghan-dinkin.

　　Gloss: Say-2SG-IMP　director　are-since

　　Trans: Say something, since you're the director.

A further 3 tokens of CS in Season 1, Episode 1 were excluded because they were what I'll refer to as "mock code-switching". They are produced when one of the characters (Baimao) is teaching others how to pronounce a certain Chinese Mandarin word, as shown in (2) below:

The scene: Baimao (Ahbash) and Daodao (Ushuhpichaq) encounter a gang on the street. One gang member threatens and provokes them.

(Ahbash and gang members, Season 1, Episode 1, 1:38)

(2) Gang member 1: kø-si-tæ　　　　yangzangni,　kø-si-tæ　　　yangzan [jɑŋzɑŋ]-ni, hæ!

　　Gloss:　　　　Show-3SG-IMP　what you got show-3SG-IMP　what you got

Trans: Show me what you got, huh? Show me what you got!

(Here, "yangzang [jɑŋzɑŋ]" is a loan word from Chinese word "样子", meaning "appearance, gesture, manner, ability or posture.")

Ahbash: 样子 [jàŋzə]！

Gloss: Appearance

Trans: Appearance!

(Here the show immediately switched into the next scene after the provocation of the gang member where Ahbash starts to teach the gang how to pronounce the Chinese word "样子 [jàŋzə]" correctly)

The gang: 样子!

Gloss: Appearance

Trans: Appearance!

Ahbash: 样子!

Gloss: Appearance

Trans: Appearance!

The gang: 样子!

Gloss: Appearance

Trans: Appearance!

Ahbash: *样子!*

Gloss:　Appearance

Trans: Appearance!

The gang: *样子!*

Gloss:　Appearance

Trans: Appearance!

Ahbash: *四声样!*

Gloss:　fourth tone Appearance

Trans: fourth tone "*样*"!

The gang: *四声样!*

Gloss:　fourth tone Appearance

Trans: fourth tone "*样*"!

Here, the *Anar Pishti* writers are using this scene to mock Xinjiang people who always use Chinese loan words and code-switch between Uyghur and Chinese Mandarin. Such people are normally considered "less" Uyghur by the general public, especially people who have a Uyghur education background rather than Chinese Mandarin education background after Mandarin replacing Uyghur as the MoI in many schools. If, as I hypothesized, earlier seasons are not heavily influenced by government censorship, it's not surprising that this kind of mocking is not seen again after Season 1.

The final CS dataset consists of 155 tokens of CS observed on *Anar Pishti.* The dependent variable for this study is the rate of CS per minute per season. To calculate this rate, I divided the total number of CS by the aggregated length in minutes of the three episodes in each season, which includes all the time used for silence and advertisements.

5.1.1.2.   Independent variables

All tokens were coded for three independent variables: Season (1-6), Type (inter-sentential, intra-sentential, inter-speaker), and Direction (Uyghur-to-Chinese, Chinese-to-Uyghur, Mix).

5.1.1.2.1. CS type

I coded the CS tokens by Type to see whether there is a gradual structural integration of Mandarin into Uyghur in CS from earlier seasons to later seasons. Intraspeaker instances of code-switching were coded as either *intrasentential* CS (as in (3)) or *intersentential* CS (as in (4)). It has to be noted that I did not specifically code for System vs Content switches, but I will return to these concepts in the discussion chapter.

(3) Intra-sentential CS:

(Ahbash, Season 4, Episode 1, 2:44)[6]

Ahbash: 火锅-gha   chilap yey-dighan 金针菇 waghu.

Gloss: hotpot-into dip   eat-to   enoki mushroom

Trans: Enoki mushroom, which you dip into hotpot to eat.

---

[6] Throughout the paper, examples from the *Anar Pishti* data are notated as follows: Character name, season number, Episode number, Start time of example within the episode in minutes and seconds.

(4) Inter-sentential CS:

(Boss, Season 4, Episode 1, 10:38)

Boss: 这公司　　有那么多人 上班，它不缺你一个我告诉你！ Ishlimisængmu buldu!

Gloss: this company has many people work it-NEG lack you I tell you work-neg-1SG

Trans: There are a lot of people working in this company, it makes no difference whether you are here or not! You don't have to work here!

Cases where one speaker uses language A and the next speaker uses language B were coded as *interspeaker* CS, as in (5).

(5) Inter-speaker CS:

(Ushuhpichaq and Ahbash, Season 4, Episode 1, 3:15)

Ushuhpichaq: Biræ ay　　yitæmdikæn u?

Gloss:　　one month last-QM　　it

Trans:　　Can it last for one month?

Ahbash:　　充电宝.

Gloss:　　power bank

Trans:　　 Power bank.

5.1.1.2.2. Direction of code-switching

All tokens of CS were coded for Direction of code-switching to see if the code-switches on the show become gradually dominated by Mandarin over time and replace Uyghur. If the switch was

from a matrix/main clause[7] in Uyghur to a word or subordinate clause in Chinese Mandarin, it was coded as a Uyghur-to-Chinese (UG_CH) token, as in (6) and (7). If the switch was from a matrix/main clause in Chinese Mandarin to a word or subordinate clause in Uyghur, it was coded as a Chinese-to-Uyghur (CH_UG) switch, as in (8) and (9). It has to be noted that "matrix clause" in the note below does not refer to Matrix Language (ML), but I will return to the concept in the discussion chapter.

(6) UG_CH (Intra-sentential CS (the matrix clause is in Uyghur):

(Ahbash, Season 4, Episode 1, 2:44)

Ahbash: 火锅 gha   chilap yey-dighan 金针菇 waghu.

Gloss: hotpot-into dip   eat-to   enoki mushroom

Trans: Enoki mushroom, which you dip into hotpot to eat.

(7) UG_CH (Inter-sentential or Inter-speaker CS: preceding sentence is in Uyghur and following sentence is in Chinese Mandarin):

(Ushuhpichaq and Ahbash, Season 4, Episode 1, 3:15)

Ushuhpichaq: Biræ ay   yitæmdikæn u?

Gloss: one   month last-QM   it

Trans: Can it last for one month?

---

[7] Here matrix clause is defined as a clause that structurally surrounds an embedded clause, and main clause is defined as the clause that contains the main subject and the main verb.

Ahbash: *充电宝.*

Gloss: power bank

Trans: Power bank.

(8) CH_UG (Intra-sentential CS: the matrix clause is in Chinese):

(Ximu, Season 6, Episode 1, 7:43)

Ximu: *这个* adilæ *太好吃了， 你吃不吃?*

Gloss: this adilæ   very delicious you eat-NEG-eat

Trans: adilæ (a brand of cookie) is very delicious, do you want some?

(9) CH_UG (Inter-sentential or Inter-speaker CS: preceding sentence is in Chinese Mandarin and following sentence is in Uyghur):

(Boss, Season 4, Episode 1, 10:38)

Boss: *这公司有那么多人上班,   它不缺你一个我告诉你!* Ishlimisængmu buldu!

Gloss: this company has many people work it-NEG lack you I tell you work-neg-1SG

Trans: There are a lot of people working in this company, it makes no difference whether you are here or not! You don't have to work here!

In inter-speaker code-switching cases where the switch is from a code-switching sentence to a Uyghur sentence, it is coded as Mix_UG as in (10). If it is from a code-switching sentence to a Chinese Mandarin sentence, then it is coded as Mix_CH as in (11).

(10) Mix_UG (Inter-speaker CS: preceding sentence is a CS sentence and following sentence is in Uyghur):

(Ushuhpichaq and Ahbash, Season 3, Episode 1, 11:19)

Ushuhpichaq: mæn quyash-qa 开关 bekit-watqan.

Gloss: 1SG sun-for   switch build-PROG

Trans: I was building a switch for the sun.

Ahbash: yenila burunqi ish-im   yahshimunimæ.

Gloss: it seem previous job-1SG better

Trans: It seems that my previous job is better.

(11) Mix_CH (Inter-speaker CS: preceding sentence is a CS sentence and following sentence is in Chinese Mandarin):

(Ximu and Teacher, Season 6, Episode 1, 7:43)

Ximu: 这个 adilæ 太好吃了， 你吃不吃?

Gloss: this adilæ very delicious you eat-NEG-eat

Trans: adilæ (a brand of cookie) is very delicious, do you want some?

Teacher: *西木，你又在吃零食?！   那是什么?*

Gloss: Ximu   you again PROG eat snack that is what

Trans: Ximu, you are eating snacks again?! What's that?

5.1.2. Results

5.1.2.1. CS Type

Table 4 below shows the percentage of each type of CS across all six seasons. It's calculated by dividing the number of each type of CS by the total number of CS.

Table 4 Number and proportion of CS by Type, all seasons

CS type	CS_N	%
Intra-sentential CS	111	79.9
Inter-sentential CS	5	3.6
Inter-speaker CS	23	16.5
Total	139	100

Table 4 above shows that most of the CS occurring in the show is intra-sentential CS (79.9%), with inter-sentential CS being the second and inter-speaker CS being the least frequent. This result, on the surface, seems to suggest that integration of Chinese language into *Anar Pishti* becomes more complex over time, but I will discuss more about what this result indicates in chapter 6.

5.1.2.2. CS Direction

Table 5 below is the percentage of each type of CS direction in all six seasons.

Table 5 Number and proportion of CS by Direction

CS direction	Number of CS	%
UG_CH	117	84.2
CH_UG	9	6.5
Mix_UG	13	9.3
Mix_CH	0	0
Total	139	100

Table 5 above shows that most of the CS in all six seasons is from Uyghur to Chinese Mandarin, with from CS sentence to Uyghur being the second, and from Chinese to Uyghur being the third. And there is no CS that is from a CS sentence to a Chinese sentence. Again, I will talk more about it in the discussion chapter.

5.1.2.3. CS/Min by season

For each of the six seasons of *Anar Pishti*, Table 6 shows the number of CS, the aggregate time length of the first three episodes and the frequency of CS per minute.

Table 6 Number of Code-Switches Per Minute by Season

	Seasons						
	S1	S2	S3	S4	S5	S6	Total

Table 6 (cont'd)

N CS	1	5	39	29	44	21	139
Time (in minutes)	37.05	50.18	47.57	43.25	46.82	55.65	280.52
CS per min.	0.03	0.10	0.82	0.67	0.94	0.38	0.50

We see that there is a dramatic increase starting from Season 3 until Season 5. There is a very dramatic drop from Season 5 (0.94 CS/min) to Season 6 (0.38 CS/min) in terms of the frequency per minute of CS. This is because there are scenes in Season 6 where there is only Chinese Mandarin dialogues are used instead of simple Chinese lexical insertions (to Uyghur sentences), which leaves out the possibility of occurring of CS. I calculated the length of the pure Chinese Mandarin scenes as a percentage of the length of each episode in which they occur. On average, pure Mandarin dialogue accounts for 24.4% of episode time as shown by Figure 5 below.

Figure 5 Percentage of pure Chinese Mandarin Dialogue in first 3 episodes of S6
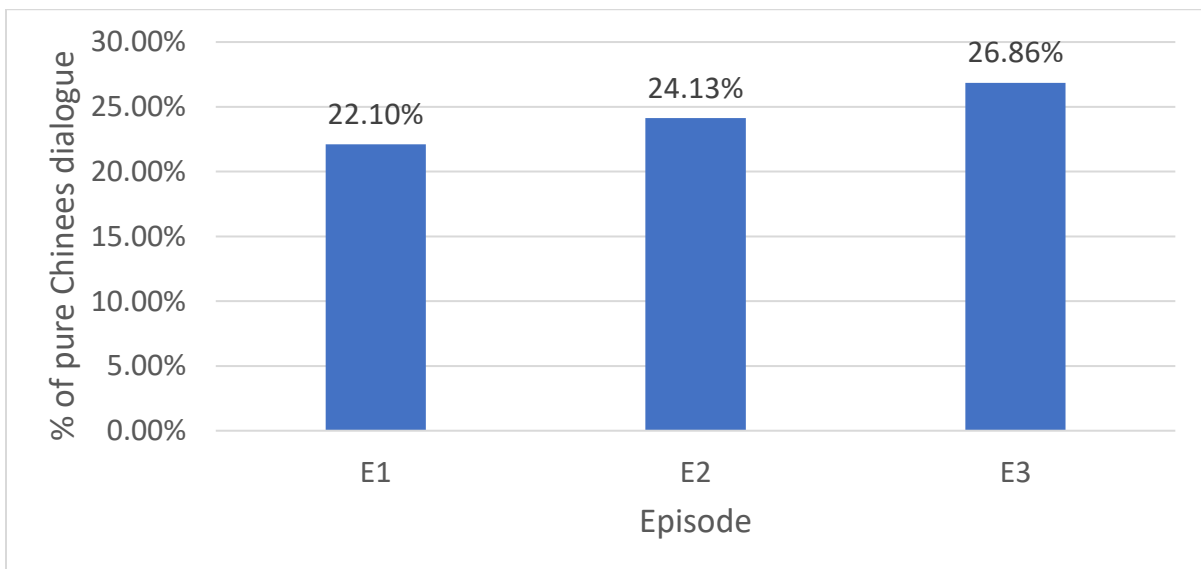
Why is there such a large proportion of Chinese Mandarin-only dialogue in Season 6? It is partly because in this season, three Chinese Han characters are added as long-term fixed characters for the show. But the *number* of regular Chinese characters is not in itself what gives rise to the sudden inclusion of so much Chinese-only dialogue. Crucially one of the new characters is a Teacher character, and as such they are an authority figure. Contrast this with Season 1, in which there is only one Chinese Han character as a tattoo artist (not as an authority or a fixed character) with strong Xinjiang Chinese accent. Authority figures have greater license to speak at length than other individuals. And as mentioned in Chapter 2, Chinese Mandarin is now the MoI in Xinjiang in most schools.

In addition, unlike other normal scenes in the show, which are always different from each other with different story lines and different characters (played by those main actors) and never appear repeatedly within different episodes and seasons, those scenes which are entirely in Chinese Mandarin always appear regularly and repeatedly in each episode of Season 6. And more interestingly, one of those scenes has a particular name, "Little Chinese Culture Class" ("中华文化小课堂", look at figure 6 below), in which the Teacher character always only talks to one main Uyghur character, Ximu, and teaches him about Chinese Han culture (e.g. playing couplet games ("对对联")) and Chinese law (e.g. Chinese traffic law), even though there's another Chinese Han character always appearing with Ximu in the scene.

Figure 6 "中华文化小课堂" ("Little Chinese Culture Class")

Nevertheless, this raises the question: Does the proportion of Chinese on *Anar Pishti* increase partly because there are more Han Chinese characters on the show over time? The answer to this question will be shown in the next section.

5.2. Chinese spoken by Uyghur characters

In this section, to test my Hypothesis 3, namely, to see whether the presence of a Chinese character in a dialogue or scene in the show (either as an addressee or auditor) actually influences the proportion of Chinese spoken by the Uyghur characters, I focus on the analysis of Chinese utterances produced only by Uyghur characters in different circumstances in each season.

5.2.1. Data and method

I extracted all the Chinese utterance spoken by Uyghur characters from the dataset (n=654).

For the purposes of the analysis, I operationalized *addressee* as the single individual to whom the speaker was directing their utterance. Where there was ambiguity – such as when a character seemed to be addressing two or more other characters at once, I operationalized all other characters in the sketch as *auditors*, regardless of whether they were 'ratified' to be in the interaction or not, i.e. I did not distinguish between within-scene auditors and within-scene overhearers. Also, I consider the viewing audience as *auditors*; however, since the show is ultimately still a Uyghur show and almost no Han Chinese in Xinjiang as well as in China in general can speak Uyghur, I assume all viewing audience are Uyghurs, which is why my analysis focuses only on the *auditors* in the scenes where there are Uyghurs and Han Chinese.

In addition, apart from the mean length of Chinese utterances and proportion of Chinese utterances spoken by Uyghur characters, I am not going to discuss other potential independent variables that might predict Chinese use by Uyghur characters such as topic of the utterance, whether the utterance is serious or funny, sex of the speakers, social context/domain of use, etc. I leave this for future work.

5.2.2. Results

By excluding the utterances spoken by Uyghur characters to non-Uyghur addresses (n=175) from the total number of Chinese utterance spoken by Uyghur characters (n=654), I first analyzed the utterances produced by Uyghur speakers alone without any addressee plus the utterances to Uyghur addressees (n=479), and then calculated the mean length of such utterances in each season

(Table 7 and Figure 6). Table 8 and Figure 7 also display the utterances as a proportion of all speech in each season.

Table 7 Mean length of Chinese utterances by Uyghur speakers to Uyghur addressees

Season	N of utterances	Average length (S)
S1	4	0.76
S2	26	0.66
S3	72	0.50
S4	100	1.81
S5	140	2.21
S6	137	1.88

Figure 6 Mean length of Chinese utterances by Uyghur speakers to Uyghur addressees
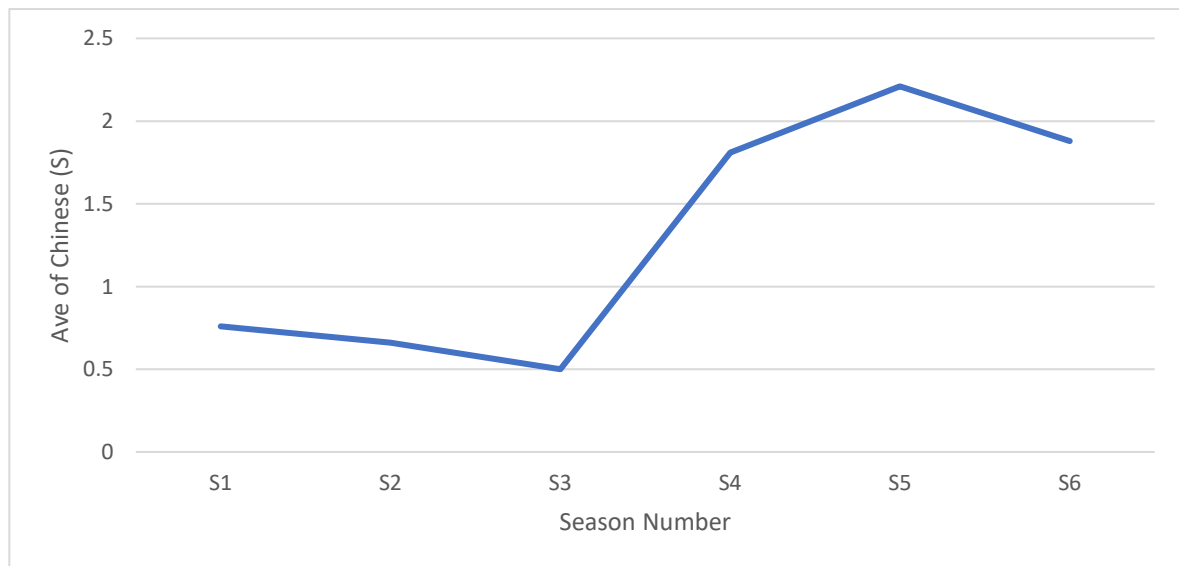
Table 8 Proportion of Chinese utterances by Uyghur speakers to Uyghur addressees as a

proportion

Season	Total Chinese by Uyghur to Uyghur (s)	Total speech (s)	Proportion of Chinese by Uyghur to Uyghur (%)
1	3.04	1809.96	0.17
2	17.16	2140.1	0.80
3	36.04	2656.92	1.36
4	181.08	2700.15	6.71
5	309.29	3015.98	10.26
6	258.18	4380.2	5.89

Figure 7 Proportion of Chinese utterances by Uyghur speakers to Uyghur addressees as a
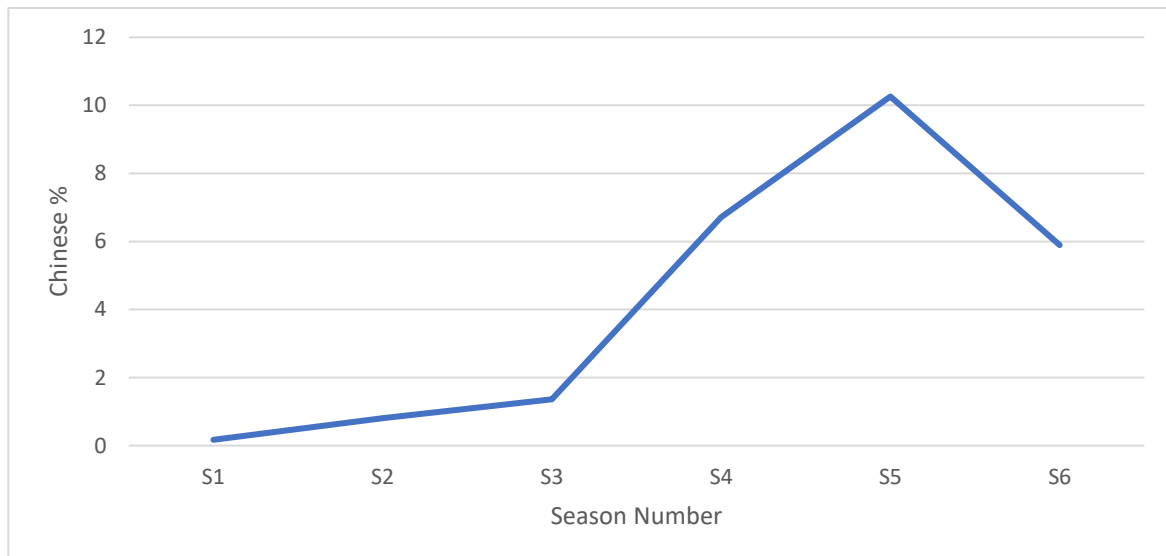
proportion

The proportion of Chinese and the average time length of Chinese utterances produced by Uyghur speakers increase notably starting from Season 4 till Season 5, but they decrease from Season 5 to 6; however, both measures of Season 6 are still much higher than the ones of the first three seasons.

Since the number of Han Chinese characters on *Anar Pishti* increases over time, it was important to rule out their influence as auditors on the speech of Uyghur characters. That is, Uyghur characters might be addressing each other more frequently in Chinese to accommodate to co-present Chinese Han characters (since neither the show's creators, nor the audience, nor the fictional Uyghur characters would be likely to assume that Han Chinese people would be able to understand the Uyghur language).

The analysis suggests that we can rule out this explanation. I constrained the dataset to the 433 utterances without any Chinese auditors and calculated the mean length of such utterances in each season (Table 9 and Figure 8). Table 10 and Figure 9 also display the utterances as a proportion of all speech in each season. Table 9 (and Figure 9) and Table 10 (and Figure 10) show that the proportion of Chinese and average time length of Chinese utterances produced by Uyghur speakers increase notably starting from Season 4 till Season 5 but decrease from Season 5 to 6; however, both measures of Season 6 are still higher than Season 1, 2 and 3.

Table 9 Mean length of Chinese utterances by Uyghur speakers to Uyghur addressees with no Chinese auditors present

Season	N utterances	Average length
S1	4	0.76
S2	26	0.66

Table 9 (cont'd)

S3	72	0.50
S4	100	1.81
S5	120	2.06
S6	111	1.94

Figure 8 Mean length of Chinese utterances by Uyghur speakers to Uyghur addressees, with
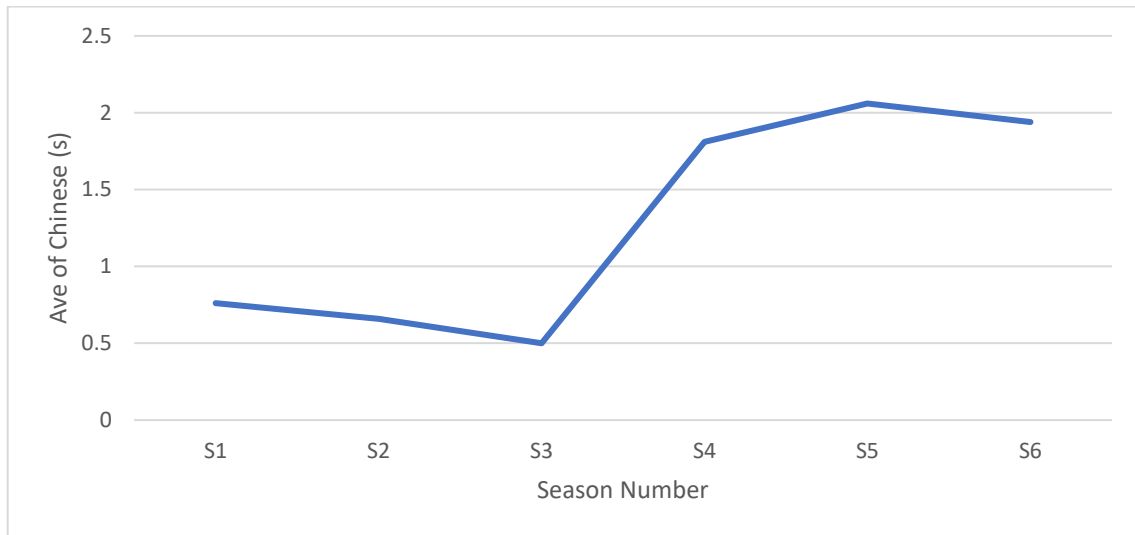
no Chinese auditors present

Table 10 Chinese utterances by Uyghur speakers to Uyghur addressees as a proportion of all

speech without Chinese auditors present

Season	Total Chinese by Uyghur to Uyghur (s)	Total speech (s)	Proportion of Chinese (%)
1	3.04	1809.96	0.17
2	17.16	2140.1	0.80
3	36.04	2656.92	1.36
4	181.08	2700.15	6.71

Table 10 (cont'd)

| 5 | 246.63 | 3015.98 | 8.18 |
| 6 | 215.84 | 4380.2 | 4.93 |

Figure 9 Chinese utterances by Uyghur speakers to Uyghur addressees as a proportion of all speech without Chinese auditors present
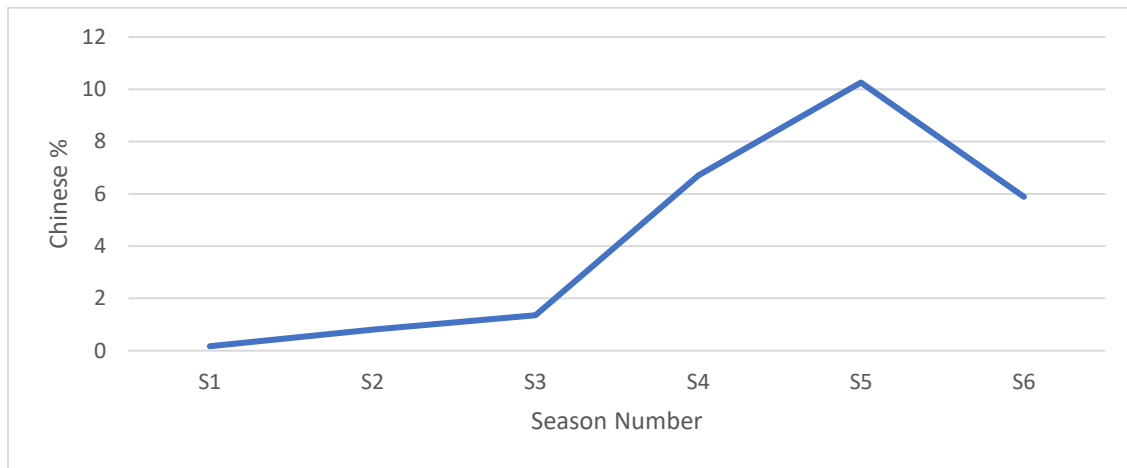
These results indicates that even without Chinese addressees and auditors, Uyghur characters produced notably more amount of and more complex Chinese in later three seasons comparing to earlier three seasons. This means that the increasing proportion of Chinese used on *Anar Pishti* is not because there are more Han Chinese characters on the show over time, but because there has been more government censorship since 2017.

5.3. Summary

In summary, the results of the CS analysis indicate that CS frequency goes up starting from Season 3, which is inline with the results of the amount of Chinese utterances used in the show overtime. The results also indicate that most of the CS produced in the show is intra-sentential CS, which will be discussed in the next chapter to see whether this result supports Hypothesis 2.

The results of the proportion of Chinese utterances produced by Uyghur characters in different situations show that after 2017 Uyghurs start to increasingly use more Chinese utterances to address each other even when there are no Han Chinese in the scenes. This means that the main reason of this notable difference between earlier and later three seasons is because of the influence from the Chinese government as the overhearer of the show, not from the Han Chinese addressees or auditors. Therefore, I believe the results show that the effect of the Chinese government as the overhearer overrides the effect of the Chinese addressees and auditors in the show, which is in line with the Hypothesis 3.

6. Discussion, conclusion, and implications

This thesis has investigated the nature and amount of Chinese language use on the Uyghur online comedy show *Anar Pishti* across six seasons (2016-2020). My research asked whether language use on *Anar Pishti* really did change over time, and if yes, whether that change could be ascribed to Chinese government media censorship (or even the mere threat of such censorship).

In this section I reiterate my hypotheses (Chapter 3), summarize results, and discuss what I found.

6.1. Cultural change and language change on *Anar Pishti* over time

To answer my research question, I had to first determine whether my intuitions about language and culture on *Anar Pishti* could be confirmed, which led to this hypothesis:

H1: The amount of Chinese cultural and linguistic representation will increase in *Anar Pishti* after 2017.
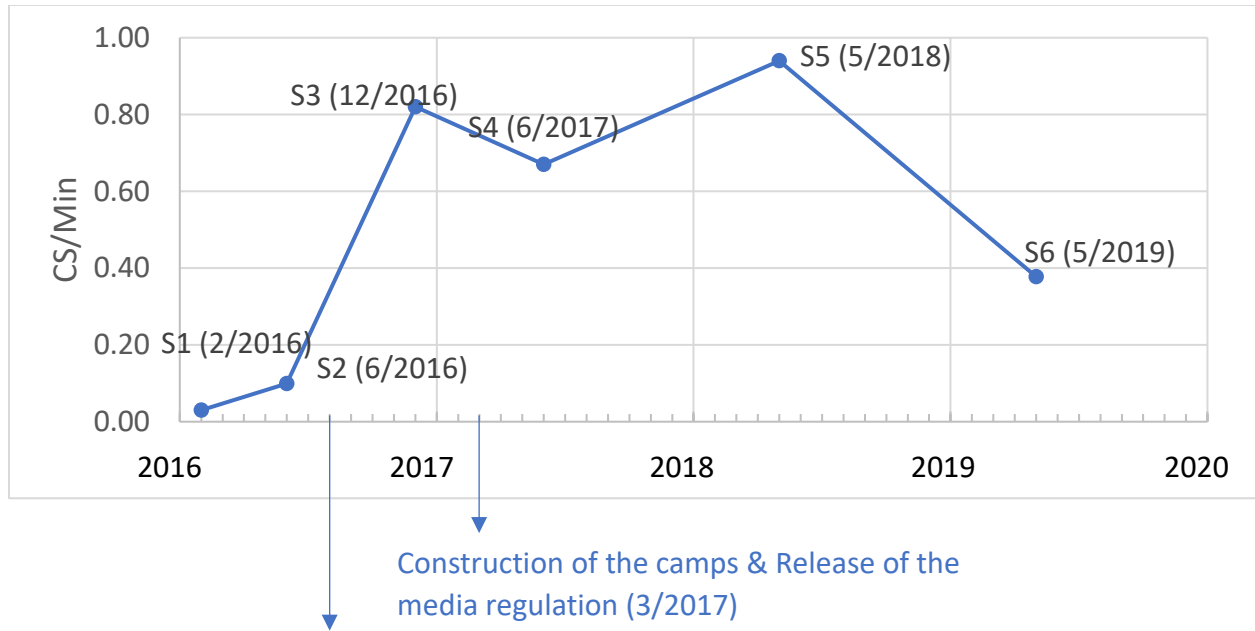
Overall, quantitative analysis of the data confirms Hypothesis 1. There was an abrupt uptick in the proportion of Chinese language use on *Anar Pishti* after 2017 (Chapter 4, section 4.2.2), and in the mean length of Chinese utterances (Chapter 4, section 4.3.2). In addition, the semiotic changes from Season 1 to 6 indicate the increasing amount and penetration of Chinese semiotic elements over the later seasons in the show.

There was also a jump from occasional lexical insertions to frequent all-Chinese utterances. Notably, this occurred not simply as a result of adding Han Chinese characters and settings, but because *Uyghur* characters were given more dialogue in Chinese. This can be seen in Figure 10, which graphically displays the results from Table 6 (Chapter 5, section 5.1.2.3) of code-switching

frequency per minute, season by season. The numbers in parentheses after each season number indicate the month and year of the season's release, e.g. Season 3 was released in December 2016. Code-switching occurs frequently at a rate of around 0.8 switches per minute in Seasons 3, 4, and 5, but drops dramatically in Season 6. The drop is due to the inclusion, in Season 6, of all-Chinese-language scenes, which meant that switches from Uyghur to Chinese were much less frequent (as discussed in Chapter 5, section 5.1.2.3).

Figure 10 is annotated to show the correlation between the increase in code-switching on *Anar Pishti* and two relevant social events in Xinjiang. The first is the appointment of Chen Quanguo (8/2016), the party secretary of Xinjiang, who is responsible for the political climate change in Xinjiang and the passing date of the censorship regulations. The second is the release in March 2017 of Chinese government regulations (described in Chapter 2, section 2.3.1) of "extremist" media content. Around the same time, the government constructed the "re-education camps" into which many Uyghur citizens have since been interred.

Figure 10 Timeline of each Season's release and other events

This means that there's a strong correlation between the increasing amount of semiotic representation of Chinese authority, and a heavy Chinese authoritative presence in Xinjiang. One interpretation is that government censorship –or the threat of censorship—led the *Anar Pishti* writers to incorporate more Chinese language and semiotic representation into the show.

But the correlation could be spurious. Or it can indicate that the writers of the show are reflecting an acceleration of the language shift in Xinjiang due to the dramatic political change in the post-2017 external political environment (one similar example will be the situation of the Dutch language spoken in France, mentioned in Chapter 3, section3.2.2). Code-switching is not only a behavior on the individual level, determined by individual choices. It can also reflect societal multilingualism and can be associated in some cases with community language shift. Romaine stated that "language shift generally involves bilingualism (often with diglossia) as a

stage paving the way to eventual monolingualism…" (1994, p.50), yet here I consider whether the rise in Chinese language use on *Anar Pishti* is due to societal language shift beyond the show. Helpfully, David (1996, p.58-64) provides 9 different criteria for determining whether CS is a sign of an ongoing language shift. Three of them are closely related to the current project. Note that David uses the term "ethnic language" to refer to speakers' first language (L1), in contrast with the societally more dominant (in terms of power or size of population) language, which David calls the "non-ethnic language".

David (1996) criterion 1

"Code switching can be seen as a symptom of language shift when the use of code switches is triggered by an imperfect knowledge of the ethnic language."

Anar Pishti is a Uyghur show and characters/writers don't code-switch to Mandarin because of limited knowledge in Uyghur. Even though the characters/writers increase their frequency of CS over the six seasons, it's not likely to be due to a rapid decrease (in just a few years) in their knowledge of Uyghur.

David (1996) criterion 9

"Language shift is also indicated if the ethnic language is used less sparingly in code switches and the dominant language is the non-ethnic language."

If the code-switches on *Anar Pishti* had become gradually dominated by Mandarin over time, namely, the direction of CS had become mainly from Mandarin to Uyghur, then this would be an indicator of language shift more than censorship. But as shown in Table 5 in Chapter 5, section 5.1.2.2, switches from Mandarin to Uyghur are extremely rare in all seasons of the show. This too

suggests that a censorship interpretation of the correlation between CS increase and government repression is a better fit for the data than a language shift interpretation.

To further rule out language shift as an explanation, we now consider my second hypothesis, and the last of the three relevant David (1996) criteria.

H2: Integration of Chinese language into *Anar Pishti* will not become more complex over time.

The empirical evidence points toward confirmation of this hypothesis. Almost 80% of code-switches in the data were intrasentential, which is superficially suggestive of highly intimate integration of the grammars of Uyghur and Chinese Mandarin. Poplack (1980), for example, places intrasentential switches higher on an integration scale than inter-sentential switches (Poplack, 1980, p.614). However, as I mentioned earlier in chapter 3, based on the Matrix Language Frame (MLF) model created by Myers-Scotton (1997), for the integration of language A into language B to be "complex", in an intrasentential code-switching sentence, the System Morphemes (e.g., function words, inflectional morphemes, etc.) have to come from language A. In other words, language A has to transform from being the Embedded Language (EL) that consists of Content Morphemes (e.g., nouns, verbs, adjectives, and some prepositions) to being the Matrix Language (ML) that consists of System Morphemes in a code-switching sentence. But, based on my observation, within most of the intrasentential CS cases in the *Anar Pishti* data, the System Morphemes come from Uyghur, not Chinese Mandarin. This means the Matrix Language (ML) is still Uyghur, and the integration of Chinese Mandarin is still "superficial". More specifically, as shown in the table 11 below, most of the intrasentential CS in *Anar Pishti* are like the example in the left-hand column, where the System Morphemes are from Uyghur and the speaker inserts a Chinese content morpheme (e.g., a noun 开关 "switch"). Only a minority (n=2) are similar to the example in the

right-hand column, where the speaker inserts a System Morpheme (the past tense morpheme) from Chinese Mandarin. And even when the mean length of Chinese utterances increases in later seasons (as shown in Chapter 4, section 4.3.2), the longer Chinese words/phrases still consist of Content Morphemes, namely, they are still simple lexical insertions.

Table 11 Two different types of intrasentential code-switches

Chinese Content Morpheme insertion	Chinese System Morpheme insertion
Season 3, Episode 1: CS: mæn quyash-qa 开关 bekit-watqan. Gloss: I sun-for switch build-PAST-PROG Translation: I was building a switch for the sun.	Season 3, Episode 3: CS: mæn 赞了 qighan ikkingla-ning Gloss: 1SG like-PAST do 2PL-POSS CS: ræsimi-gæ. Gloss: picture-DAT Translation: I liked both of your posted pictures.

Therefore, the evidence does not fit David's fifth criterion:

David (1995) criterion 5

"In order to ascertain the extent of language shift or maintenance, the degree of use of the ethnic language in code switches should be determined. If such switches are made up of minimal linguistic items from the non-ethnic language while the dominant language in the discourse remains the ethnic language, then it would be valid to conclude that the ethnic language is not threatened."

CS to Mandarin on *Anar Pishti* does not begin as inserting "minimal linguistic items" from Mandarin, becoming gradually more linguistically integrated over time. And in Season 6, the

characters are simply speaking Chinese, with no code-switched elements from Uyghur. Xinjiang's creators and actors grew up in a school system in which Chinese Mandarin was the MoI; they could have written a large proportion of *Anar Pishti'*s dialogue in Chinese Mandarin from the start of Season 1 if they chose to. They did not make this choice. Further, there is no evidence that language shift in Xinjiang underwent a dramatic acceleration starting in 2017 that would correlate with the *Anar Pishti* findings.

In sum, it's unlikely that the rise in CS to Chinese, and the rapidly increasing overall use of Chinese in *Anar Pishti,* is due to the writers unconsciously reflecting their own (or Xinjiang's) language shift to Chinese. But it would still be good to have stronger potential evidence that the show's increase in Chinese language is prompted by government pressure. This brings us to my third hypothesis.

H3: *Anar Pishti* characters will increasingly design their speech for government overhearers rather than other characters in the show.

To test this hypothesis, I reasoned that the best way to isolate government influence from language shift or the writers' creative choices was to see whether overhearer effects increasingly outweighed addressee and auditor effects, in violation of Bell's audience design model (1984). I isolated the speech of Uyghur characters speaking Chinese to other Uyghur characters, which increased over the seasons. Crucially, there was a big increase from Season 4 onwards in Uyghur characters addressing other Uyghur characters in Chinese, even when no Han Chinese characters were co-present in the sketch (Chapter 5, section 5.2). I interpret this to mean that the *Anar Pishti* writers are designing Uyghur characters' speech for the possibility that unaddressed, unratified government watchers ('overhearers') are monitoring the show.

6.2. Other potential explanations: Creative choice and social context

Using Chinese Mandarin in *Anar Pishti* could also be merely a creative choice made by the screen writers and the director of the show. Portraying foreign languages in films and novels is not uncommon in the creative industry. For instance, the English novel *Bodega Dreams* by Ernesto Quiñonez (shown in Figure 11 below) has many instances of Spanish because the novel is mainly about life experience of Latino people in America.

Figure 11 *Bodega Dreams* by Ernesto Quiñonez Figure 12 Comedy-drama TV series *Ramy*
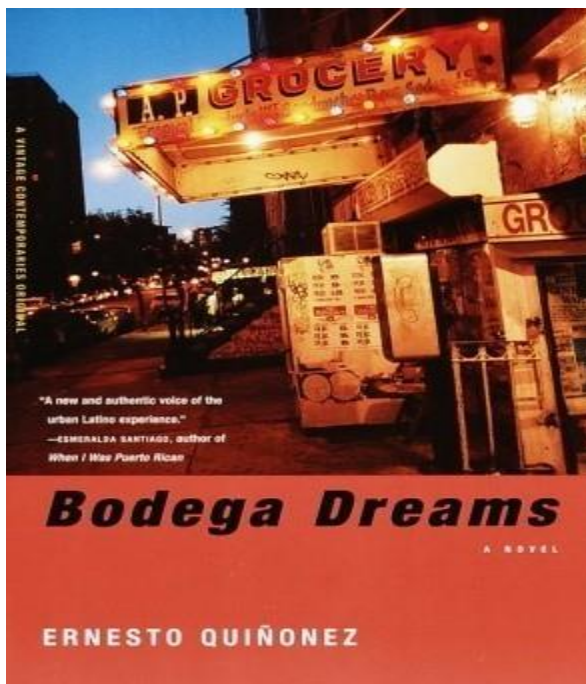
The American English comedy-drama TV series *Ramy* (Figure 12) also has Arabic in it because the show depicts the life of American Muslims. The use of foreign languages in these kinds of creative works often play a role of attracting a broader audience or making the work more authentic for a local audience, and the creators of *Anar Pishti* might have made the same decision for the same reason. Further support for this possibility comes from my observations of *Tohtima*

Tohti, which is also a Uyghur comedy online show. Like *Anar Pishti,* it also mainly targets the young generation of Uyghur people in Xinjiang. The first season of the show was released on April 15th in 2017 on the online platform, Tencent Video ("腾讯视频"), and one more season was released in 2018. However, even though the first season was released in 2017, unlike *Anar Pishti,* it doesn't seem to have many Chinese sentences based on my observation of the first three episodes. Perhaps the creators of *Tohtima Tohti* made a decision on artistic grounds not to reflect the reality of Uyghur people using Chinese at times in everyday life.

However, this interpretation still cannot rule out the influence from governmental censorship on *Anar Pishti.* This is because the first two seasons of *Anar Pishti* include very few instances of CS (more like *Tohtima Tohti*) but the frequency of CS increases dramatically around 2017 when the crack-down on Uyghur Muslims happened in Xinjiang. If the Chinese government's censorship is targeting all Uyghur online media, including Uyghur shows, then it's hard to explain why *Tohtima Tohti* is not being influenced as much as *Anar Pishti*. One possible explanation is that the Chinese government only targets popular Uyghur shows. Based on the "collective action potential" theory (King, Pan & Roberts, 2013), the targets of Chinese government censorship are those who have the potential to create collective actions outside the internet, like protests. Given that *Anar Pishti* has much higher popularity than *Tohtima Tohti*[8], it's reasonable to infer that the Chinese government viewed *Anar Pishti* as having much more potential to cause collective actions outside the online environment among Uyghurs in Xinjiang than *Tohtima Tohti*.

Nevertheless, one can still argue that using Chinese Mandarin to superficially nod to the censors without actually changing much of the linguistic fabric of *Anar Pishti* is itself perhaps a

---

[8] Official viewing figures for *Anar Pishti* and *Tohtima Tohti* are not available, I base my judgment of their relative popularity based on the number of reviews they received on Douban (the biggest Chinese online database of information related to film, TV shows, books, music, etc.). The first season of *Anar Pishti* received 403 reviews, while the first season of *Tohtima Tohti* only received 3.

creative choice. But even from this perspective, one cannot rule out the influence of censorship since they do it to be in line with the censors in the first place.

In addition, based on the language attitude studies mentioned above (Cao & Wang, 2009; Sun, 2012; Elterish, 2016), Uyghurs are more likely to speak Chinese Mandarin in a social environment/context related to authority or government. Therefore, one could ask: could it be that use of Chinese on *Anar Pishti* increases just because there happen to be more such social contexts in the later seasons? The answer is probably "no" since based on my own observation, the Chinese use in later seasons occurs in a variety of social contexts, not just contexts where there is an authority.

6.3. Further implications for future studies

For future studies, one can analyze the data with different independent variables such as topic of the utterance, whether the utterance is serious or funny, sex of the speakers, social context/domain of use, etc. to avoid potential confounds.

Furthermore, in order to see whether the Chinese government censorship is only targeting Uyghur population or all ethnic minority populations in China, more similar online shows using different ethnic minority languages can be studied in the future.

6.4. Significance of the thesis

The study done in this thesis is significant for the following two reasons.

First, for its novelty. The study done in this thesis is novel due to: 1) its theme since very few sociolinguistics studies related to government oppression, especially government censorship, have been done in Xinjiang because of its political restrictions in the region; 2) its contribution to the

sociolinguistics studies related to TV and online shows since not many studies related to Uyghur language and Uyghur online shows have been done within the field; 3) its relevancy to the ongoing political change and oppression in Xinjiang, China. Therefore, the questions posed in the thesis and the answers to these questions will be of great interest for any researchers/regular readers who are interested in censorship studies, media studies, and Xinjiang contemporary politics.

Second, for its sociopolitical purpose. Another goal of this project is to cast important light on Xinjiang 's current political situation, how Muslims are treated in China and how they're targeted in Chinese censorship system, and demonstrate the religion, cultural and language erosion that Uyghurs are experiencing right now in China. Ethically, this project shows that censorship not only filters information but also is able to take away people's right to practice their religion, culture and eventually eliminate their language. It indicates that censorship is ethically wrong not only because of epistemological restrictions, but also because it can actively erase people's identity or even their existence.

BIBLIOGRAPHY

Androutsopoulos, J. (2012). Repertoires, characters and scenes: Sociolinguistic difference in Turkish-German comedy. *Multilingua*, 31(2/3), 301–326.

Androutsopoulos, J. (Ed.). (2014). *Mediatization and sociolinguistic change* (Vol. 36). Walter de Gruyter GmbH & Co KG.

Auer, P. (Ed.). (2013). *Code-switching in conversation: Language, interaction and identity*. Routledge.

BBC News. (2009). *Xinjiang arrests 'now over 1,500.* Retrieved March 12, 2021, from http://news.bbc.co.uk/2/hi/asia-pacific/8181563.stm

BBC News. (2014). *Deadly China blast at Xinjiang railway station*. Retrieved March 12, 2021, from https://www.bbc.com/news/world-asia-china-27225308

BBC News. (2018). *Lunar New Year: Chinese TV gala includes 'racist blackface' sketch*. Retrieved March 12, 2021, from https://www.bbc.com/news/world-asia-china-43081218

BBC News. (2020). *'The Uighurs and the Chinese state: A long history of discord.* Retrieved March 12, 2021, from https://www.bbc.com/news/world-asia-china-22278037?intlink_from_url=https://www.bbc.com/news/topics/cpz1y9ney3mt/uighurs&link_location=live-reporting-story

Bednarek, M. (2012a). Constructing 'nerdiness': Characterization in The Big Bang Theory. *Multilingua*, *31*(2), 199-229.

Bednarek, M. (2012b). "Get us the hell out of here": Key words and trigrams in fictional television series. *International Journal of Corpus Linguistics*, *17*(1), 35-63.

Bell, A. (1982). Radio: The style of news language. *Journal of Communication*, *32*(1), 150-164.

Bell, A. (1984). Language style as audience design. *Language in Society*, *13*(2), 145-204.

Bleichenbacher, L. (2012). Linguicism in Hollywood films? Representations of and audience reactions to multilingualism in mainstream movie dialogues. *Multilingua*, *31*(2/3), 155–176.

Blom, J., & Gumperz, J. (1972). Social Meaning in Linguistic Structures: Code Switching in Northern Norway. In Gumperz, J. & Hymes, D. (Eds.), *Directions in Sociolinguistics: The Ethnography of Communication* (pp. 407-434). Holt, Rinehart, and Winston.

Bubel, C. (2006). *The linguistic construction of character relations in TV drama: Doing friendship in Sex and the City* [Unpublished doctoral dissertation]. University of Saarbrücken.

Bullock, B. E., & Toribio, A. J. E. (2009). *The Cambridge handbook of linguistic code-switching*. Cambridge University Press.

Burnett, H. (2017). Sociolinguistic interaction and identity construction: The view from game-theoretic pragmatics. *Journal of Sociolinguistics*, *21*(2), 238-271.

Byram, M. & Hu, A. (Ed.). (2013). *Routledge Encyclopedia of Language Teaching and Learning*. Routledge.

Censorship (2020, November 19). In Wikipedia. Retrieved March 13, 2021, from https://en.wikipedia.org/wiki/Censorship.

Chen, X., & Ang, P. H. (2011). The Internet police in China: regulation, scope and myths. In Herold, D, & Marold, P. (Eds.), *Online Society in China* (pp. 40-52). Routledge.

Chen, H., & Mao, H. (2013). Measuring language attitudes based on matched-guise technique. In Du, W. (Eds.), *Informatics and management science IV* (pp. 773-779). Springer. https://link.springer.com/chapter/10.1007/978-1-4471-4793-0_99

Ethnologue: Languages of the World. (n.d.). *Chinese, Mandarin.* Retrieved December 15, 2020, from https://www-ethnologue-com.proxy2.cl.msu.edu/language/cmn

Côté, I. (2015). The enemies within: targeting Han Chinese and Hui minorities in Xinjiang. *Asian Ethnicity*, *16*(2), 136-151.

Cowlishaw, G. K. (2000). Censoring race in 'post-colonial' anthropology. *Critique of Anthropology*, *20*(2), 101-123.

David, M. K. (1996). *Language Shift Among the Sindhis in Malaysia* [Unpublished doctoral dissertation]. University Malaya.

Davis, E. V. W. (2008). Uyghur Muslim ethnic separatism in Xinjiang, China. *Asian Affairs: An American Review*, *35*(1), 15-30.

De Jonghe, A. (1967). *De taalpolitiek van Willem I in de Zuidelijke Nederlanden (1814-1830): De genesis der taalbesluiten en hun toepassing*. Darthet. Department of Xinjiang United Front. (2017, June 19). *Identifying Religious Extremism*. Xinjiang Documentation Project. Retrieved March 13, 2021, from https://xinjiang.sppga.ubc.ca/chinese-sources/online-sources/identifying-religious-extremism/

Dillon, M. (2003). *Xinjiang: China's Muslim far northwest*. Routledge.

Dimova, S. (2012). English in Macedonian television commercials. *World Englishes*, *31*(1), 15–29.

Dorian, N. C. (1973). Grammatical change in a dying dialect. *Language*, 413-438.

Dorian, N. C. (1977). The problem of the semi-speaker in language death. *International Journal of the sociology of language*, *12*(1), 23-32.

Dorian, Nancy. 1982b. Linguistic models and language death evidence. In L. Obler and L. Menn (Eds.), *Exceptional Language and Linguistics* (pp: 31-48). Academic Press.

Dynel, M. (2011). Stranger than fiction? A few methodological notes on linguistic research in film discourse. *Brno Studies in English*, *37*(1), 41–61.

Eberhardt, M., & Downs, C. (2015). "(r) You Saying Yes to the Dress?" Rhoticity on a Bridal Reality Television Show. *Journal of English Linguistics*, *43*(2), 118-142.

Nijmegen: Max Planck Institute for Psycholinguistics, The Language Archive. (2022). *ELAN* (Version 6.4) [Computer software]. Retrieved June 2, 2022, from https://archive.mpi.nl/tla/elan

Elterish, A. B. (2016). Language Use Among Uyghur Students in Xinjiang, PR China. In *Language change in central Asia* (pp. 197-220). De Gruyter Mouton.

Emery, L. (2002). Censorship in contemporary art education. *International Journal of Art & Design Education*, *21*(1), 5-13.

Encyclopedia Britannica. (2022). *Xinjiang*. Retrieved June 2, 2022, from https://www.britannica.com/place/Xinjiang/Climate#ref71457

Fasold, R. W. (1984). *The sociolinguistics of society* (Vol. 1). Wiley-Blackwell. Gal, S. (1979). *Language shift: Social determinants of linguistic change in bilingual Austria*. Academic Press.

Hei, K. C. (2002). Code switches: A measurement for language shift? A case study of a Malaysian Chinese family. *Methodological and Analytical Issues in Language Shift Studies*, 91-111.

Stamou, A. G. (2014). A literature review on the mediation of sociolinguistic style in television and cinematic fiction: Sustaining the ideology of authenticity. *Language and Literature*, *23*(2), 118-140.

Kam, S., & Clarke, M. (2021). Securitization, surveillance and 'de-extremization' in Xinjiang. *International Affairs*, *97*(3), 625-642.

Kamberi, D. (2015). Uyghurs and Uyghur identity. *Radio Free Asia*.

King, G., Pan, J., & Roberts, M. E. (2013). How censorship in China allows government criticism but silences collective expression. *American Political Science Review*, *107*(2), 326-343.

Kwon, N. A. (2013). Collaboration, coproduction, and code-Switching: Colonial cinema and postcolonial archaeology. *Cross-Currents: East Asian History and Culture Review*, *2*(1), 10-40.

Leibold, J. (2020). Surveillance in China's Xinjiang region: Ethnic sorting, coercion, and inducement. *Journal of Contemporary China*, *29*(121), 46-60.

Language Development Project. (n.d.). *4. Utterances.* Retrieved June 5, 2022, from https://ldp-uchicago.github.io/docs/guides/transcription/sect_4.html.

Mandala, S. (2007). Solidarity and the Scoobies: An analysis of the -*y* suffix in the television series Buffy the Vampire Slayer. *Language and Literature*, *16*(1), 53-73.

Milroy, L., & Muysken, P. (1995). Introduction: Code-switching and bilingualism research. *One speaker, two languages: Cross-disciplinary perspectives on code-switching*, 1-14.

Mufwene, S. S. (2002). Colonization, globalization and the plight of 'weak' languages. *Journal of Linguistics*, *38*(2), 375-395.

Muysken, P., & Muysken, P. C. (2000). *Bilingual speech: A typology of code-mixing*. Cambridge University Press.

Myers-Scotton, C. (1993). Common and uncommon ground: Social and structural factors in codeswitching. *Language in Society*, 475-503.

Myers-Scotton, C. (1997). *Duelling languages: Grammatical structure in codeswitching*. Oxford University Press.

National Radio and Television Administration. (2007, December 20). *Regulations on the Administration of Internet Audio-Visual Program Services*. Website of the Central People's Government of the People's Republic of China. Retrieved September 12, 2022, from http://www.gov.cn/flfg/2007-12/29/content_847230.htm

National Radio and Television Administration. (2014, March 19). *Notice of Strengthening Regulations on Internet Audio-Visual Programs like Online Shows and Micro-Movies*. Website of National Radio and Television Administration. Retrieved September 12, 2022, from http://www.nrta.gov.cn/art/2014/3/19/art_113_4861.html

People's Daily Online. (2017). *Xinjiang Uyghur Autonomous Region Regulation on De-extremification ("新疆维吾尔自治区去极端化条例")*. Retrieved September 12, 2022, from http://xj.people.com.cn/n2/2017/0330/c186332-29942874.html

Planchenault, G. (2012). Accented French in films: Performing and evaluating in-group stylizations. *Multilingua*, 31(2/3), 253–275.

Poplack, S., Sankoff, D., & Miller, C. (1988). The social correlates and linguistic processes of lexical borrowing and assimilation. *Linguistics*, 26, 47-104.

Quaglio, P. (2009). *Television dialogue: The sitcom Friends vs. natural conversation* . John Benjamins Publishing.

Reichelt, S., & Durham, M. (2017). Adjective intensification as a means of characterization: Portraying in-group membership and Britishness in Buffy the Vampire Slayer. *Journal of English Linguistics*, *45*(1), 60-87.

Rodríguez, PA. (2013). Violence resistance in Xinjiang (China): tracking militancy, ethnic riots and ¿knife-wielding¿ terrorist (1978-2012). *Current History Online* , (30), 135-149.

Romaine, S. (1994). *Language in society: An introduction to sociolinguistics*. Oxford University Press.

Rong, M. A. (2009). The development of minority education and the practice of bilingual education in Xinjiang Uyghur Autonomous Region. *Frontiers of Education in China*, *4*(2), 188-251.

Solomon, M. (1978). Jimmy Carter and PlaYboY: A sociolinguistic perspective on style. *Quarterly Journal of Spee*ch. *64*(2), 173-182.

Stamou, A.G. (2011). Speech style and the construction of social division: Evidence from Greek television. *Language & Communication, 31*(4), 329–344.

Sun, M. (2012). Xinjiang Weihan Yuyan Jiechu De Shehui Bianliang Fenxi —— Yi Yiningshi Wei Lie. *Journal of Beifang University of Nationalities*, (2), 29-34.

Tagliamonte, S., & Roberts, C. (2005). So weird; so cool; so innovative: The use of intensifiers in the television series Friends. *American Speech, 80*(3), 280-300.

The Economist. (2011). *Let them shoot hoops*. Retrieved July 21, 2019, from https://www.economist.com/asia/2011/07/30/let-them-shoot-hoops

The Guardian. (2014). *Urumqi car and bomb attack kills dozens*. Retrieved July 21, 2019, from https://www.theguardian.com/world/2014/dec/08/eight-sentenced-to-death-xinjiang-attacks-china

The State Council Information Office of the People's Republic of China. (2021). *Xinjiang Population Dynamics and Data*. Retrieved June 5, 2022, from http://www.scio.gov.cn/zfbps/32832/Document/1713594/1713594.htm

Trotta, J. & Blyahher, O. (2011). *Game done changed*: A look at selected AAVE features in the TV series *The Wire*. *Moderna Språk, 105*(1), 15–42.

Tsiplakou, S. & Ioannidou, E. (2012). Stylizing stylization: The case of *Aigia Fuxia*. *Multilingua*, 31(2/3), 277–299.

Uyghur (2020, December 15). Ethnologue: Languages of the World. Retrieved July 21, 2019, from https://www-ethnologue-com.proxy2.cl.msu.edu/language/uig

Wang, Y. (2012). *新疆少数民族汉语教学理论与实践研究* [The Theory and Practice of Chinese Language Education for Ethnic Minorities in Xinjiang]. 新疆教育出版社.

Weller, T. (2012). The information state: An historical perspective on surveillance. In Ball, K., Haggerty, K. & Lyon, D. (Eds.), *Routledge handbook of surveillance studies* (pp. 57-63). Routledge.

Willemyns, R. (1997). Language Shift Through Erosion: The Case of the French-Flemish'Westhoek'. *Journal of Multilingual and Multicultural Development*, *18*(1), 54-66.

Winfield, B. H. (2012). Censorship. *Oxford Bibliographies*. Retrieved July 3, 2019, from https://www.oxfordbibliographies.com/view/document/obo-9780199756841/obo-9780199756841-0020.xml?rskey=9cJ6Zu&result=3&q=censorship#firstMatch

Wittern-Keller, L. (2008). *Freedom of the Screen: Legal Challenges to State Film Censorship, 1915-1981*. University Press of Kentucky.

Wozniak, A. M. (2015). River-crabbed Shitizens and missing knives: A sociolinguistic analysis of trends in Chinese language use online as a result of censorship. *Applied Linguistics Review*, *6*(1), 97-120.

Xinjiang. (2023, April 12). In *Wikipedia*. Retrieved July 4, 2019, from https://en.wikipedia.org/wiki/Xinjiang

[肥州]. (2018). 易中天 贺卫方 秦晖 吴思 陈丹青 舒国治 想象下一个十年[完整版]（理想国沙龙 2010）[Yi Zhongtian He Weifang Qinhui Wusi Chendanqing Shuguozhi imagine the next decade [full video] (Ideal Country Salon 2010)]. *YouTube*. Retrieved December 12, 2021, from https://www.youtube.com/watch?v=wK_sdRR3MAE&t=2112s&ab_channel=%E8%82%A5%E6%B4%B2

Zeng, J. (2020). Artificial intelligence and China's authoritarian governance. *International Affairs*, *96*(6), 1441-1459.

曹湘洪, & 王丽. (2009). 多元文化背景下的语言选择——以乌鲁木齐市城市居民为例 [Language choice in multi-cultural environments - take Urumqi citizens for an example]. *云南师范大学学报: 哲学社会科学版*, *41*(6), 16-23.