6-30-2022

The Making of Microcelebrities on *Douyin:* Multimodal Design and Online Identities in Informal English Instruction Videos

Yaqian Jiang
University of South Florida

Scholar Commons Citation

The Making of Microcelebrities on *Douyin*:

Multimodal Design and Online Identities in Informal English Instruction Videos

by

Yaqian Jiang

A dissertation submitted in partial fulfillment
of the requirement for the degree of
Doctor of Philosophy
Department of World Languages
College of Arts & Sciences
University of South Florida

Major Professor: Camilla Vásquez, Ph.D.
Wei Zhu, Ph.D.
Amanda Huensch, Ph.D.
Brandon Tullock, Ph.D.

Date of Approval:
July 8, 2022

Keywords: Self-branding, online identity construction, multimodal analysis, social semiotic approach, authenticity, commodification

DEDICATION

I dedicate this dissertation to my beloved grandpa, Xianrong Wu, who inspired my

passion for learning and is constantly supporting me with his unconditional love.

ACKNOWLEDGEMENTS

First and foremost, I would like to express my sincere gratitude to my major advisor, Dr. Camilla Vásquez. Without her guidance and dedicated involvement, this dissertation project would not have been possible. I am extremely grateful for her invaluable support, patience, inspiration, and encouragement during my Ph.D. study. She saw my potential when I did not see it myself and helped me grow as a researcher with her immense knowledge and plentiful experience more than words can describe.

I would also like to extend my heartfelt thanks to my committee members, Dr. Amanda Huensch, Dr. Wei Zhu, and Dr. Brandon Tullock for their valuable time and thoughtful feedback throughout the process, and Dr. Ippokratis Kantzios for chairing the dissertation defense.

I am fortunate and grateful to be the recipient of the Dissertation Completion Fellowship grant from the University of South Florida. This grant played an important role in helping me complete my dissertation and graduate on time.

My appreciation also goes out to my LALS friends and colleagues, particularly Antonella Gazzardi and Özge Güney, for their moral support. Special thanks to my friend, Liz Sanches, for proofreading my transcripts, for providing her cultural insights that helped me analyze my data, for listening when I needed to vent, and for cheering me up when I felt overwhelmed. I am thankful to Dr. Judith Bridges, supervisor of undergraduate linguistic courses, for her generous help and continuous support of my teaching and my professional growth. Thanks should also go to my students who I had a great pleasure working with. I feel blessed for having the opportunity to grow with them together.

Finally, I also wish to thank my family for their unfailing support. I am grateful to my husband, Dr. Ning Wang, for his understanding, love, and care during the compilation of this dissertation.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

As the online education market continues to expand globally, teaching and advertising have converged in digital spaces in unprecedented ways. Driven by the attention economy, independent online educators must compete to cultivate a consumer base. To attract followers, they must engage not only in online teaching activities but also in self-branding practices. This dissertation explores how microcelebrities who teach English on *Douyin* utilize linguistic and other semiotic resources for both instructional and self-branding purposes.

To address this topic, I collected 84 informal English language instruction videos created by 14 different *Douyin* microcelebrities. Taking a social semiotic approach to multimodal discourse (Jewitt, 2009, 2016; Kress, 2010), this study conceptualizes short videos as multimodal ensembles in which different modes (e.g., speech, text, image, sound, layout, color, gesture) of communication interact to generate meaning as an entire unit. The social semiotic approach emphasizes the agency of content creators, who take on the role of "designers," selecting from available modal resources with a keen awareness of the interactional requirements of the digital environment in which their artifacts are situated.

My analysis shows that the *Douyin* microcelebrities self-promote through constructing a multifaceted, hybrid identity that is consistently represented in their videos and profile pages. By strategically exploiting multimodal resources available to them, these content creators craft popular and appealing English instruction videos, while at the same time, promoting themselves and attracting more followers on *Douyin.* The study identifies strategies for creating engaging online learning/advertising content for different audiences with media technologies. It also adds

to a growing body of research on digital genres and authenticity in digital self-branding. Furthermore, the project sheds light on the broader media landscape and user-generated participatory culture in China.

CHAPTER ONE: INTRODUCTION

China's Short Video Industry and Microcelebrities

China has undergone an explosive growth in the number of mobile internet users in

recent years. According to the *Statistical Report on Internet Development* in China released from

China Internet Network Information Center (CNNIC, 2020), the proportion of Chinese online

users accessing the internet through cellphones takes up 99.3% of the total number of the 897

million netizens in March 2020, far exceeding the proportion of users of desktops, laptops, TVs,

and tablets. The rapid growth of mobile netizens in China can be attributed to two major reasons

(Li, 2019). First, the major telecommunication enterprises have invested tremendously in the

construction of mobile network infrastructure nationwide in order to increase mobile network

coverage and to lower the cost of internet data plans. Second, many Chinese phone brands have

gained growing favor with consumers, providing affordable domestically produced smartphones.

With nearly all Chinese internet users accessing internet via their mobile phones, China has been

witnessing a surge in the popularity of short videos apps since 2016, which take up a growing

share of users' online data and time. As of March 2020, China's user population accessing short

videos was 773 million, accounting for 85.6% of all netizens. Some popular short videos apps in

China include *Douyin*, *Kusaishou*, and *Xigua Video*. As Li (2019) pointed out, in addition to the

widespread availability of smartphones and the considerable expansion of the 4G mobile

internet, Chinese internet users' increasingly fragmented time online is another contributing

factor to the spectacular growth in consumption of short videos in China. According to a survey

conducted by the China Youth Daily, over 76% of participants aged between 18 and 35 believed

1

they were living an accelerated life because of heavy workloads and the information age ("Over 76% young Chinese", 2019). With a fast-paced life, leisure time can be quite scattered and tends to be split into small chunks. As a result, the time spent online for leisure becomes increasingly fragmented. Since short videos are mainly played on mobile devices, they are suitable for viewing during small chunks of time with relatively few restrictions on location, such as the time spent waiting in line, on the bus, taking small breaks, and so forth. Additionally, online users are frequently multitasking when it comes to consuming content as well. Chatting with friends and checking emails are some of the everyday tasks they perform. These activities reduce their attention threshold, which also leads to their preference for shorter videos over longer ones, which require a greater commitment of time. A short video that provides concise, direct content is much more likely to be consumed in its entirety. Furthermore, the simple and easy operation of short videos apps has also contributed to the growing consumption of short videos. In most cases, after entering the applications where the short video occupies the screen, viewers only need to swipe up to go to the next video.

Using tracking technologies and artificial intelligence, short video platforms constantly collect data on their users and then recommend and deliver videos that are likely to be of interest to that particular viewer. The recommendation algorithm is built upon the labels assigned to each user in the backend of these short video apps, according to the users' browsing histories and operational behavior. This personalized distribution of short videos motivates users to continue viewing the content. In addition to the use of distribution personalization mechanisms to maintain users' interest in the content, short video platforms also deliver popular videos and videos of different content and categories to keep the viewing experience fresh.

From the perspective of video content creators, it is easier to manage and produce videos in short forms than those in longer forms. Short videos offer more flexibility in content and style and tend to focus on a single topic. As such, video creators have more opportunities to experiment with different styles and concepts, providing their viewers with variety while still delivering information and entertainment. In addition, the production of short videos requires less time and fewer professional skills in video shooting and editing. The lower threshold for video making has attracted numerous grassroots content creators. In fact, the creators of short videos are mostly ordinary people, who are both content producers and consumers. This dual role users play is also referred to as *prosumers*, which is a product of the blurring distinctions of production and consumption brought by the participatory affordances of contemporary media technologies. The content production mode, a combination of the Professional Generated Content (PGC) and the User Generated Content (UGC) encourages users to create while watching videos (Xiao et al., 2019). Specifically, ordinary users are encouraged to make their own versions of popular videos through imitation and adaption, which leads to increased user interactions and engagement in online "participatory culture", a culture which sets "relatively low barriers to artistic expression and civic engagement" and provides "strong support for creating and sharing one's own creations" (Jenkins, 2009, p. 3). In contrast, longer videos are less appealing to both viewers and creators due to the demand for more content and formalized structure, as well as more labor and professional skills in making. However, it is worth noting that the duration of a short video is vaguely defined as each short video sharing platform has their own rules concerning time duration. Typically, the length of short videos is less than 5 minutes.

The rapid growth of short video applications in China has contributed to tremendous economic gains of China's short video industry. According to Statista[1], a statistics portal providing market and consumer data, the revenue of China's short video market hit 130 billion yuan (approximately 20 billion USD) and was projected to amount to 386 billion yuan (approximately 59 billion USD) by 2022. Some main ways through which short video apps gain cash flow include paid advertisements and the commission fees it charges content creators. Content creators can earn cash through means such as 1) paid endorsements, 2) driving traffic to their personal store, an embedded e-commerce feature, where they sell the products and services of third-parties or their own commodities, 3) live streaming to advertise products as well as to receive virtual presents and donations from followers, and 4) even selling their vetted accounts which have amassed a considerable number of followers.

Despite the multiple ways that content creators can make money on short video platforms, one thing is certain: they need to grow their fanbase before they can earn any "serious" money. As such, perhaps not surprisingly, monetary gains are the strongest motive for content creators who aspire to garner internet fame and accumulate large audiences, or in other words, who desire to become "microcelebrities" (Senft, 2008). The advent of social media has led to the creation and growth of microcelebrities worldwide. Although the emergence of microcelebrities can be traced back to the late 1990s on Bulletin board systems (BBS) as the dominant social networking platform at that time, microcelebrities in the Chinese digital context are a relatively recent phenomenon as 2016 marks the surge of microcelebrities in China (Li, 2020). Since then, Chinese microcelebrities, with the rapid expansion of the short video and livestreaming industry, have achieved tremendous commercial successes (Li, 2019).

---

[1] https://www.statista.com/statistics/874562/china-short-video-market-size/

This recent rise of microcelebrities in China has inspired many ordinary people to seek fame and fortune on social media sites. The online practices that online users engage in to amass followers are pursued in response to the "attention economy", a marketing concept that determines economic values based on the amount of public attention one attracts (Fairchild, 2007). The number of followers as well as "likes" and comments received for a piece of content all serve as highly visible metrics of social media success (Marwick, 2015). Knowing that the amount of viewers' attention they captivate has a direct impact on their popularity and income, microcelebrities and "wannabe" microcelebrities often find themselves engaged in fierce competition, using social media to create continued streams of content to maintain their audience's interest as best as they can. As will be elaborated upon the literature review section (Chapter 2), becoming a microcelebrity also requires self-commodification to create a public persona. A carefully crafted and maintained online image is used as an essential technique to attract attention and publicity.

Douyin and *Douyin* Short Videos

Douyin (known as *TikTok* internationally) is a social network application for user created short-form mobile videos. First launched in the Chinese market in September 2016, *Douyin* has quickly become one of the most popular Chinese short video applications. According to the 2019 Douyin data report[2], the number of Chinese domestic Daily Active Users (DAU) of this fast-growing video-sharing internet application soared to over 400 million on January 5th, 2020.

Seeing the tremendous success *Douyin* has gained in China, its Chinese parent company *ByteDance* soon decided to expand its international market. After *ByteDance* acquired *Musical.ly*, a popular lip-syncing video app in the United States but founded in China, and

---

[2] Retrieved from https://cbndata.com/report/2168/detail?isReading=report&page=1&readway=stand

merged it with *Douyin*, the company launched *TikTok*, the international version of *Douyin*, in the United States in August 2018. User accounts of *Musical.ly* were consolidated into *TikTok* accordingly. By 2019, *TikTok* had accumulated 1.5 billion downloads on the App Store and Google Play and was the third most downloaded non-gaming app of the year worldwide. In addition to the massive traction *TikTok* has gained among American teenagers, the application's popularity has expanded around the world including Europe, Brazil, and India[3]. According to the U.S. based application analytics firm *Sensor Tower*, *TikTok* overtook *WhatsApp* and *Facebook* to become the most downloaded non-game application globally in February 2020.

Despite the enormous popularity *Douyin* and *TikTok* have enjoyed, the business model of these two versions of the application are quite different between the Chinese and the U.S. markets. While the majority of videos on *TikTok* primarily focus on lip-syncing and entertainment, there is a much wider range of content on *Douyin*. Topics such as major life events, travel, everyday life, food, and traditional culture are just a few among many popular categories. To explore the types of content that *Douyin* users frequently engage in, Lu and Lu (2019) conducted 28 semi-structured interviews with active users. Their analysis showed some of the content that the interviewees enjoyed the most included videos showing kindness or positive thinking, of knowledge sharing, of science, of education and language learning, of arts and crafts, and of professional and social skills. On the contrary, videos with a focus on education, learning, professional skills, parenting and so forth are scarce on *TikTok* (Shutsko, 2020). One major reason for the discrepancies of content in these two applications is the different user demographics. While *TikTok* users are predominantly adolescents or Zillennials (Shutsko, 2020), *Douyin* has a more balanced user distribution across several age groups from younger

---

[3] TikTok was banned in India on June 29, 2020 as a result of a violent border dispute between India and China.

generations to middle-aged adults (CNNIC, 2020). Moreover, the number of users who are 60 years and older is quickly growing on *Douyin*, and a few of them have even gained tremendous popularity and amassed millions of fans ("A new target market", 2020).

With users from all age groups viewing and sharing content of diverse topics, *Douyin* has not only become a social media platform for sharing life and humor, but it is also being widely used to gain knowledge and skills. According to the 2019 *Douyin* data report, 14,890,000 knowledge-based videos were created and shared in that year. Many of these videos attracted numerous views. One of the specific examples the report provided was that 130 million Douyiners watched chemistry class videos made by a chemistry teacher. The data report also listed the top five most popular knowledge categories of videos on *Douyin*, which are cooking, language learning/teaching, subjects taught in school, professional skills, and science. In this project, my focus is on the English language teaching videos created by some popular Douyiners who may be considered microcelebrities (Senft, 2008). As a former language teacher who taught English in both China and the U.S. and as an applied linguist, I am naturally interested in content with a language pedagogical focus and produced in the English language teaching context of China. However, the most important reason for choosing the English language teaching videos on *Douyin* as the focus of this research is because English language teaching and learning is a very popular category of *Douyin* videos and has attracted enormous attention from users.

As a mobile application, *Douyin* is available on both Android and iOS. Similar to other short video apps, *Douyin* presents a simple user interface (see Figure 1) to allow an easy access to its content. Upon entering the application, what is presented by default is the *Douyin* feed called 推荐 (meaning: recommended) when the language of the app is Chinese and "For You"

when the app language is English[4]. The *Douyin* feed is mainly in the form of short videos and occasionally in the form of live streams, both of which are more or less personalized content recommended by the *Douyin* algorithm based on a combination of factors including video popularity, the user's previous viewing history, and their interactions with videos. A simple swiping up action on the screen is all that a viewer needs to do to watch the next video. The Tab "For You" appears at the top of the screen, together with another two tabs 同城 (meaning: same city) / "Nearby" and 关注 (meaning: follow) / "Following" to the left. Users can tap on "Nearby" to view the videos provided by users nearby or select "Following" to check the content created by people whom they are following. On the top left is the tab called "直播" (meaning: live streaming), which directs users to view live streams.

---

[4] As an active *Douyin* user, I noticed that recent versions of *Douyin* (including version 13.6.0 which was released on November 24, 2020, and later versions) no longer have English as one of the app languages. Previous versions allowed users to set either simplified Chinese or English as the language of the app. However, more recent versions have removed the language settings, and Chinese as the default language cannot not be changed.
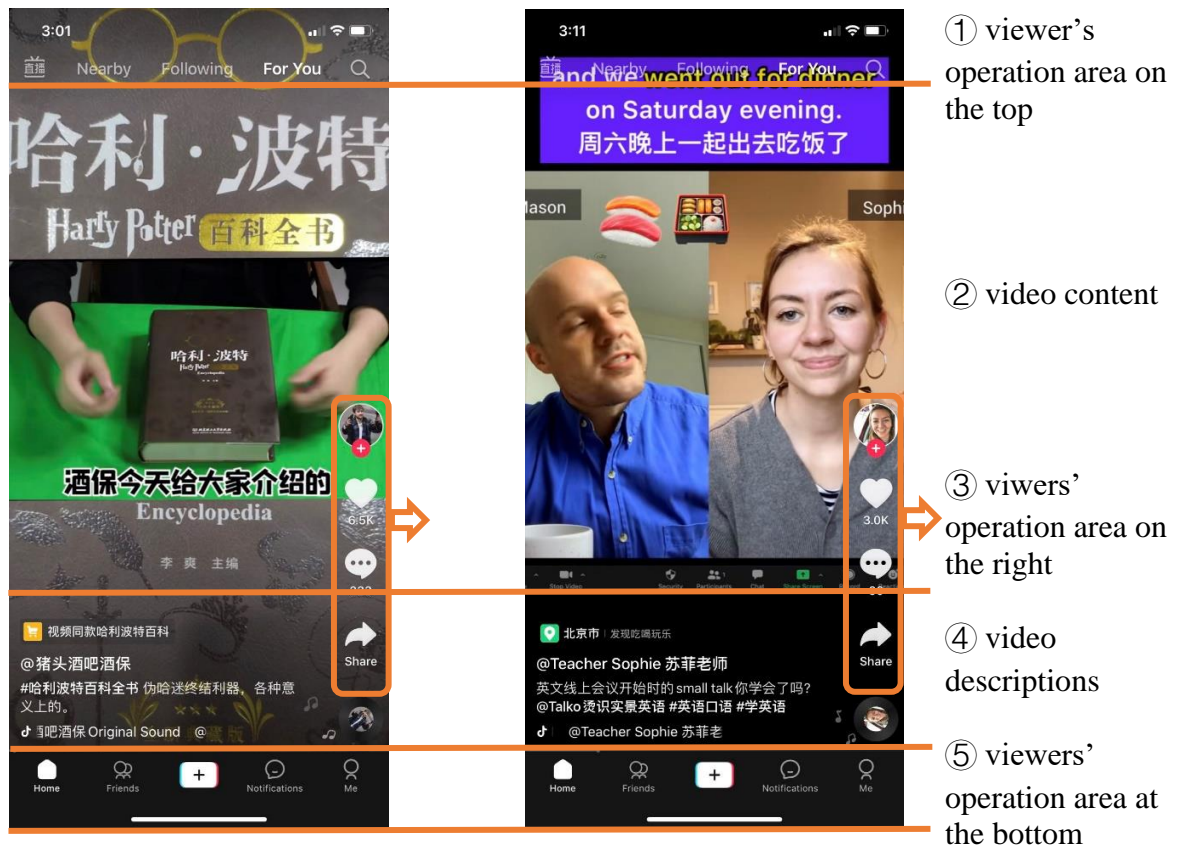
Figure 1. Examples of user interfaces of "For You" on *Douyin*

When viewing videos on *Douyin*, what occupies the center of the screen are the videos. Compared to traditional videos that are almost always presented horizontally on the screen, most videos on *Douyin* are presented vertically to accommodate people's viewing habits on cellphones and as a result of using cellphones as the device to shoot videos. Video description and other relevant information such as background music, location, and product links are placed near the bottom of the screen. On the right side slightly lower toward the bottom is a viewers' operation area, where viewers can see the profile picture of the content creator (with a red "+" icon displayed if the account is not yet in the user's following list), a "like" button they can tap to give a like, a "comment" button, and a "share" button arranged in a row. Another users' operation

area is at the bottom, where five tabs are displayed. The tab in the middle, which features a "+"

icon, is the feature that people can use to shoot photos and videos.

When users tab the "+" icon at the bottom, the interface for shooting and editing videos

and photos will appear (See Figure 2). The app-made videos can last 3 minutes maximum.

Videos more than 3 minutes need to be filmed outside of the *Douyin* app using a smartphone's

built-in camera or a different device such as a digital camera. In this case, videos can be as long

as 15 minutes. However, most videos on *Douyin* remain short in length, ranging from a few

seconds to 2-3 minutes, with the consideration of viewers' short attention spans. Compared to the

video length limit that *Douyin*'s built-in video editor sets, users' desire for better quality and

editing of the videos is a more important reason that they choose not to use *Douyin*'s built-in

video editor to create videos. Alternatively, digital cameras, video editing apps (e.g, *Jianying*,

Videoleap, *VUE*), and video editing software (e.g., Adobe Premiere Pro, CyberLink

PowerDirector) are welcomed by more professional content creators who have more strict

requirements for filming and editing. Right before the final step to upload the video, *Douyin*

Users can choose background music for their videos here.
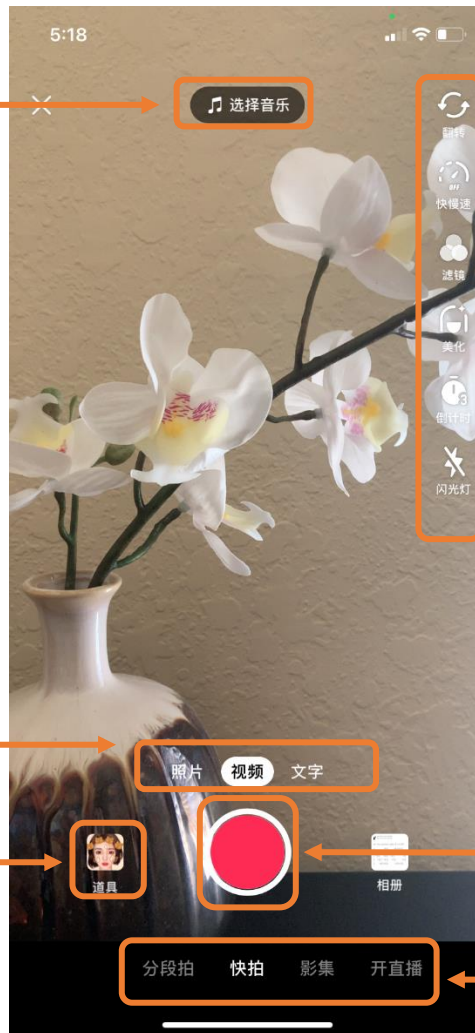
The control panel on the right includes features "Flip" that flips the camera, "Speed" that adjusts the video speed, "Beauty" that smooths lines and imperfections in one's face, "Filters" that changes the feel and color of the video, "Timer" that allows users to delay the shutter release and set the time to stop filming automatically, and "Flash" that turns on the flashlight on the phone.

content options (from left to right): photos, videos, text.

Add visual effects (e.g., beautifying appearances, changing backgrounds) before recording a video if desired.

Press the red button to start recording and press it again to stop recording.

Users can select to film videos with the pause/resume camera feature and of different lengths, create videos/photo slideshows using templates, or go live streaming.

Figure 2. Douyin's user interface of built-in video editor

provides users with options to add written descriptions, hashtags, and reference to other users using "@". Users can also add a thumbnail to the video, indicate their locations, and change the privacy setting for the video.

The myriad video filming and editing features available to users provides them with countless possibilities to design and create content attractive to viewers. However, not every

content creator is successful in amassing a large audience enabling her to turn her account into a revenue stream. As Marszałek (2020) pointed out, a *Douyin* user needs 1,000 followers to acquire the e-commerce feature for her account and at least 10,000 followers in order to qualify her account for "Influencer Personal Verification". The verified account will gain access to *Starmap*, *Douyin*'s internal platform that links influencers to commercial brands. Driven by the desire for financial success and meanwhile being aware of the competition for attention and microcelebrity status, many users are doing the best they can to create short videos that are attention-grabbing with the help of technology. Since English language learning has gained a crucial status in China, it is not surprising to find that there are many *Douyin* accounts that attract followers through their English language instruction videos.

English Language Mobile Learning in China

Digital technologies, such as mobile devices, affordable internet data, and high speed internet access along with various online platforms have created tremendous opportunities for internet users to reconfigure knowledge production and dissemination (Sauro & Zourou, 2017). The expanded repertoire of technology available to internet users enables them to transform how languages are taught and learnt in various settings. People are no longer confined to a designated space (e.g., classrooms) and time (e.g., language course meeting times) in which to learn a new language. Digital technologies allow learners with diverse linguistic, cultural, and socioeconomic backgrounds to learn a language, or "bits" of language (Blommaert, 2010) anywhere and anytime they like. The blurring of the distinction between everyday life and language learning, along with the penetration of mobile technologies into people's everyday lives, constitutes a key feature of the modern language learning landscape (Ho, 2019). Moreover, the outbreak of Coronavirus disease 2019 (COVID-19) in early 2020, has forced the sudden transition from in-

person classes to remote learning worldwide. As a result, there has been a significant surge in online learning and teaching on digital platforms since COVID-19.

With China's rapid economic growth in the recent four decades, learning the English language is seen as the key for global access and social mobility in China. On the national level, English proficiency has long been promoted as indispensable for the nation's modernization and internationalization, although English is seldom used for daily communication (Gao, 2012). On the individual level, proficiency in English can bring various economic, social, and educational opportunities for the betterment of personal well-being. For example, English has become a required skill that most large corporations and multinational companies are expecting of employees. As a result, many working adults seek ways to learn and improve their English in order to gain a competitive advantage. In addition, the desire to travel to English language speaking countries serves as another strong language learning motivation (Cai, 2017). As Merriam and Bierema (2014) suggest, the L2 motivation of adult learners originates from their hope for a better life in various aspects including career, lifestyle, and social networks. Moreover, proficiency in English is also seen as a passport to higher education at home or abroad and social prestige (Hu, 2005).

The past a few decades have witnessed persistent efforts to English language education in the formal education system. Policies on English Language Teaching (ELT) have been driven by the perceived importance of English to national modernization, a desire to catch up with developed countries, and a pressing sense of urgency (Hu, 2005). English language education has been a compulsory course in mainstream education from secondary schools to universities ever since 1978 and further expanded to primary schools in the new millennium (Lam, 2005). At

present, English has enjoyed the largest population of learners in comparison with other foreign languages taught in China (Xu, 2010).

Because of the escalating demand for a workforce with relatively high English proficiency, the Chinese government has not only invested in English education but also established various test systems and standards (Kang & Lin, 2019). The College English Tests (CET)-4 and -6, for example, are frequently listed as prerequisites in job advertisements. Another popular standardized test is the Business English Certificates (BEC), which is required by some companies for their employees to obtain a business-related English language qualification (Cheng, 2008). The Test of English for International Communication (TOEIC) is also well-known among Chinese working adults. The test is specifically designed to measure the everyday English language skills of people working in an international environment and has been extensively used for hiring and promoting in many foreign enterprises (Kubota, 2013). Moreover, international standardized tests of English language proficiency primarily designed for L2 speakers wishing to enroll in English-speaking universities, such as the Test of English as a Foreign Language (TOEFL) and the International English Language Testing System (IELTS), have been increasingly used in measuring the English proficiency of job applicants. As Kang and Lin (2019) argue, English language learning has become one of the most popular subjects in the lifelong learning curriculum of many people, and this phenomenon has led to the proliferation of English learning programs in both the public and private education sectors in China.

Despite the prominence national policies attach to English and the tremendous effort invested in English language education, there remain a host of issues with English language learning in China. One of the major concerns is that the professional competence of the teaching force is still relatively low and cannot sufficiently support the envisioned goals of English

14

language education. In addition, there is an acute shortage of trained teachers. Furthermore, the unequal distribution of and access to English teaching/learning resources between major cities and underdeveloped areas has widened the gap in the quality of English language education. Also, for adult learners in particular, finding time and opportunities to learn English in the traditional school education system can be quite challenging due to work and family responsibilities.

The rapid development of mobile technology in China is occurring at the right time to help learners resolve some of the issues with English language learning mentioned above. As Kang and Lin (2019) highlight, mobile technologies offers several important affordances for informal language learning that have attracted an increasing number of leaners. First, the portability and instant power-on capability of mobile technologies enables language learners to take advantage of their fragmented time. This is particularly useful for learners who are unable to dedicate large chunks of time to study on a daily basis because of family commitments, occupational responsibilities, and other social obligations. Second, the on-demand capability of mobile technology allows learners to take the initiative in identifying their learning needs, choosing materials and resources for learning, implementing specific learning strategies and so forth. For instance, it is completely learners' own decision to select which English language teaching videos to watch and which ones to skip, as well as how seriously they treat the content. Third, mobile technology accommodates multimodal learning preferences by presenting learning materials in different formats to appeal to a wide range of learning styles. As such, the mobile network makes it easy for physically isolated learners to virtually connect and interact with learners and experts around the world and to access unlimited learning resources.

In this project, I am concerned with online users' practices with social media technology in "designing" short videos of English language teaching. Although pedagogical short videos are hosted on popular video-sharing platforms, content creators take on the role of "designers" of the learning environment and the advertising message. Design, in a multimodal social semiotics sense, refers to "the situated process in which a sign maker chooses semiotic resources and possible arrangements for semiotic entities to be produced to meet particular social functions or purposes" (Jewitt, 2016, p. 73). Taking a multimodal approach, this study addresses the questions of how and for what pedagogical purposes the online content creators use the digital technologies and affordance of Chinese social media platform, *Douyin*.

Statement of the Problem

So far in this introduction, I have discussed the rapid expansion of China's short video industry and the explosive growth of microcelebrities in Chinese digital environments. I also introduced *Douyin* as a popular short video sharing app in China and the countless possibilities it affords users to design and create content, which subsequently leads to the rise of user-generated participatory culture and microcelebrity phenomenon in the attention economy. I also discussed English language teaching and learning in contemporary China and how digital technologies and the current COVID-19 pandemic have changed the Chinese ELT landscape. What consequently emerges, however, are some lacunae in the understanding how online users on Chinese social media sites attract followers through the teaching of English. I elaborate the gaps below.

Driven by the attention economy, microcelebrities on *Douyin* actively engage in various activities such as self-branding, constructing a public image, creating videos appealing to viewers' interests, teaching, and so forth in order to grow a fan base to monetize. The advancement of digital technologies allows users to draw on their entire linguistic and semiotic

inventories to create content that they believe to be most attention-grabbing to viewers. Consequently, it is not surprising that users develop their own unique style in designing and creating online content, which may help distinguish them from other online content creators and thus attract more followers. Nevertheless, thus far, only limited research focuses on the design of materials from a multimodal perspective which brings to the fore the fact that, in addition to language, there are other kinds of resources that content creators use to make meaning and to achieve their communicative purposes. In addition, the semiotic resources used in meaning making need to be understood in their historical and socio-cultural context where people draw on available linguistic and other semiotic resource in their environment. This is because meanings are "culturally made, socially agreed and socially and culturally specific" (Kress, 2009, p. 59). However, the multimodal meaning making practice on Chinese social media sites, particularly in the form of short videos, remains under researched, despite the rapid expansion of China's short video industry and microcelebrities. This gap is important to be addressed to understand some of the ways in which online users' multimodal and discursive practices are unique to the Chinese digital environment. Furthermore, how online users draw on multimodal resources to engage in self-branding and online identity construction practices in the specific socio-cultural context of China remains relatively unexplored. This area deserves scholarly attention because of its potential to contribute a better understanding of persuasive strategies appealing to Chinese online consumers.

Furthermore, the proliferation of short video applications in China has encouraged new forms of entertainment and information sharing practices. As Sauro and Zourou (2017) point out, while language learning in classroom settings and other formal instructional contexts have attracted a considerable amount of scholarly attention, relatively fewer empirical studies have

addressed online informal language learning beyond the physical boundaries of classrooms.

Previous research on the use of technologies for language teaching and learning has focused on

issues such as usability (see, e.g., Stevenson & Liu 2010), effectiveness (see, e.g., Macaro et al.

2012), and how digital technologies have been changing learners' learning strategies (see, e.g.,

Qian et al., 2018). However, few studies have addressed teacher agency or how language

teachers/users of a specific language interact with new technologies and their related affordances

to deliver language instruction, particularly through the form of short videos. As Ho (2020) calls

for, (short) videos as a form of language instruction deserve more attention from future research

as the affordance of (short) videos has the ability to influence how teaching and learning could

be done in the future. Moreover, Li et al. (2018) further emphasized the need to investigate

teachers' practices with technological affordance in order to better understand how and for what

pedagogical purpose the technology has been and will be used.

Purpose of the Study

The study reported upon here addresses the two gaps highlighted above through

qualitative analysis. Taking a social semiotic approach to multimodality (Kress, 2010; Kress &

van Leeuwen, 2001, 2006), I explore how content creators of popular *Douyin* accounts manage

and marshal multimodal resources and affordances of the platform in order to create English

instruction videos that attract numerous followers, which in turn enables them to become

microcelebrities. I first question how the content creators of pedagogical short videos of English

self-promote on the profile pages of their *Douyin* account. Through a content analysis, I next

explore the topics, language skills, and teaching styles/methods of the videos selected from some

poplar accounts. Finally, I interrogate the multimodal strategies adopted by content creators in

English language teaching videos for meaning-making and discursive practices situated in the

Chinese digital context. Specifically, I examine the linguistic and other semiotic elements integrated into online informal language teaching videos and how these resources are used for instructional and self-branding purposes. Moreover, I investigate types of modes used, their arrangements in videos as multimodal constructions, and how those modes and arrangements make meaning and possibly appeal to viewers. Additionally, I address the question of what modes and resources are most facilitated by the *Douyin* platform and what forms are less or least facilitated.

Overall, this study aims to investigate *Douyin* users' deployment of multimodal features for self-branding purpose through the teaching of English. Seeing the exponential growth of the Chinese short video industry and the rise of microcelebrities, I was inspired to explore Chinese microcelebrities interaction with social media technology and their use of semiotic resources. I then narrowed down my focus further to *Douyin* microcelebrities who provide informal English language instruction through short videos not only because of my linguistic and language education background, but more importantly, because English teaching and learning is a very popular category of content on *Douyin* that attracts millions of viewers.

Research Questions

In order to analyze multimodal meaning-making, microcelebrities' self-branding practices, and the design of informal English instructional videos in my dataset, I ask the following major research questions (RQs) and sub RQs in this study:

1. How do creators of informal language instruction videos on *Douyin* project themselves and advertise their products/services on their profile pages?

2. What kinds of English language topics or other topics are addressed in the videos and in what ways are these topics addressed?

a. What English language skills are being taught?

b. What content topics are discussed?

c. What are the teaching styles and methods used?

3. What multimodal resources are employed by content creators in their videos?

a. How do they use multimodality to construct English expert identities?

b. How do they use multimodal resources to communicate other personae that they present on the platform?

c. How do they exploit multimodal resources to teach English?

d. What multimodal strategies do they use to engage in self-branding and advertising?

In RQ 1, I address how content creators on *Douyin* project a carefully edited public image on their profile page to self-promote and advertise their products/services. In RQ 2, I focus on the content of sampled *Douyin* videos of English language teaching to understand what these videos are about. For example, what are the most discussed topics in these videos? Are there as many videos focusing on grammar as there are videos that center on pronunciation, vocabulary, situation/cultural aspects? In RQ 3, I examine content creators' selection and arrangement of multimodal resources to design instructional content and to self-advertise. In this question I seek to connect semiotic meaning making practices with online identity construction and self-branding in the Chinese digital context.

Significance of the Study

This study aims to make an empirical contribution to the understanding of Chinese multimodal digital discourse practices, in general, and to the multimodal construction of short videos on *Douyin*, specifically. Furthermore, the present research uncovers the widespread nature of self-branding and online identity construction practices amongst online users in China.

Simultaneously, this study takes up the call made by several computer-mediated communication (CMC) scholars (e.g., Ho, 2020; Kress, 2010; Li et al., 2018; Zhang & Cassany, 2021), who indicated that much more needs to be done to explore how constantly evolving technologies are used for purposes like education and advertisements, and for audiences in various socio-cultural contexts. Additionally, this study contributes to the current research on "participatory culture", a culture that provides "strong support for creating and sharing one's creations" and values individual contributions (Jenkins, 2009, p. 3).

CHAPTER TWO: LITERATURE REVIEW

The purpose of this chapter is to offer a comprehensive overview of research on multimodality, online identity construction, informal language learning, and commodification of English in online context. This will begin with an introduction to multimodality (Jewitt, 2009, 2016; Kress, 2010) as the overarching analytical framework.

Multimodality, Digital Technologies, and Design

Jewitt (2009, 2016) defines multimodality as an inter-disciplinary approach that attends systematically to all kinds of meaning-making forms (e.g., image, gaze, layout, color) that people use to communicate and represent as well as the interactions between these forms. From a multimodal perspective, language is just one of many communication forms and should not be viewed as the most significant *mode* of communication. Rather, all modes can contribute to the representation and communication of meaning and each mode has its distinct meaning-making potential. To identify whether a resource serves as a mode in a specific community from a social semiotic view, that resource needs to meet the following criteria listed by Kress (2009, 2010). First, the resources must be consistently used by people in that community to convey relatively fixed meanings. Second, the resource serves the *ideational function* of indexing meanings about states, actions, and event, the *interpersonal function* of representing meanings about relations of speakers, and the *textual function* of forming entities with internal coherence to convey message. Thus, each mode is shaped within its historical and socio-cultural context to realize social functions. Take typography as an example. Typography attends to the visual properties of written text including font size, font color, font direction, font style, bold, legibility (as a result of

combination of size, color and background) and so forth (Pauwels, 2012). To examine whether typography is a mode, one needs to consider in what context typography is used. Narrowing down to the discourse community of English language teaching/learning on the *Douyin* platform in the Chinese online context, typography meets the requirements for being a mode. It has been consistently used by content creators of English language teaching/learning in their videos to convey meanings such as information value and prominence. For the ideational function, using typographic means to highlight specific words in English subtitles helps frame the video as instructional, and delivering instruction is an action and event in the world. The information salience ascribed to the highlighted text also signals the social interaction between the knowledge disseminator and the learner, through which typography meets the requirement of the interpersonal function. Furthermore, the elements in typography are arranged coherently to convey meaning within their environment. The typographic resources used for an effect of emphasis function well in the videos of informal English instruction on *Douyin*. Since typography fulfills all the three functions in the given context, it is a mode. In short, a recourse becomes a mode when the community deploys it with "discernable regularity, consistency and shared assumption" about its meaning-making potentials (Kress, 2009, p. 59).

With the understanding that modes need to be examined within the community they are situated, whether a resource is a mode is not a question with one simple answer. The status of being a mode does not necessarily transfer from one cultural and historical context to another. I treat font size and font color as semiotic resources which the mode of typography can be realized as, while a typography artist might consider these two resources to be different modes. Similarly, one community may construct language as a semiotic mode that can be realized as speech and writing, but another community might treat speech and writing as two separate modes.

To some scholars of multimodality, modes can be classified based on the medium side into five "super modes" (i.e., visual mode, auditory mode, tactile mode, gustatory mode, and olfactory mode) to address five human senses, and each super mode includes a variety of sub-modes (Pauwels, 2012). For example, the visual mode as a super mode encompasses image, written text, layout, and so forth as sub-modes. From a social semiotic perspective, however, each mode differs in its distinct, culturally specific means of meaning making and therefore does not share an overarching category with other modes. For instance, although speech and writing are often placed under the label of language, the way they use affordances for making meaning is different. Speech is sequenced in time, which is linear. This organizing principle is common in all cultures. While writing is often displayed to imitate the linearity of speech, it is also organized by the logic of space to at least a certain extent (Kress, 2009). Furthermore, the meaning-making potential of the graphic component in writing can vary depending on cultural and historical context. Compared to alphabetic writing that represent sounds directly, Chinese written characters are mainly logograms that represent a word or a morpheme. There is inherent meaning in Chinese characters that is realized visually.

The different affordances (e.g., time, space) of modes lead to different ways for making meaning. As such, the distinct meaning-making potential of each mode has a tremendous influence on the sign-maker's choice(s) of modes in each specific communicative event. In addition to meaning-making potentials, modal preferences of societies, cultures, and social domains is another reason for the variation in choice of modes (Kress, 2010). In other words, the selection of modes is socially and culturally specific. Take how movies and TV programs are visually presented in China as an example. Most TV shows and movies have onscreen subtitles in Chinese. One reason is that subtitles help audiences understand regional dialects represented

in the shows and movies. On the other hand, subtitles also make the content more accessible to the Chinese speakers of different dialects because all Chinese dialects share the same writing system.

Although each mode of communication has its own means to make making, modal resources rarely function as separated entities. Instead, different modes are almost always meshed together in multimodal ensembles, in which they interact to represent and communicate meaning in its entirety. The selection and arrangement of modes in a multimodal ensemble is closely tied to the interplay between modes and is therefore essential for meaning making. Thus, modality analysis has a strong emphasis on the interplay between different modes. As Jewitt (2009) asserted, "the relationship across and between modes in multimodal texts and interaction are a central area of multimodal research" (p. 17).

The contemporary media landscape is characterized by the use of current media technologies, multimodality, and user-generated content. Digital technologies have enabled various modes such as sound, image, movement, and spatial arrangement, to be extensively used in individual information sharing and knowledge production in online space. When knowledge can be made by any user anywhere with the multiplicity of semiotic resources at one's disposal to realize meanings, it is not surprising to see a great variation in the choice of mode (e.g., how much writing, speaking, or visuals are deployed?) as well as of genre (e.g., is the content delivered through a lecture or a vlog?). The process of selecting mode and genre for a meaning-making purpose is closely related to the notion of multimodal design, which is defined as the act of presenting, realizing, and sometimes (re-)contextualizing social positions and relations, as well as "knowledge in specific arrangements for a specific audience" (Kress, 2010, p.139). Digital technologies provide new possibilities for multimodal design, allowing a great diversity

in the ways individual interact with different modes. Meanwhile, the design is conditioned by the affordances and constraints of each specific social media site. Furthermore, the multimodal design of content by individuals is situated within its historical, social, and cultural context and therefore may be useful to reflect social and communicative conventions.

When crafting pedagogical short videos, video makers have to analyze the affordances and constraints of delivering lessons using videos and make numerous decisions on how to maximize learning opportunities within the constraints of the video medium (Ho, 2019). Additionally, they need to carefully assess what semiotic resources should be used, as well as when and how, in order to better engage in self-branding and advertising practices to foster an audience and to earn revenue. Using the explanation provided by Tagg (2015), "affordances" refers to possibilities that users see of what they can do with technological tools. Depending on the users' technical skills, their previous experience using technologies of similar kind, and their purpose of communication, they may or may not use a specific technology and they may or may not use the same tool in the same way. For instance, in the case of videos, one content creator may use text superimposed on the video to deliver language instruction while another vlogger chooses to record a voice over for her video. On the other hand, "constraints" refer to what tools prevent us from doing (Ho, 2019). In the case of short videos, some constraints could be the lack of real-time and face-to-face interaction between video maker and audience, as well as the restrictions on the video length. Creators of online instructional content as designers of learning environments make meaning linguistically as well as semiotically, orchestrating various resources from their entire meaning-making repertoires. The selection of semiotic resources to create pedagogical videos plays an important part in the design of the multimodal learning environment and resource. As designers of an informal digital learning environment, video

makers of language instruction content not only need to gain an understanding of the affordances and constraints of the platform, but they also have to strategically integrate the multimodal resources that together maximize both learning and advertising effects. As Ho (2019) argued, online content creators need to carefully manipulate the modal intensity of different resources in different parts of the lesson by making certain aspects salient and others less so. In her study of a popular YouTube English teacher's video lessons, Ho observed that the transitions between language instruction and role play in the video are realized through the alternation between the intensity of embodied (e.g., gesture, eye contact, facial expressions) and disembodied modes (e.g., setting of the scene, placement of objects). Instead of using linguistic means, such as written text or spoken language to signal the start and the end of instruction or role play (e.g., stopping the lesson and stating "next you will see a role play"), the YouTube teacher observed by Ho ensures the smooth flow of the lesson through the strategic choice of modal intensity of various semiotic resources in different parts of the lessons. Norris (2004) argued that when analyzing multimodal interactions, communicative modes take on different levels of intensity depending on the nature of the communication. For instance, in a telephone conversation, the intensity of spoken language is the strongest as it is the mode that the interlocutor at the other end of the telephone directly reacts to. Digital technologies such as videos and websites offer a wide range of semiotic resources for content creators to select and deploy, leading to greater variation in the content, style, and aesthetics of sign-making practices (Adami, 2018). Online content creators of English teaching materials can use short videos as a means to construct a public identity, to establish their linguistic and pedagogical expertise by using a combination of semiotic resources, and to distinguish themselves from their competitors. In online spaces, particularly in the marketplace of video-sharing platforms, establishing a unique identity is the

key to captivate public attention, which helps content creators stand out from the crowd (Ho, 2019). The amount of attention, realized in the forms of number of followers, likes and comments, directly influences one's popularity and subsequently determines her value in the market. As such, successful self-branding and product promotions is dependent on the identity construction and representation that ride the tide of the attention economy. Similar to the multimodal design of pedagogical videos, content creators strategically select and manipulate semiotic resources to be used in their profile pages in order to maximize the self-branding and advertising effects. As shown in the subsequent analysis, profile pages are also utilized as an important means for users to establish a public identity, attract attention, and advertise their products and services.

Identity and Identity Construction

The notion of identity, according to communication scholars, is concerned with who people are (Tracy, 2002). In the fields of sociolinguistics and discourse analysis, there has been a dichotomy between the ways to conceptualize identity, which are often referred to as "essentialism" and "non-essentialism". According to Mary Bucholtz (2003), essentialism argues that

> the attributes and behavior of socially defined groups can be determined besides explained by reference to cultural and/or biological characteristics believed to be in inherent to the group. As an ideology, essentialism rests on two assumptions: (1) that groups can be clearly delimited; and (2) that group members are more or less alike (p. 400).

From an essentialist point of view, identity is relatively fixed because the characteristics shaping identity are viewed as culturally or biologically given in advance and are separated by

sharp boundaries. However, recent scholarship on identity has begun to shift its focus to address the performative perspective of identity, and therefore adopt a non-essentialist approach. In contrast to essentialism, non-essentialism sees identity as socially constructed, fluid, and changing from one discourse context to another. Sociologist Stuart Hall (1996) defines identities as:

> about questions of using the resources of history, language and culture in the process of becoming rather than being: not 'who we are' or 'where we came from', so much as what we might become, how we have been represented and how that bears on how we might represent ourselves (p. 4).

Two key concepts that help better understand identities are sameness and difference. Sameness refers to the perception that people who are identified as similar imagine themselves as a group, while difference concerns with perceived distinctions between individuals and othering (Bucholtz & Hall, 2004). It should be highlighted that sameness and differences are by no means pre-determined by nature. Rather, they are products of social interactions and subject to an individual's own interpretations. For instance, how people within a community classify themselves (i.e., avowed identities) may be completely different from the identity categories an observer imposed on them (i.e., ascribed identities), especially when the observer is not a member of the group. Even within a group, how each member identifies themselves can vary greatly from one another. Furthermore, depending on the time and context, people may choose to perform some specific identities, which might or might not be perceived as the same thing by the other party involved in the interaction. Therefore, despite the useful perspectives on identity that sameness and difference provide, the subjective nature of these two concepts has made them inadequate to address the power dimensions in identities (Bucholtz & Hall, 2004).

The choices people make to affiliate themselves with a particular group or distance themselves from that group is essentially driven by agency and power. As Bucholtz and Hall (2004) point out, "social grouping is a process not merely of discovering or acknowledging a similarity that precedes and establishes identity but, more fundamentally, of inventing similarity by downplaying difference" (p. 371). When people choose to obscure specific differences among members in order to establish a group, they usually do so in order to empower themselves. On the contrary, when distinctions are highlighted and different identity groups are formed as a result, these groups are in most cases hierarchically organized as a result of power inequities, which can influence individuals' identity choices. Such ideological ranking renders the more powerful identities normalized and default, which are described as unmarked. In contrast, other identities become marked and more recognizable. Distinctions based on how marked identities are different from unmarked identities therefore become socially evaluated to be deviant and inferior. For example, Hill's (1998) study on "Mock Spanish", which describes the practice of Whites incorporating Spanish language features into English for a jocular or pejorative effect, demonstrated that White as a race and the English language is normalized in the public space of the U.S., while other racial identities (e.g., Chicano and Latino people) and other languages are highly visible and are being constantly judged.

Scholars in applied linguistics share a general understanding that because identity is multilayered and complex, studies on a particular dimension of identity cannot be done without exploring other dimensions. For instance, when racial identity in the U.S. is the focus of a research study, there might need to be some discussion of gender and nationality identity dimensions, and how these different dimensions interact. However, relatively limited research has addressed the intersections of different identity dimensions (Block & Corona, 2016) with

only a few exceptions. For instance, Roux and Peck's (2019) investigation of an online

marketing campaign targeting young women in South Africa uncovered racial inequality hidden

under the commodification of gender empowerment of female identity. With a similar focus on

racial discrimination, Mortensen and Maegaard (2019) demonstrate that the representation of

Greenlandic people in Danish history and contemporary media as exotic and uncivilized is

realized through interweaving valorizations of race and sexuality, which is mediated by unequal

distribution of power between the colonizing nation of Denmark and the colonized nation of

Greenland. In addition to the multi-layered nature and complex intersections of various

dimensions, identity is also fluid, ephemeral, and formed in interactions.

To explore the interrelations between language and identity, Buchotlz and Hall (2010, p.

19-26) proposed a framework for the analysis of identities in linguistic interactions. Their

framework consists of five principles. First, 'the emergence principle': Identity is not a pre-

existing product, but a social and cultural phenomenon emerging during interactions. Second,

'the positionality principle': Identity work in an interaction encompasses different kinds of

stances, which refers to the social, cultural or interactional positions interlocutors take depending

on the identity they perform. The interactional stances that speakers express at a particular point

of the communicative event may quickly shift, alter, or be abandoned as the conversation moves

on. For instance, a customer may express her dissatisfaction with the price of a laptop as a

student, but then shift to a more positive stance toward the product because of its excellent

performance from a researcher's point of view. These interactional positions, despite their

transience, frequently appear simultaneously in a single interaction. As such, a comprehensive

understanding of identity is inseparable from the analysis of multiple positions exhibited in

discourse. Third, 'the indexicality principle': Identity is formed in interaction through explicit or

implicit associations between language use and social meaning, which are informed by the socio-cultural beliefs and values of speakers. Fourth, 'the relationality principle': Identity is intersubjectively forged in relation to other available identity positions and other speakers. Rather than being absolute, the social meaning of identity is relative, taking in consideration the dichotomous view of similarity/difference, genuineness/artifice, and authority/legitimacy. Fifth, 'the partialness principle': Identity is essentially situated in interactions and discourse contexts, and therefore is fluid and everchanging. Consequently, identity is partial rather than complete as the construction of it is always an ongoing progress.

There has been a shift in sociolinguistics from viewing language as a single entity that indexes different kinds of identities in a fixed manner to conceptualizing language as a repertoire encompassing linguistic and other multimodal resources for communication and identity construction (Rymes, 2010). Bucholtz and Hall (2004) further highlight that language is a complex semiotic system because of its capacity to convey both referential or semantic meaning and contextual or social meaning. For example, the contemporary slang term *poggers* is used to express excitement at the semantic level, but this expression is also associated with online gaming culture and therefore can be used to index an identity relevant to gaming.

From a social semiotic view, identity construction and representation is a process of realizing meanings of identity through the deployment of semiotic resources. The rapid development of social media in the twenty-first century has significantly changed the way people interact and communicate. The digital world has transformed language in many ways, some of which include the inventions of new vocabulary, genres, and styles; the integrations of multimodal affordances such as visual, aural, gestural, and spatial models; the increasing use of translanguaging practices that make meaning based on a holistic linguistic and semiotic

32

repertoire and so forth (Darvin, 2016). What accompanies the transformation of language in the digital era is the transformation of identity. When performing and negotiating identities in online contexts, linguistic and other semiotic elements are the two major resources people draw upon for identity production that is accomplished as the outcome of culture. Furthermore, as people's perceptions of their identities and relations to the world constantly shift depending on the international and social context, it is common to observe that online users come up with different ways to manipulate language to perform dynamic and multiple identities that are confined by "fixed localities, patterns or cultural traditions" (Darvin 2016, p. 526). Bucholtz and Hall's framework (2010) for the analysis of speakers' identity construction in discourse is particularly useful in exploring identity work in digital space because of its potential for investigating interactional stance-taking in digital environments. With an emphasis on semiotic associations between language and social categories, as well as on various positions and roles of interlocutors, Bucholtz and Hall's framework (2010) offers guidance on investigating multilayered and transient identities in online multimodal contexts.

In term of identity types, the discourse analysts Karen Tracy (2002) divided identities into three categories. The first kind, master identities, refers to those aspect of personhood that are presumed to be relatively stable and unchanging: ethnicity, age, national and regional origins. The second type of identity, interactional identity, describes specific roles that people take on in a communicative event in relation to other people involved in the interaction. Interactional identities may be formulated at different levels of abstraction: they may be formulated at the level of social roles (e.g., parent, student, doctor), or they may be formulated to address the particular discourse actions a person is doing (e.g., storyteller, listener, caller, call-taker). The third category of identity is personal identities, or relational identities, which included

33

personality aspects of self, the kinds of relationship people have with others (e.g., warm, condescending, hostile), as well as the stances people take. Although he uses different labels, the sociolinguist Don Zimmerman's conceptualizations of identity types are not significantly different from Karen Tracy's. In Zimmerman's (2014) framework Tracy's master identities are referred to as transportable identities, emphasizing the fact that some specific facets of self are visible to others and are carried from one social context to another. To address the different levels associated with Tracy's notion of interactional identity, Zimmerman further classified this identity type into discourse identities and situated identities. Discourse identities addresses the type of discourse activity projected and recognized by participants regarding what they are doing interactionally in a particular spate of talk. Discourse identities can shift and form layers in a conversation. For instance, when an English teacher starts her lesson by telling a personal story about herself, her discourse identity at that moment is storyteller. However, if she raises a question for her students during or after her storytelling, her discourse identity shifts to a questioner while she is asking the question. Zimmerman's notion of situated identities, on the other hand, refers to one's role regarding the overall communication event based on socially distributed knowledge. For example, although the English teacher shares her personal story and asks questions, the overall event the teacher engages in is teaching, during which her situated identity is a language teacher.

The identity framework used in this study adopts is a combination of Karen Tracy (2002) and Don Zimmerman's (2014) conceptualizations of identity types. I use the categories of transportable, situational, and discourse identities from Zimmerman, but I also include the personal identity category from Tracy in my framework because it allows me to examine the interpersonal relationships between microcelebrities and viewers, which tend to be an important

34

aspect of self-promotion. In addition, Ho and Tai's (2020) framework of expert identity of online English language teaching will be used as well. This framework encompasses three aspects: multimodal design knowledge, pedagogical knowledge, and linguistic knowledge. This framework will be used to explore how expertise in informal online English teaching is built on the orchestration of the three aforementioned aspects of knowledge.

Self-branding and online identity performance

Becoming a microcelebrity requires a certain degree of self-commodification to create a public persona. The microcelebrity self is carefully constructed and performed to attract attention and publicity. This type of identity construction practice is inherently tied to the notion of self-branding. Marwick (2013) defined self-branding as "primarily a series of marketing strategies applied to the individual" and "the strategic creation of an identity to be promoted and sold to others" (p. 166). Marwick (2013) further noted that self-branding is both a set of practices and a mindset, in which individuals regard themselves as a human brand and as salable commodities. In the digital era, it is no coincidence that the promotion of the self on a wide scale is heavily dependent on the features of internet communication technologies which provide affordable means of information distribution.

Internet technologies have made it possible for people to apply sophisticated branding strategies, and social media platforms have allowed individuals with internet access to present their positive images to the world at a relatively lower cost. For instance, a Twitter user may describe themselves using words "a happy scholar, educator, author" in their bio to establish a positive and professional self-representation. Moreover, the use of advertising and marketing terminology to describe a set of skills and tasks in a catchy and appealing way can also be widely applied to online self-branding work and is easily consumable by potential customers (Marwick,

2013). For example, several of the Douyiners who I have observed state "a TESOL certificate" on their profile page to demonstrate their qualifications for teaching English. The meaning of TESOL, which is an abbreviation for 'Teaching English to Speakers of Other Languages' may not be widely known to *Douyin* viewers if they do not specialize in a field related to language education. However, in those profile pages, no further information is provided to clarify what TESOL is. Instead, short and simple phrases such as "have a TESOL certificate" are presented to convey a sense that the content creator is professional enough to get certified and is familiar with terminology in their own field. The lack of explanation of what TESOL is helps create a sense of divide between language experts and learners, and therefore the language teaching skills indexed in the statement "having a TESOL certificate" are used to gain attention. Furthermore, coining slogans to describe oneself and emphasizing the positive is also an important technique of self-branding (Marwick, 2013). Take the expression "When life gives you lemons, make lemonade out of it", from the profile page of a *Douyin* content creator of English language teaching videos as an example. This catchy message not only conveys the Douyiner's optimistic outlook on life, but also indicates her competence in speaking and teaching idiomatic English. In addition to linguistic elements, social media affordances allow users to draw on various other semiotic resources to engage in meaning-making, self-branding, and advertising practices. Marwick (2013) identified a set of characteristics of a successful personal brand, which includes a distinct username, multiple social media accounts, the distribution of content using the internet, and the promotion of this content using social media. Self-branding involves not only creating an image of oneself, but also making that image visible to others. With the help of the internet, it is possible to disseminate personal information and content on an exponentially greater scale than in the past. Networked visibility allows practitioners to envision – and sometimes achieve –

tremendous fame and wealth on a national or even a global scale, which serve as the primary inspiration of many online users who wish to acquire microcelebrity status.

To seek public attention and achieve microcelebrity status, practitioners often engage in the labor of constructing specific identities as a unique selling point. One of the most important aspects of identities that microcelebrities present, according to previous research in self-branding, is authenticity. In a general sense, scholars' conceptualizations of an authentic identity centers on the characteristics of realness and genuineness (e.g., Duffy & Hund, 2019; Duguay, 2019; Khamis et al., 2017). A number of previous studies have investigated online users' self-presentation strategies in relation to authenticity. For instance, Berryman and Kavka's (2017) analysis of a popular YouTuber, Zoe Sugg, and her videos showed authenticity practices can involve establishing intimacy with viewers through broadcasting private life and displaying ordinariness as an everyday person. Focusing on the same YouTuber and her self-branding strategies, Jerslev (2016) presented similar findings that Zoe Sugg performs authenticity through enacting an intimate and ordinary persona in her blog posts and vlogs. The analysis explored intimacy in relation to Zoe's intimate way of addressing her followers and ordinariness along with the disclosure of private self using both linguistic and visual resources. Furthermore, Jerslev (2016) notes that performing connectedness and availability to viewers are also important techniques to establish authenticity. Taking a slightly different perspective, Duguay (2019) explored queer women's self-branding practices including posting videos and photos on Instagram and Vine. She discovered that authenticity can be performed through enhancing intimacy with followers, of which the practices include self-revelation of sexual identity as well as generating and sharing flirty and sexual content. In a similar vein, Raun (2018) examined the microcelebrity strategies a popular transgender YouTuber implemented in her vlogs. He argued

that the YouTuber's exposure of her private life (e.g., audio-visually documenting her medical procedures of transition) and her interaction with the audience through vlogs (e.g., direct address, close-up gaze, ask for advice and opinions) are some of the practices through which the transgender vlogger evokes intimacy, which is a key strategy of self-promotion. Nevertheless, different from some of the previous literature that conceptualized intimacy as a means to perform authenticity, Raun views intimacy and authenticity as two different microcelebrity techniques. To elaborate, Raun treats intimacy as the genre that is constructed and evoked through language and various other modal resources, and this genre is anticipated by the viewers. Meanwhile, he also conceptualizes intimacy as a capital relying on affective labor. However, Raun does not see that intimacy as a self-commodification strategy necessarily undermines authenticity. Instead, he argued that authenticity can be branded by performing a "public private self", a self with partial disclosure of personal life to the audience. In sum, research on self-branding practices has found that performance of microcelebrities rests on and features the construction of authenticity, intimacy, ordinariness, and constant presence and accessibility, despite being discussed in different terms and classified in different ways.

However, many previous studies on self-branding discovered that performers rely on an imaginary sense of what consumers desire to remake themselves into commodities, creating and performing identities exclusively designed for a specific online environment (e.g., Blommaert, Lu, & Li, 2019). Thus, the tension between self-promotion and authenticity seems apparent. Specifically, microcelebrities may find themselves trapped in a dilemma where they "struggle between presenting authentic selves to build followings, and cashing in on the resulting popularity which risks discrediting those authentic selves" (Tan et al., 2020). Marwick (2013) also added that the business-targeted self-presentation contradicts the notion of authenticity. To

be considered "real", one also needs to show a consistent self-presentation in front of others.

However, consistency can be notoriously difficult to sustain (Khamis et al., 2017). Many fame

seekers may be encouraged to engage in emotional labor (known as labor that regulate one's own

emotion to project a desired image) and self-surveillance (understood as the action of monitoring

one's own behavior with consideration of others' opinion) to ensure an appropriate branded

persona (Marwick, 2013).

In contrast to the rich literature in self-branding practices in Western-originated social

media sites, relatively limited research has investigated the online performance of

microcelebrities in other online platforms with only a few exceptions, one of which centers on

the South Korean context. In this study, Song (2018) explored the self-branding strategies of

broadcasting jockeys (BJs) on *AfreecaTV*, a popular South Korean live streaming platform,

implemented in their live streaming. Different from presenting the self as a real and

unpretentious person through portraying ordinariness and intimacy – as has shown to be the

common in social media platforms such as YouTube and Instagram – a great number of BJs on

AfreecaTV achieve their microcelebrity status by performing staged, exaggerated, aggressive,

and eccentric personae. Focusing on the microcelebrity phenomenon in the Chinese context, Li

(2019) also argued that not all online influencers seek fame through portraying themselves as

real people and revealing their private lives, especially when fashion, beauty, and lifestyle are

not the core topics of the online content they generate. Li's analysis of satirical and ludicrous

short videos created by Papi Jiang, a Chinese microcelebrity, reveals that Papi performs an

exaggerated and carefully designed persona in her videos in order to playfully present her

criticisms on the life of young urban generations in China. In addition, unlike many

microcelebrities who invest heavily in enhancing their connectedness with their followers, Papi

spends limited effort in maintaining close connections with viewers. Her interaction with her followers in the comment sections is extremely infrequent. Also, she explicitly acknowledged in her interviewers that she preferred to stay away from followers' comments as much as she could to avoid being affected by negative opinions. Still within the Chinese online context but narrowing the focus down to the self-branding practices on *Douyin*'s rival, *Kuaishou*, another popular Chinese short video platform, Tan et al. (2020) found many microcelebrities on this social media sites relied on exaggerated, abjective, and distasteful performance (e.g., putting one's face into a cow's rump) in their live streaming and videos to seek attention. This type of self-presentation was particularly attractive to viewers of lower class and socio-economically marginalized groups. Centering on online live streaming practices in particular, Blommaert et al. (2019) pointed out that some Chinese female streamers present themselves as flirtatious yet not too overtly sexualized in order to attract followers and fan donations. These identities are performed exclusively online and therefore, are edited rather than real.

Based on the existing literature, there seems to be some salient distinctions in the self-branding practices of microcelebrities in different social media platforms and in different genders. To better understand microcelebrities' online discursive practices, more research is needed to focus on various media sites that are situated in specific socio-cultural contexts. To address this lacuna, the present study provides a systematic analysis of identity construction of online English language teachers/users on *Douyin*. I consider self-branding techniques, including performing ordinariness, enhancing intimacy, and increasing accessibility as important features that contribute to an authentic persona. Ordinariness refers to being a real and an everyday person. Intimacy is associated with the practices of disclosing private information and staying connected with the audience. Accessibility means being available to the audience and

encouraging interactions between the viewer and the influencer. Furthermore, I view the construction and performance of expertise as another key strategy to portray authenticity in this present study. Since the analysis of the current research focuses on online English language teachers' microcelebrity practices, one important dimension involved in online identity work is the construction of a "real" English language teacher identity. I argue that demonstrating expertise in the English language is essential for content creators on *Douyin* to brand themselves as qualified and professional language instructors that viewers can trust.

Microcelebrity, influencer, and social media influencer marketing

The two notions of microcelebrity and influencer are similar. Coined by Theresa Senft (2008), the term microcelebrity refers to "ordinary internet users who accumulate a relatively large following on blogs and social media through the textual and visual narration of their personal lives and lifestyles, engage with their followings in 'digital' and 'physical spaces, and monetize their following by integrating 'advertorials' into their blogs or social media posts and making physical paid-guest appearances at events" (Abidin, 2016, p. 3). On the other hand, influencers, sometimes referred to as online/digital/social media influencers, are defined as "people who built a large network of followers and are regarded as trusted tastemakers in one or several niches" (De Veirman et al., 2016, p. 1). As both terms describe online users who have gathered large numbers of followers and actively engage in the promotion of brands and commodities, I will use microcelebrity and influencer interchangeably in this project. Marwick (2013) conceptualized microcelebrity as a continuum based on the ways of achieving internet fame, ranging from ascribed to achieved. Microcelebrity can be ascribed to people whose fame is recognized often through celebrity media such as paparazzi photos and user-generated internet memes. Therefore, ascribed microcelebrity is almost equivalent to a

41

mainstream celebrity and assigned a high level of status. They not only enjoy popularity, but also are treated with the celebrity-fan relationship of distance and aggrandizement. In contrast, achieved microcelebrity is "a self-presentation strategy that includes creating a persona, sharing personal information about oneself, constructing intimate connections to create the illusion of friendship or closeness, acknowledging an audience and identifying them as fans, and strategically revealing information to increase or maintain this audience" (Marwick, 2013, p. 117). These achieved microcelebrity practices rely on social media platforms and technologies to establish an online presence and to live a public life. The deployment of self-branding strategies is crucial for them to amass fans and to increase popularity. The *Douyin* microcelebrities that are the focus of the current project belong to the achieved microcelebrity type. The current study operationally defines a *Douyin* microcelebrity/influencer as any Douyiner who has at least 100,000 followers, a threshold presented by Marszałek (2020). In this project, I will look at some "successful" content creators of pedagogical videos of English and how they are using all the resources in their linguistic and semiotic inventory to attract followers. The operationalization of "successful" content creators of English instructional videos are the ones who have achieved *Douyin* microcelebrity status (i.e., who have garnered 100,000 followers) and who have amassed more followers than most of other microcelebrities in this genre (i.e., English language teaching videos) on *Douyin*.

Informal Language Learning

Research on language teaching and learning tends to classify learning into three major categories: 1) Formal learning, which involves certain types of language classes in the traditional school system, including both face-to-face classes and online courses; 2) Non-formal learning, which relies on professionally produced learning materials but takes place in a non-academic

context; and 3) Informal language learning, which "occurs naturalistically, using resources not specifically tailored for educational purposes and which are situated outside of any institutional context" (Socket, 2014, p. 11). Different from the first two types of learning, informal learning is unscheduled and impromptu, and it may take place with varying degrees of intentionality (Stevens, 2010). L2 learners who engage in informal language learning may or may not actively exploit the learning potential from prospective L2 resources or pay explicit attention to language acquisition. More often than not, language acquisition through informal learning is non-intentional and accidental rather than conscious and intentional.

One notion relevant to informal language learning is "out of class learning" (OCL) as both concepts attend to learning outside the classroom. Benson (2006) distinguished OCL from informal learning by pointing out that "out of class learning has been used somewhat narrowly to refer to the efforts of learners taking classroom-based language courses to find opportunities for language learning and use outside class" (p. 26). This approach focuses mostly on deliberate attempts to improve one's L2 rather than unconscious or incidental processes. Furthermore, the emphasis on the specific type of L2 learners who are enrolled in classroom-based language courses in an academic context tends to exclude the informal language learning practices of many other types of L2 learners and users. Examples of OCL of English are often situated in English as a second language (ESL) contexts rather than English as a foreign language (EFL) ones, as there are usually fewer OCL opportunities in EFL than in ESL contexts (Sockett, 2014).

In the context of *Douyin*, however, language learning content is often intentionally crafted by content creators rather than occur naturalistically. On the other hand, these content creators of English language instruction vary greatly in their professional expertise and credentials. Therefore, language learning on *Douyin* does not completely fit in any of the

43

reviewed learning types. In this dissertation project, I expand the term "informal language learning" to refer to learners' interactions with language learning content that is purposefully designed by both amateurs and professionals on the *Douyin* platform.

Online informal language learning

The rapid development of internet communications technology has provided language learners with new conditions in which language teaching and learning can take place. What has accompanied the fast-growing online services of recent years is an increasing number of people engaging in informal language learning activities outside of the classroom setting using the internet (Godwin-Jones, 2018). As pointed out by several scholars (e.g., Kubota, 2011; Sockett, 2014), L2 learner motives in accessing L2 materials and interacting with L2 communities in online contexts is not limited to learning, but also entertainment and leisure.

With a growing interest in digital technology-enhanced language learning and teaching beyond the classroom, scholars in the field of applied linguistics have investigated various major social media platforms to explore how learners utilize a wide range of available digital affordances to facilitate their beyond-classroom language acquisition. The social media sites that have been the focus of previous studies include YouTube (Arndt & Woore, 2018; Benson & Chan, 2010; Benson, 2015a, 2015b; Chik & Ho, 2017; Wang & Chen, 2020), Facebook (Sockett, 2014), Reddit (Isbell, 2018), and other digital contexts, such as gaming (Kuure, 2011), language learning social network sites (Chik & Ho, 2017) and specific local online discussion forums (Sun et al., 2017). As shown in the previous research, one major characteristic of online informal language learning is the great variety of approaches and materials internet users can adopt, such as interacting in the L2, watching instructional videos, and engaging in online learning communities. This type of learning practice is also referred to as *computer-assisted language*

learning "in the digital wilds", which is defined as "informal language learning that takes place in digital spaces, communities, and networks that are independent of formal instructional contexts" (Sauro & Zourou, 2017, p. 186).

Online videos as resources for learning

Learning via watching online videos as an informal language learning practice is gaining increasing popularity. In particular, the genre of video blogs (vlogs) is playing an influential role in the digital era (Ho, 2020). The practice of vlogging has been around for about a decade, and the purpose of vlogging is diverse. In general, vlogs can be considered simply as the video version of text-based blogging on the internet (Moor et al., 2010). However, vlogs with a purpose of teaching, which are referred to as "pedagogical vlogs", consist of a repertoire of genres and activities (Ho, 2019). Lange (2007) argued that the notion of "vlogging" should be treated as "an umbrella term that covers a wide number of genres, including everything from short video footage of spontaneous, real-life, personal moments, to scripted and preplanned 'shows' with characters, narratives, and professional acting" (para. 2). As one subcategory of vlogs, pedagogical vlogs are also a collection of genres in which instructional practices are multifaceted and fluid. For instance, the language instruction component in some videos is embedded in vloggers' activity of story-telling or sharing life experiences through the use of language and other semiotic resources. Some other pedagogical vlogs are more pedagogically oriented, such that the vloggers try to more or less re-enact classroom teaching through the medium of video, positioning themselves as teachers, and providing more explicit language instruction.

There are two features of pedagogical vlogs that distinguish this specific multimodal genre from other types of pedagogical videos. First, pedagogical vlogs are created by one or two

dedicated vloggers who post video lessons online regularly. The second characteristic is that there are frequent interactions between the vlogger(s) and audience through the comments section of the video-hosting sites (Ho, 2019). Despite the variations between different types of vlogs, vlogging bears the potential to reach a large audience through online users' active circulation and vloggers' promotion. According to Ho (2019), this affordance of vlogging warrants more research as it has the ability to influence how teaching and learning could be done in the future.

While previous studies that explore language learning practices in the wild are mostly learner-centered, focusing on learners' "self-directed learning, agency, and autonomy" (Sauro & Zourou, 2019, p. 2), few studies have addressed teacher agency or how language teachers draw on online multimodal resources to engage in various discursive practices, such as delivering language instruction, demonstrating expertise, and promoting services and products. However, this is important to better understand because the pedagogical use of technology can influence students' learning experiences and outcomes. With the ongoing technological advances and the distinct rise of online learning during the COVID-19 pandemic, the trend of online education is likely to continue even after the pandemic is over. Exploring how teachers can harness semiotic resources to design online content with a pedagogical focus can inform current and future teaching and learning practices that use technology. Moreover, the global expansion of the online education market and the burgeoning phenomenon of microcelebrities have the potential to impose a convergence of advertising and teaching in forms such as online videos. Therefore, exploring users multimodal meaning-making practices in online instructional videos can also allow us to better understand the self-branding strategies and online identity performance of microcelebrities.

Despite the necessity to understand online teachers' multimodal strategies for content design, only a few studies have investigated this under-researched area. One of the studies that addressed this research gap is Ho's (2020) analysis of a series of YouTube videos created by an online English instructor to teach language and non-verbal skills for job interviews. A close analysis of the multimodal design of these videos demonstrates that the YouTuber relied on her entire repertoire of multimodal resources, including spoken and written language, embodied modes (e.g., gestures, eye contact, facial expressions), and disembodied modes (e.g., screen layout, placement of objects), to create her video lessons. In a similar vein, Ho and Tai (2020) examined two online English teachers' strategic integration of multilingual and multimodal resources in their YouTube videos in relation to constructing expertise in online language teaching, which is manifested through the demonstration of multimodal design knowledge, pedagogical knowledge, and linguistic knowledge. Likewise, Yi and Jang (2020) observed the adoption of translanguaging practices and pedagogy in two English teachers' YouTube video lessons designed for Korean English learners during the COVID-19 pandemic. These studies suggest that teaching languages via creating online videos, particularly under the current COVID-19 pandemic that has forced the transition from in-person classes to remote learning, is gaining increasing popularity. As Ho (2019) points out, more research on online instructional videos is needed in order to explore how these online contents with a pedagogical focus can influence the current and future teaching and learning practices.

<center>**The Commodification of English[5]**</center>

Contemporary processes associated with globalization have contributed to a burgeoning focus on the commodification of languages in sociolinguistic research. However, the concept of "language commodification" still lacks a unified definition. Research examining language commodification tends to conceptualize language as a commodity in two (often competing) ways: 1) language as a valuable skill that is a saleable commodity itself, and 2) language as a symbol of authenticity that adds economic value to a product (Heller, 2010). The former strand of research has generated plenty of discussion, particularly on the commodification of the English language and associated ideologies in various geographical contexts. For instance, Relaño Pastor (2015) pointed out that the pervasive ideology of the commodification of English in Spanish/English bilingual programs in the Madrid region perceives English as a marketable skill that guarantees one's competitiveness in a multilingual European job market. Tan (2008) identified a similar ideology of English language commodification in the Malaysian context— English is highly valued for instrumental reasons, including access to jobs and knowledge in the era of globalization. Focusing on the industrial context, Rahman (2009) discovered that call center workers in Pakistan acquired varying degrees of near-native American or British English accents in order to pass as native speakers of English. The acquired accents were not only commodified and used as a sales strategy but were also regarded as correct and prestigious among call center workers to reproduce the language ideologies in Pakistan. Situating language commodification in a globalized age where an American-sounding variety of English is popularized, Blommaert (2010) argued that selling language courses in American accent is

---

[5] Portions of this section have been previously published in Jiang, Y. (2022). Teaching English on *Douyin*: Language commodification and translanguaging. In S. Li (Ed.), *Linguistic creativity and language commodification: English in China*. Routledge. And has been reproduced with permission from Taylor & Francis Group.

selling a package that consists of the language itself as well as its indexical meanings. Tupas (2008) took a step further to argue that the commodification of English is, in fact, the commodification of different English language proficiencies depending on "the varying monetary, instrumental and symbolic values attached to them by the market" (p. 98). The marketing of English in a globalized context are layered with two different scale-levels: one at a general level where the English language is purchased, and one at a smaller scale where a more specific set of "linguistic, pragmatic and metapragmatic skills" that index particular sociolinguistic registers are promoted to specific groups (Blommaert, 2010, p. 36).

When a language is conceptualized as symbolic capital or, in other words, a sign of authenticity, in the process of language commodification, it is either a language other than English, or it is a specific variety of English that is under research scrutiny. For instance, in tourist spheres, Leeman and Modan (2009) found that Chinese-language shop signs in Washington DC's Chinatown not only provide information about products and services to speakers of Chinese but also serve a symbolic function to represent Chinese culture and convey a sense of exoticness to non-Chinese visiting Chinatown. With respect to varieties of English, King and Wicks (2009) explored how Vernacular Newfoundland English (VNE) adopted in a car commercial was commodified as a means of promoting a product. They observed that the process of VNE commodification is associated with the stereotypes of Newfoundland culture as being remote and traditional. In a similar vein, Johnstone (2009) scrutinized the process of how the local speech of Pittsburgh was linked with Pittsburghese identity and culture through the production and consumption of T-shirts with printed Pittsburgh speech. Specifically, the production and circulation of Pittsburghese shirts provide opportunities for Pittsburgh speech to be disseminated and consumed, and subsequently enregistered as a recognizable dialect. The

indexical value of this dialect, as a representation of authenticity, is then further exploited and commodified on the T-shirts to add additional value to the product.

These two models of language commodification mostly remain parallel and seldom join hands in the same study. One major reason is that they lead to competition over "who defines what counts as [a] legitimate and commodifiable language" (Heller, 2010, p. 103). To elaborate, when language is perceived as an essential skill that leads to an individual's economic success, language is a commodity that people can purchase. Thus, anyone who speaks this language, whether it being as a first or second language, regardless of accent and dialect variations, can claim ownership of it. However, when a language is treated as a symbol, there is no extra value in the language itself. Rather, the language becomes associated with a certain way of being an indexical nature. In this case, the language or dialect is usually commodified on products as a symbol of authenticity to increase the value of the product. In this model of language commodification, those who are viewed as native speakers of that language or dialect are usually the ones who are considered eligible to evaluate if the language used on a product is authentic or not, thus excluding other users of that language from being legitimate speakers.

With the ongoing process of economic and cultural globalization, the English language as a lingua franca is widely exploited as symbolic capital (Bourdieu, 1991) and is increasingly commodified (Heller, 2010). Along with this trend is a growing emphasis on English learning in many EFL countries. China is not an exception. Like many other EFL contexts, the expansion of English language education in China is situated within neoliberalism, which is an economic and social doctrine asserting that free market is the only acceptable means to guide all human activities and that states should regulate and promote the market to ensure its prosperity (Holborow, 2015). Neoliberalism in China specifically, as Xiong and Yuan (2018) pinpoint, is

50

contextualized in a bourgeoning "'socialist market economy' where two-thousand-year long authoritarianism encounters individualism and market fundamentalism" (p. 103). As a result, neoliberal discourses treating English learning as a life changing opportunity associated with upward mobility is ubiquitous in English language education in China (Xiong & Yuan, 2018). Although English in the global process of commodification situated in such a neoliberal ideology is valued as a highly desirable skill that leads to economic profit in the market, English as a skill alone is seldom perceived to be sufficient in assigning the speaker value in the global market (Park & Lo, 2012). To make an individual's language skill marketable, the person needs to be able to understand and control "what counts as legitimate language and who counts as legitimate speakers of any given language" (Heller, 2010, p. 108). As a result, ELT in China has become involved in attempts to define what legitimate English is and who can be considered legitimate speakers of English. Considering the rapid growth of online informal English teaching as an increasingly common commercial practice on *Douyin*, the present study explores how *Douyin* influencers who present themselves as English language experts commodify and promote their own English and English skills through multimodal meaning-making strategies. The focus on the multimodal and discursive practices users engage in to assign value to their own variety of English is useful to better understand the processes of English commodification in the Chinese online context.

Authenticity and native-speakerism in English language education

The concept of authenticity and native-speakerism are useful in investigating how English is commodified in the context of *Douyin*, as they are closely tied to the questions of who is a legitimate speaker of English and what counts as "legitimate" English. The notion of authenticity in ELT research refers to being genuine and original (Buendgens-Kosten, 2014),

51

which "often carries either implicit or direct reference to the culture(s) in which the target language is used as a first language" and "generally gravitates towards 'native-speakers'" (Lowe & Pinner, 2016, p. 32). Specifically, when a speaker, language use, teaching methods, or materials are identified as authentic, what is implied is the concept of a *native speaker*, which is not only associated with one's linguistic and cultural background but can also be intrinsically linked to race. According to Lower and Pinner (2016), one common perception of an authentic English teacher is a white person who speaks English as their first language. The English of the "inner circle", typically British and American English, and their speakers, especially Caucasians, have often been endorsed as target models in English teaching in many EFL contexts (e.g., Wang, 2015; Yoo, 2014). The privilege of native speakers and their English contributes to the formation of a pervasive culture of native speakerism in these and other contexts (Wang, 2015).

Viewing English as an essential tool to achieve success in the ongoing process of globalization, many EFL countries including China and Japan attach great importance to English and English education. Along with the growing emphasis on English learning in these countries is the widely circulated ideology of native speakerism, which adopts a dichotomous view of teachers as "native speakers" who are culturally superior and "non-native speakers" who are culturally inferior and sees native speakers of English as ideal English speakers and teachers (Holliday, 2005, 2006, 2017). Consequently, acquiring "native-like" English becomes the ultimate goal of English education (Choi, 2016). Although this "neo-racist" ideology (Holliday, 2017) has been problematized in an expanding body of research by applied linguists for several decades, native-speakerism nevertheless remains entrenched in the field of English education. The negative perceptions of *non-native speaker* teachers shared by students and recruiters inevitably lead to professional discrimination. Previous studies have observed that the

marketability and employability of native English-speaking teachers are in stark contrast to those of non-native English-speaking teachers (Braine, 2010; Dervić & Bećirović, 2019; Selvi, 2010). This employment discrimination against those labeled as non-natives is associated with the perception of what accounts for an authentic speaker of English (Lowe & Pinner, 2016). From a traditional view, authenticity in ELT research is often judged to be something absolute and inherent, with little space for ambiguity or negotiation (Patterson, 2020). In the present paper, however, I adopt Lowe and Pinner's (2016) conceptualization of authenticity as "the way an individual sees themselves in relation to the various contexts in which they exist and are required to use language for the social production of meaning" (p. 32) in order to capture its dynamic nature. It should be noted that the notion of authenticity in an ELT context is different from the type of authenticity in identity performance and self-branding strategies of microcelebrities as discussed in previous section. Thus, there are two types of authenticity involved in this present study: 1) authenticity as genuineness, the "real" person in the sense of microcelebrities who perform connectedness, accessibility, affiliation, ordinariness, and intimacy, and 2) authenticity in ELT as a legitimate speaker of English. In the present research, I operationalize authenticity in ELT by looking at how *Douyin* microcelebrities of informal English instruction videos demonstrate their English competence and expertise in language teaching.

While previous studies have extensively investigated how the prevailing ideology of native speakerism as well as the discourse of nativeness and authenticity affects EFL teachers' professional identities (Braine, 1999; Houghton & Rivers, 2013; Medgyes, 1994), limited attention has been given to how English language users and learners have negotiated the ideology of native speakerism and their identities in relation to the construction of "non-native" speakers as deficient and inferior. The current project aims to address this research gap by

exploring how Chinese microcelebrities on *Douyin* who create English language instruction videos construct and negotiate their public identities as English language users/teachers/experts/learners in relation to native speakerism. In so doing, the research is useful in revealing users' perception of prevailing ideologies of English and how these ideologies are reinforced or contested in multimodal practices in this Chinese digital context.

Translanguaging in digital contexts

The term *translanguaging* was first coined by Welsh educationalist Cen Williams to refer to the strategic use of both Welsh and English as a pedagogical practice in classroom settings (Williams, 1994). Since then, the notion of translanguaging has been extensively researched in classroom contexts (e.g., Creese & Blackledge, 2015; García & Li, 2014) and in the everyday language practices of bilingual/multilingual speakers (e.g., Li, 2016, 2018; Otheguy et al., 2015). Translanguaging differs from code-switching in that it does not simply describe the process of alternating between two or more linguistic codes but also captures the more fluid and flexible nature of multilingual speakers' language use (García & Li, 2014). From a translanguaging perspective, speakers' language practice is based on a unitary linguistic repertoire rather than several independent language systems. Therefore, the features that translanguagers adopt and deploy from their holistic repertoire for the meaning-making practice cannot be easily sorted into a named language by its traditional definition (García & Li, 2014). Furthermore, García and Li (2014) noted that in addition to written and verbal languages, other semiotic resources (e.g., image, sound, gestures) are also an essential part in the making of one's translanguaging practice. This is particularly evident in online environments, where multimodal content has become ubiquitous. Digital technologies and affordances provide online users with countless opportunities to make full use of their linguistic and multimodal resources as a totality for

translanguaging practices. At the same time, translanguaging in online space illustrates the notion of "linguistic superdiversity" coined by Blommaert and Rampton (2011) to describe the fluid, complex, and unpredictable nature of the multilingual environment of the internet.

Many studies on translanguaging have explored how it is taken up in practice in various contexts. Pedagogically, translanguaging has been shown to be an effective teaching/learning method, as it not only facilitates learners' understanding of the subject matter by encouraging them to use their entire linguistic repertoire, but it also engages and empowers students through identity construction and alteration of the power relations between teachers and students (Creese & Blackledge, 2015; García & Li, 2014). Furthermore, in both educational settings and everyday contexts where translanguaging practices are often more involuntary and spontaneous, translanguaging is found to challenge monolingual ideologies by subverting language hierarchies and assigning more value to the non-dominant languages (García & Li, 2014; Li, 2016, 2018; Otheguy et al., 2015). Translanguaging is also an important resource for linguistic creativity in both online and offline contexts (Lee, 2015; Li, 2016, 2018; Song, 2018; Zhang & Ren, 2020). For example, in linguistic landscapes studies, scholars have identified translanguaging in shop signs to communicate to multilingual consumers and attract their attention (Cook, 2013; Jufferman, 2013; Song, 2018). In terms of informal online language teaching contexts, Ho and Tai (2020) observed that a translanguaging pedagogy was adopted in online English teachers' YouTube videos as a strategy to establish instructors' expertise. Similarly, Yi and Jang (2020) identified several translanguaging practices in ELT videos on YouTube, including the use of subtitles in two languages, mixing languages in a sentence, and transliterations. The present study, with its particular focus on an informal online language teaching context, continues to

55

enrich the discussions of translanguaging as a pedagogical approach as well as a strategy of

language commodification.

CHAPTER THREE: METHODOLOGY

A Social Semiotic Theory of Multimodality

The overarching methodological framework used to analyze the profile pages of the *Douyin* accounts and the pedagogical short videos in this project is a multimodal analysis. According to Blommaert (2005), discourse comprises "a general model of semiosis, i.e., meaning symbolic behavior" (p. 2). Traditionally, discourse analysis has attended primarily to linguistic resources (Fairclough, 2001; Kress & van Leeuwen, 2006). With the advent and wide application of digital technologies, the use of various modes of discourse for communicative purposes has become more salient and led to a growing scholarly attention. Following this trend, the subject of discourse analysis has expanded from linguistic texts to images, color, sound, typography, gazes, facial expressions, gestures, spatial positioning of people or objects inside the image, and so forth (Jewitt, 2009; Kress, 2010). A multimodal perspective attends to the entire range of meaning-making forms, referred to as modes, as well as to the interrelations of these modes. Its focus on different modes allows the investigation of communication and interpretation to go beyond written and spoken language as the only resource of meaning making, and take into consideration other modes such as gesture, eye contact, images, sounds, color, layout, camera angle, and so forth (Jewitt, 2016).

Jewitt (2009) explained that there are three primary approaches to conducting multimodal analysis. The first is the social semiotic approach, which emphasizes the role of choices in rendering meanings in specific socio-cultural and historical context. The second one is multimodal discourse analysis (MDA), and this approach focuses on meaning interpretations.

57

The subtle difference between the social semiotic approach and MDA is that the former attaches more importance to the "sign-maker" than the latter. The last approach comprises multimodal interactional analysis which "addresses a dimension of the social semiotic that conventional multimodal analysis does not seem to commonly address and focuses on how multimodal texts are interfaced with and mediated by people" (Jewitt, 2009, p. 33). In this project, the social semiotic approach is adopted to address the choices of semiotic meaning-making resources made by content creators of English language teaching videos on *Douyin*.

A social semiotic approach to multimodality is also referred to as a "social semiotic theory of multimodality", "Multimodal Social Semiotics theory", or "Social Semiotics and Multimodality (Kress, 2010). What lies at the core of this approach is the notion of *sign*, which is a fusion of the *signified* (the meaning) and the *signifier* (the physical form). Signs exist in all modes and are defined as systems formed by the combination of socially and culturally shaped meaning-making resources (Jones, 2013). These meaning-making resources are also referred to as semiotic resources. Social semiotic theory rests on the assumption that signs are always newly created rather than pre-established in social interaction (Kress, 2010). The productions of signs are motivated by the *interest*, defined as "the articulation and realization of an individual's relation to an object or event" in a given social context (Kress, 1993, p.174). Therefore, the relation between a form and a meaning is motivated and never arbitrary.

The social semiotic approach considers communication to be semiotic work and multimodal, mediated by all kinds of modes (Kress, 2010). A linguistic account of speech or communication is limited in the way that it does not give equal attention to the meaning-making potentials of the modes other than 'spoken or written language'. Even the most sophisticated speech transcription that includes intonation, pauses, non-verbal elements, and so forth can miss

many other elements that convey meaning. However, a social-semiotic multimodal account of meaning attends to all signs in all modes to explore what meaning is conveyed, treating signs as a multimodal ensemble that generate meaning as an entire unit. From a multimodal social-semiotic perspective, all modes are equal in their capacity of meaning representation and communication, although each mode has its own way to realize meanings. Therefore, language is no longer viewed as the central means fully capable of communicating all meanings. Rather, it is simply one mode among others that have meaning-making potentials. A social semiotic approach seeks to explore three main aspects: the *interest* of the sign-maker, the environment in which meaning is generated, and the meaning and the semiotic/cultural resources available to the sign-makers (Kress, 2010). Compared to a linguistic, pragmatic, or a sociolinguistic approach, a social semiotic theory of multimodality can inform us:

> about interest and agency, about meaning(-making); about processes of sign-making in social environments; about the resources for making meaning their respective potentials as signifiers in the making of signs-as-metaphors; about the meaning potentials of cultural/semiotic forms (Kress, 2010, p. 59).

As numerous modes, including image, speech, writing, sound, layout, gaze, color and so forth, have been widely deployed in the contemporary digital space, the selection and configurations of various multimodal resources on social media can have unique features. For the creators of the digital content, what they need to take into account during the design of the content is not just the choices of resources but also the social relations they aim to establish. Some common modes in digital contexts include speech, writing, image, layout, and gesture, and these modes frequently interact and operate in ensembles to construct meanings and to communicate complex ideas that are socially informed.

In this dissertation project, the social semiotic approach to multimodality (Kress, 2010; Kress & van Leeuwen, 2001, 2006) is used as an analytical framework for investigating the multimodal constructions of *Douyin* videos motivated by the content producer's *interest*, or in other words, their purpose of accruing followers. The findings of this research can provide insights on the multimodal meaning-making practices as well as the discursive and communicative practices situated in the Chinese online context.

Research Questions

In order to analyze multimodal meaning-making, microcelebrities' self-branding practices, and the design of informal English instructional videos in my dataset, I ask the following major research questions (RQs) and sub RQs in this study:

1. How do creators of informal language instruction videos on *Douyin* project themselves and advertise their products/services on their profile pages?

2. What kinds of English language topics or other topics are addressed in the videos and in what ways are these topics addressed?

    a.   What English language skills are being taught?

    b.   What content topics are discussed?

    c.   What are the teaching styles and methods used?

3. What multimodal resources are employed by content creators in their videos?

    a.   How do they use multimodality to construct English expert identities?

    b.   How do they use multimodal resources to communicate other personae that they present on the platform?

    c.   How do they exploit multimodal resources to teach English?

    d.   What multimodal strategies do they use to engage in self-branding and advertising?

<center>**Data Sampling and Management**</center>

Selection of *Douyin* accounts

   With the aim of examining how *Douyin* users employ multimodal features in order to self-promote through the teaching of English, I sampled 300 accounts that have a focus on English language teaching and/or learning. As an active *Douyin* user, I am fully aware that the content *Douyin* "pushes" to my personal feed is the result of the *Douyin* recommendation algorithms, which are built upon a combination of factors including information from videos (e.g., hashtags, sounds, topics) that I previously spent time viewing, my past interactions with content on the platform (e.g., videos and sounds I liked or shared, accounts I followed, comments I made, content I create, keywords I searched), and popular content on *Douyin* in general. The main purpose of this recommendation algorithm is to offer users personalized content that they are interested in to enhance their satisfaction and dependance on the platform. As such, my past behaviors on *Douyin* are converted to future activities in the form of the videos or live streaming appearing on the "For You" feed of my private account. The personalized feed of my private account includes content of a variety of categories such as travelling, pets, musical instruments, history, science, language and so forth, which I enjoy viewing. However, to sample accounts with an English language teaching focus for the present study, I created a new account instead of using my private account because I needed to train the *Douyin* algorithm to recommend content that restricted specifically to the research purpose of the current project.

   To start the sampling process, I set up a new *Douyin* account on November 8, 2019 using a different phone number, so that the videos showing up in the feed were not personalized as a result of my previous activities on this platform. In what follows, I will present a day-by-day description of how I sampled the 300 accounts relevant to English language teaching. This

<center>61</center>

description of how I trained the algorithm to find videos and accounts is relevant to the current study. To simulate the experience of a user who was new to *Douyin* and who was interested in English language teaching and learning videos, I spent one hour each day during the week of November 8th 2019 to November 14th watching the *Douyin* feed on the "dummy" account I created. When I identified any account that had a focus on English language teaching, I added that account to my "following" list. I followed three steps to identify targeted accounts. First, while watching a video, I relied on my English language teaching expertise to make a judgement about whether the video was relevant to English language teaching/learning. Some evidence I relied on to identify English instructional content included whether the vlogger indicated at the beginning of the video that she will teach something about English (e.g., a vlogger starts the video saying "Here are some food related words that are easily mispronounced"); whether the speaker(s) used English in the video with some instructional meta-language; whether the monologue or dialogue was captioned in English with highlights (e.g., larger fonts, different color) on specific words; whether the vlogger explicitly discussed English learning methods, and so forth. If I identified the video as relevant to English language teaching/learning, then my next step was to click on the account owner's profile image to view their profile page. If they self-identified as an English language teacher, highlighted their English language proficiency, and/or advertised their online English classes in their bio or cover page, my next step was to quickly check the first three videos listed under their profile (excluding the video I had viewed previously in "For You" feed) and see if at least of one of these videos focused on English language teaching/learning. If so, the account was included in my sample. If not, I checked the next three videos (again, excluding the video I had viewed in "For You feed) to see if at least two out of three focused on English language teaching. If yes, I included it in my sample. If no, I

skipped the account. However, there were instances in which the account owner did not associate themselves with English language teaching/learning on their profile page. In this case, I observed the top six videos (excluding the video I had viewed in *Douyin* feed) under their profile and used the above-described criteria to decide whether the videos of that account had a consistent focus on English teaching/learning. If at least two out of the six viewed videos are clearly English language instruction, then the account was included into the sample. Once an account met the sampling criteria, I immediately followed that account.

Following these sampling criteria and procedures, I identified 30 accounts that had a focus on English language teaching on November 8th, 2019. During the first half hour, I only came across two relevant accounts. However, as soon as I began following these two accounts, many more videos about English language appeared in my personal feed. During the second half hour on this day of sampling, I collected 28 accounts that had a focus on English language learning/teaching.

The next day during the same one-hour period, I identified another 85 accounts that had a focus on English language learning/teaching. It took a bit longer to identify if an account should be sampled than on the first day because the videos appearing in my feed seemed at first glance to involve some English language learning/teaching, but after inspecting other content on those accounts, many were actually more about traveling, English song sharing, casual English conversations between couples that were recorded for a life-sharing purpose, discussing various topics in English, and sharing inspirational or emotional English quotes. In other words, several of these were not primarily dedicated to language teaching/learning. However, the overall process of identifying and following the accounts that met the sampling criteria was relatively straightforward.

On the third day of data collection, I identified another 71 accounts that had a focus on English language learning/teaching. On the fourth and fifth days, I identified 55 and 40 new accounts, respectively. In addition to English language learning/teaching videos, there was also an increase in the number of videos showing up in my personal feed that were about teaching/learning of languages other than English, vlogs of studying abroad, travelling to different countries, and sharing various of aspects of everyday life in Western countries such as the tipping culture in the U.S. and the laid-back life Norwegians live. On day 6 I identified 11 new accounts in an hour, and on my last day of sampling, I identified 8 new accounts. During the week I sampled accounts, I observed an increasing number of videos irrelevant to language learning appearing in my feed. The topics of these videos varied (e.g., pets, restaurant reviews, famous quotes, songs, travel, lip sync, stereotypes of a particular group of people). The number of relevant accounts that I was able to sample within a one-hour period started to decrease on day 3, as can be seen in the fewer accounts I sampled during the last two days. However, even with employing this method of training the algorithm to deliver a specific type of content, it is not possible to collect all accounts that meet the sampling criteria (i.e., all accounts which center on English language teaching and/or learning) on *Douyin* not only because there are an immense number of accounts, but also because new accounts dedicated to English learning/teaching likely are added on the platform every day. Therefore, what I collected represents the popular accounts that were actively producing informal English instruction videos during the one-week data collection period.

To summarize this sampling experience, altogether, 300 *Douyin* accounts were collected in one week. On the first day of data collection, the algorithm was learning my preferences. The heaviest days of data collection were days in the middle, when *Douyin* was sending me most

accounts that were relevant. Then towards the end of the week, relevant accounts become infrequent and *Douyin* started to push less relevant content to me in the hopes that I would also like these types of content. Taking this as an indication of near saturation of data, I consider the 300 accounts I collected as sufficient for conducting the present study.

Once the 300 *Douyin* accounts were collected, I captured a screenshot of the profile page of each of the collected accounts and saved all screenshots in a folder on my personal computer with 2 backup files saved in my cloud storage. I assigned each screenshot of the profile page a number from No. 1 to No. 300 and labeled whether the account was "personal" (account run by individual users) or "non-personal" (account affiliated with an institution or a cooperation). In addition, I labeled personal accounts based on the videos of the account that involve the active presence of the vlogger. Then, I transferred this information to an Excel sheet. Moreover, the number of followers of each account, which can be found by checking the screenshot of the profile pages, was also added into the Excel sheet as a separate category. After reviewing these different types of information, I selected 14 of the total of 300 accounts for a more detailed analysis. Ten of the 14 selected accounts were the top 10 most followed, personal accounts, and involve the influencers' active presence in their videos. Accounts with a large audience base but run by institutions/corporations, or those in which videos do not involve the video maker's frequent presence were excluded from this subset. The other 4 accounts were chosen based on distinct features in their videos. These four accounts are:

(1) @嘉德仓自学 English 早 7:50 播 (Jiadecang Self-learning English, Stream Starts at 7:50 a.m.)

(2) @一森教英语(Yiseng[6] Teaching English)

---

[6] *Yiseng* is the Chinese transliteration of *Ethan*.

(3) @英国小克里(British Little Keli[7])

(4) @ 🇺🇸 Allie 爱丽 💕 (<flag of the United States emoji> Allie Aili[8] <two hearts emoji>).

The first of these, the account 嘉德仓自学 English 早 7:50 播 (Jiadecang Self-learning English, Stream Starts at 7:50 a.m.), is run by a Tibetan male, Jiadecang, with most of his videos showing himself practicing spoken English. What is interesting about Jiadecang and his videos is that he actively uses his ethnicity as a resource to attract attention. Moreover, unlike most other microcelebrities who devote tremendous efforts to establish an expert identity in English language teaching, Jiadecang identifies himself as a language learner. Nevertheless, many videos he created have an instructional component, which reflects a continued narrowing divide between experts and novices facilitated by digital technologies (Selwyn, 2015, p. 230). The videos in the second of these accounts, @一森教英语(Yiseng Teaching English), feature a mixed race young boy named Yiseng, who is half Chinese and half British. In the videos, Yiseng draws on his bilingual repertoire to teach English. Yiseng's videos were chosen for further investigation because Yiseng is one of the youngest users in the sample who teaches English on *Douyin*. The owner of the third account @英国小克里 (British Little Keli) is Chris, who was born in the U.S., but grew up in the U.K. He graduated from the University of Cambridge with a bachelors' degree in Chinese and worked in China for 6 years. Chris and his videos attracted my attention because of his multicultural background and transnational family. The fourth of these accounts, @ 🇺🇸 Allie 爱丽 💕 (<flag of the United States emoji> Allie Aili[9] <two hearts

---

[7] *Keli* is the Chinese transliteration of *Chris*.
[8] *Aili* is the Chinese transliteration of *Allie*.
[9] *Aili* is the Chinese transliteration of *Allie*.

emoji>), is created by Allie, an American teenage girl who speaks both English and Mandarin. Allie is the only teenage content creator in my sample, and the way she designs typography (e.g., font size, color) for the texts superimposed on the videos for instructional purposes is quite unique.

Sample specifics

To address my research questions, I focus on 2 data sources: 1) the 14 profile pages of the 14 selected accounts, 2) the 84 videos from the 14 selected accounts, from which a further subset of 5 videos is drawn, for the more in-depth analysis required for RQ 3.

To address RQ1 above microcelebrities self-branding and advertising practices on the profile pages, I use the dataset that consists of the screenshot of the profile pages of the 14 selected accounts. The screenshots in this dataset were taken on February 18, 2021, about one year later following the initial data collection of the 300 accounts. I chose to use updated screenshots of the profile pages to keep the data as up to date as possible and to reflect any possible new trends on the platform.

Next, for each of these 14 accounts, I sampled the first 6 videos that appeared in the video list on the profile page, resulting in a total of 84 videos. These 84 videos will comprise the dataset I will use to analyze RQ2 about English language topics or other topics addressed in the videos and the ways to address them. In many accounts, the first 1-3 videos are "pinned" videos, which means the account creators themselves chose to display the video(s) on the top of their video list appearing on the profile page. Each user can pin up to three of their videos. The videos after the pinned ones appear in reverse chronological order, with the most recent videos appearing at the top. That is to say, at least 3 of the most recent videos at the time of data collection were sampled from each account. All 84 videos were downloaded and saved on

February 18, 2021. In addition, I took a screenshot of the video list that shows the 6 selected videos. This allows me to keep a record of which videos were pinned and which were not. Furthermore, I took a screenshot of each video when I was viewing it on *Douyin* to record information including video descriptions, audio information, product link superimposed on the video that takes viewers to the vlogger's personal store, an embedded e-commerce feature, as well as the time the video was posted. Information of this type, as it is not part of the video, is missing in the downloaded version of these videos. Therefore, I saved a screenshot of each video of selection while viewing them on *Douyin* to record any information that viewers see in addition to the video itself.

For RQ3 that addresses semiotic resources content creators deploy in their videos, I selected one video out of the 6 from each account based on the following criteria. First, if there were pinned videos among the 6, the pinned video with the most likes (relevant to English language teaching) was selected. Second, if none of the pinned videos were about English language teaching, then the video among the rest of the 3 that received the most likes was chosen. Third, if there were no pinned videos, the video that accumulated the most likes and was focused on English language teaching among the 6 was selected. Altogether, 14 videos were sampled for the multimodal analysis, one from each of the 14 accounts. Finally, I selected 4 videos from the 14 videos to analyze their unique multimodal strategies used for self-branding and language teaching purposes.

The first video I chose for a detailed multimodal analysis is a pinned video from @MrYang 杨家成 (Mr. Yang Jiacheng Yang), the most followed account among the 300 accounts collected. The selected video had received more than 2 million likes by the time the data were collected, making it one of the most popular videos of the content creator as well as

one of the most popular videos of all informal English instruction videos on *Douyin*. Therefore, the video was selected for its popularity and the polarity of its content creator. The second video I chose is one of the six videos collected from @大白外教英语 (Foreign Teacher Dabai's English), which is the second most followed account in my dataset. Therefore, one reason for selecting his video is the popularity of the content creator. Furthermore, I find the split-screen feature adopted in the video particularly interesting in making a comparison of English vocabulary words with similar spellings but different meanings. The third video I sampled for RQ 3 is from Scott 学口语's (Learning Spoken English with Scott), one of the top 10 most-followed accounts. This video quickly went viral shortly after its initial posting and has become is one of the most popular videos of the content producer with over 1.1. million likes. Shortly after the video was posted and attracted considerable attention, many other content creators of English language teaching soon "copied" Scott's video script and made their own versions of prepositions and transport without referencing Scott. It did not take long for Scott to notice what had happened and he made another *Douyin* video accusing them of plagiarizing his content and clarifying that he is the original content creator. Although this information is not the focus of this study, it nevertheless speaks for the popularity of this video, which is the main reason I chose it for a detailed multimodal analysis. The last video was selected from @嘉德仓自学 English 早 7:50 播 (Jiadecang Self-learning English, Stream Starts at 7:50 a.m.), one of the four additional accounts I sampled for this dissertation project. I chose the video from this particular content creator to examine how he deploys multimodal resources to construct a hybrid identity of an amateur-expert and how he uses his Tibetan ethnicity for self-commodification.

<center>**Procedure and Data Analysis**</center>

In this section I explain my analytical procedures. I treat the 14 sampled accounts as 14 units. The RQs are organized according to the unit of the analysis.

RQ1: How do creators of informal language instruction videos on *Douyin* project themselves and advertise their products/services on the profile page?

The unit of my analysis for RQ1 is the profile page of the 14 selected *Douyin* English language instructional content creators. In order to answer RQ1, I examined the information the 14 content creators provide on their profile page systematically, including the information they include in cover photos, profile pictures, and bios. I organized information on each profile page into a separate table and explored how the information influencers choose to put on their profile pages help them with self-promotion and advertising one by one. After completing the individual case studies of the profile page, I analyzed patterns in identity construction and product advertising across the 14 accounts.

RQ2: What kinds of English language topics or other topics are addressed in the videos and in what ways these topics are addressed?

RQ2 addresses very specific content that is of interest to TESOL professionals and people interested in English language education. In order to understand how the 14 content creators of English instruction videos use resources to teach and to attract followers, it is necessary to know what the vloggers are doing in their videos. Therefore, to answer RQ2 a content analysis is performed to understand what these videos are about. Specifically, I created a table for each microcelebrity to present information including the topics, language skills addressed, as well as teaching styles and methods of each of the 6 videos from the 14 sampled accounts.

<center>70</center>

After completing the analysis of these 14 cases, I performed a cross-case analysis to look for patterns, such as the most and least common topics discussed, similarities and differences in teaching methods and styles, as well as the English skills most and least often addressed in the videos. I also compared the number of pinned and unpinned videos in relation to their topics to explore the current trend in China's online ELT market.

RQ3: What multimodal resources are employed by content creators in their videos?

The unit of the analysis for RQ3 is videos, but the focus of the analysis is the multimodal strategies content creators utilize to accomplish a list of goals including constructing expert or other types of identities, teaching, as well as self-promoting and advertising. To address RQ3, I adopted a multimodal transcription method adapted from both Matwick and Matwick (2017) and Ho (2019). Matwick and Matwick's (2017) multimodal transcription method used in the analysis of two female-hosted TV cooking shows is useful for the present research as it focuses on the interplay of images, speech, written texts, and sound. Ho (2019)'s multimodal transcription method applied to analyze a YouTuber's English video lessons highlights the structure of the video lesson, which is also useful for my purposes in order to understand pedagogical videos as a hybrid genre and activity.

Following the data transcription, I took the social semiotic approach to multimodality (Kress, 2010; Kress & van Leeuwen, 2001, 2006) to examine content creators' selection and design of multimodal resources for meaning making. I also employed an identity framework that combines Karen Tracy's (2002) and Don Zimmerman's (2014) conceptualizations of identity types to examine the public image the content creators create. I used the categories of transportable, situational, and discourse identities from Zimmerman, but I also included the personal identity category from Tracy in my framework because it allows me to examine the

interpersonal relationships between microcelebrities and viewers. When analyzing the expert identities content creators of informal English instruction videos on *Douyin* construct, Ho and Tai's (2020) framework of expert identity of online English language teaching were used as well. This framework encompasses three aspects: multimodal design knowledge, pedagogical knowledge, and linguistic knowledge. This framework was used to explore how expertise in informal online English teaching is built on the orchestration of the three aforementioned aspects of knowledge.

Evaluation of Research

Positionality

My role as the researcher in the study is mainly a non-participant observer of online interactions. While documenting and sampling *Douyin* accounts and videos that fit the parameters of the current research project, I did not engage in common communicative practices on *Douyin* such as giving likes, leaving comments or sending private messages to other Douyiners: I only followed the accounts that I included in my sample. However, it should be pointed out that my analysis and interpretation of the collected data are by no means neutral and objective. In any research, the researcher inevitably takes a certain stance or point of view in the process of analysis (Barton & Lee, 2013). There is no neutral perspective from which the phenomenon under analysis can be analyzed and interpreted because every researcher has her own background knowledge that shapes her point of view. The research findings, therefore, are influenced by the researcher's subjectivity and thus, it is important to acknowledge this subjectivity in my research.

There are three major factors that may influence the way I handled the data and lead to my subjectivity in the present study. First, as someone who was born and raised in China, I speak

Mandarin as my first language and am familiar with various Chinese socio-cultural digital practices of younger generations. My insider knowledge of the Chinese language and other meaning making resources from China is something I heavily rely on during the process of data analysis. Second, I was a former English language teacher and have experience teaching English in both Chinese and U.S. contexts. Specifically, for about 2 years in China, I taught language skills needed for taking standardized English tests including TOEFL and IELTS to learners of different age groups. I also had 3 years of experience teaching ESL learners at U.S. universities. My familiarity with English language learners as well as the English learning context in both China and the U.S. offers me multiple perspectives to examine the language teaching content on *Douyin.* In addition, I obtained my master's degree in Teaching English as a Second Language at a U.S. university and am currently pursuing a Ph.D. degree in Applied Linguistics. Therefore, my academic background will also influence the way I analyze multimodal texts, as well as evaluate and interpret content relevant to language pedagogy and acquisition. Furthermore, as a second language speaker of English, I consider myself a lifelong learner. As such, I may also bring a learner's perspective to the data analysis of the present study. Third, my familiarity with online interactions and technology on Chinese social media allow me to describe the user interfaces, popular phenomena, as well as digital affordances of different online platforms based on my personal experience as a user. I have been an active user of *Douyin* for more than 3 years. My main activity on this platform has been browsing video content created by others, checking profile pages of some Douyiners and reading users' comments. However, I also have some experience creating, editing, and sharing videos on this platform. My experience related to personal content creation and promotion on *Douyin* was partially motivated by my present research—I created three short videos initially inspired by my everyday life, but also with the

thinking of familiarizing myself with the process of making instructional videos in mind. One of the videos I created was through the *Douyin* built-in video editor. The other two videos were initially filmed using my iPhone camera outside of the *Douyin* app and my digital camera, respectively. The editing tool I used for the video I shot with my smartphone camera was *Jianying*, the official *Douyin* video editing app owned by *Bytedance*. For the video I filmed with a digital camera, I used Adobe Premiere Pro to complete some basic editing. In this project, my prior video filming and editing experience is useful for describing and analyzing content creator's use of multimodal techniques. Nevertheless, it should be noted that my experience and editing skills as a content creator on *Douyin* is very limited, and I am not a professional videographer. In addition, during the process of creating and sharing videos, I familiarized myself with *Douyin*'s marketing tool Dou+, a paid service that helps users to promote their videos and accounts. This also helps me understand whether and how the accounts I examine use this paid service to increase views and attract followers.

Ethical considerations

The dataset of the current research project consists of 14 profile pages and 84 videos from 14 selected *Douyin* personal accounts that focus on English language teaching and which involve the vlogger's active presence. After consulting *Internet Research: Ethnical Guidelines 3.0* produced by the Associated of Internet Researchers (2019), I chose not to anonymize the usernames of these accounts nor de-identify the vloggers for the following reasons. First, access to *Douyin* is publicly available to anyone. Although many users of *Douyin* have their own account, viewers do not need to set up an account to access the content on this platform. Users' *Douyin* account is also public by default. However, they can change the account to private in their settings. Private content is not visible to users to whom the account owner does not give

permission to view their content. Users can also change the private setting of specific videos they create, making the content public, visible to friends only, or completely private. Moreover, users can set the information such as the videos they liked, the accounts they follow, and the list of followers to either public or private. Thus, as viewers, the information and content they can access from all accounts that have not added them as friends is set to be publicly available. In my dataset, none of the vloggers of the 14 selected accounts have personal connections with me or have me on their *Douyin* friends list. That is to say, the data I sampled from these accounts are completely public. I did not collect any data directly from individuals, such as interviews, direct message, or email exchange. Therefore, I was working with multimodal texts, rather than human subjects in this research (Markham & Buchanan, 2012).

Second, the content creators of the 14 selected accounts can be considered to be public figures. These 14 accounts include the top 10 most followed among the 300 accounts I initially collected, having a following of at least a million people. Although the remaining 4 accounts do not have a fanbase as large as the top 10, the vloggers of these 4 accounts still managed to amass a considerable number of followers, with a following of at least 380,000 people. As Marszałek (2020) maintains, a Douyiner has achieved the microcelebrity status when they accumulated 100,000 followers. Following this criterion, the content creators of all the 14 accounts are *Douyin* microcelebrities and can be considered public figures. Furthermore, the vloggers of the sampled accounts are very intentional about being as public, as they strive to attract more followers in the attention economy, with the ultimate goal of monetizing their products and services. If someone's goal is to get as many followers as possible and to be visible/public, very intentionally promoting themselves, I assume that they would not have a problem with my research identifying them in scholarly publications. In other words, what I am doing with their

data does not contradict to their expectations for their information distributions (Nissenbaum, 2010). Moreover, referring to the guideline in *Internet Research: Ethnical Guidelines 3.0* that research should do more than "do no harm" but also benefit the people involved, I consider my study has potential to bring more attention and followers to the microcelebrities of the selected accounts, which can eventually lead to their monetary gains.

CHAPTER FOUR: RQ 1 FINDINGS

In the following three chapters, I present my analyses and findings for the three research questions. In this chapter, the analyses are presented with the goal of addressing my first research question:

RQ 1: How do creators of informal language instruction videos on *Douyin* project themselves and advertise their products/services on their profile pages?

The data analysis for RQ1 focuses on the profile page of the 14 selected *Douyin* English language instructional content creators. Specifically, I consider both textual and visual information included in their cover photos, profile pictures, and bios in their profile pages. As mentioned in Chapter 3, I first organized the information on each profile page as a separate table. Next, I followed the social semiotics approach to multimodality (Kress, 2010; Kress & van Leeuwen, 2001, 2006) to analyze content creators' multimodal meaning-making practices on their *Douyin* profile page. To explore how those *Douyin* microcelebrities create public images, I employed an identity framework that combines Karen Tracy's (2002) and Don Zimmerman's conceptualizations of identity categories, including transportable identity, situational identity, discourse identity, and personal identity. I operationally treat expert identity as a subcategory under personal identity because this identity type is relative, attending to relationships people establish with others. Additionally, I utilized Buchotlz and Hall's (2010) five principles (i.e., the emergence principle, the positionality principle, the indexicality principle, the relationality principle, the partialness principle) to investigate the characteristics of multimodal identity construction of *Douyin* content creators of informal English language instructional videos. The

remainder of this section is devoted to present the information that is relevant to self-branding

and advertising and the analysis of multimodal identity construction on the profile page of the 14

selected *Douyin* accounts one by one. The top 10 most followed accounts are presented first,

followed by the 4 additionally selected accounts that demonstrate unique features of self-

branding strategies and English language teaching. A cross-case analysis follows immediately

after that.

1) @ MrYang 杨家成 (Mr. Yang Jiacheng Yang)

The account @MrYang 杨家成 (Mr. Yang Jiacheng Yang) is the most followed account

among the total 300 accounts collected, with over 11 million followers. The owner of the

account is Jiacheng Yang, one of the most popular creators of informal English language

instruction videos on the platform. Figure 3 is a screenshot of Yang's *Douyin* profile page, on

which *Douyin* users can customize their cover photos and profile pictures, create a bio[10], display

their Influencer Personal Verification[11], and include the link to their Starmaps. Starmap is

Douyin's influencer platform that users with at least 100,000 followers can sign up for. Once

users set up a Starmap account, their account data, including content information, demographics

of followers, interaction rate, and so forth, will be shared with brands that might be interested in

cooperating with them (Marszałek, 2020). The information that Yang includes on his profile

page to self-promote and advertise his products/services is listed in Table 1.

---

[10] As of Aug 31, 2021, *Douyin* allows a self-description of 242 words maximum.
[11] A *Douyiner* has to accumulate at 10,000 followers in order to be eligible for Influencer Personal Verification.
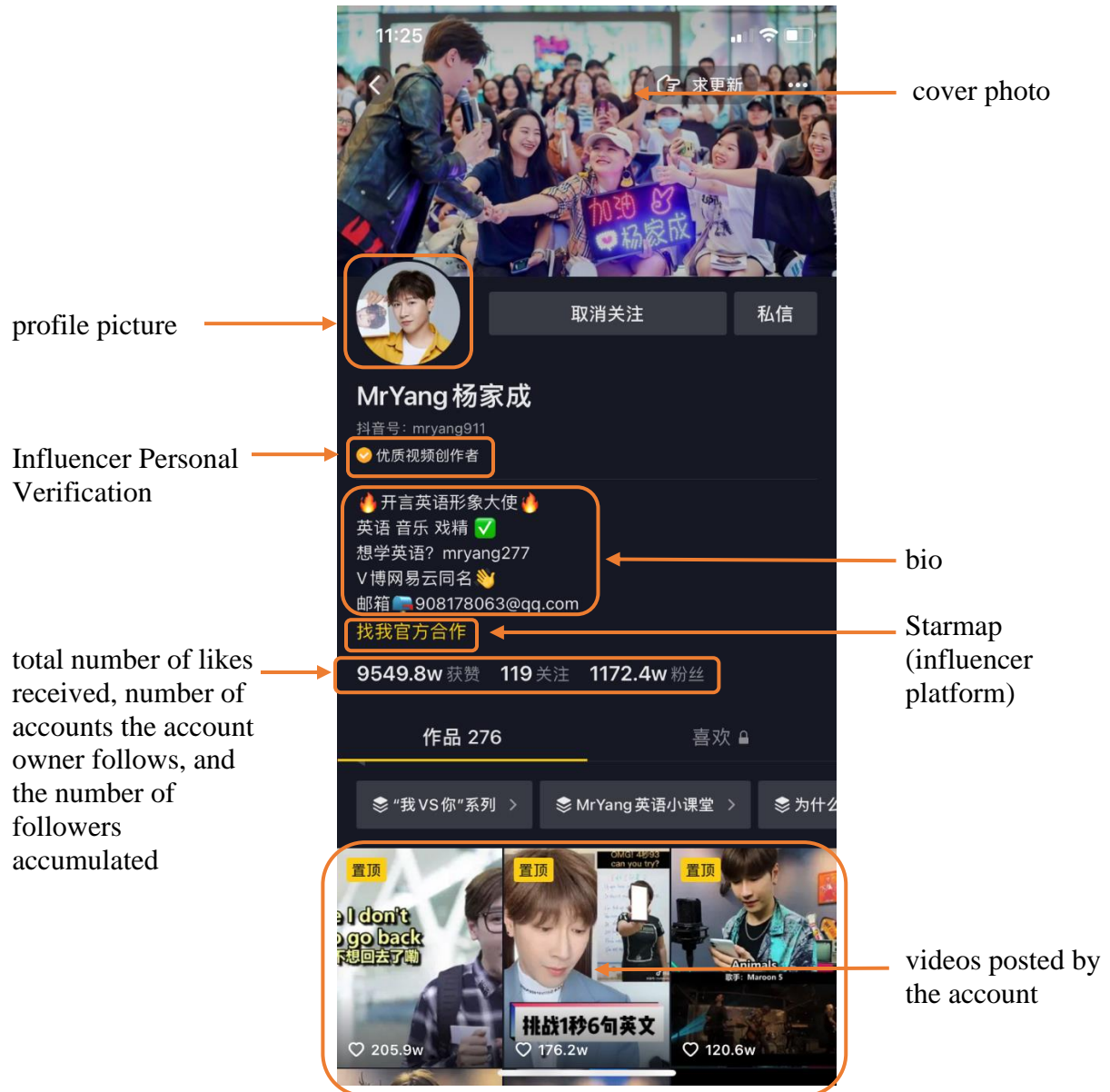
Figure 3. A Screenshot of @ MrYang 杨家成's (Mr. Yang Jiacheng Yang) Douyin profile page

Table 1. Information Yang includes on his *Douyin* profile page

Information category	Content
Cover photo	Yang is shaking hands with one of his fans; a young woman holding a sign saying "Way to go, Jiacheng Yang". The fans in the photo (more than 50 people included) appear to be mostly young Chinese females.
Profile picture	A head shot of Yang holding a book. The front cover features a photo of Yang's face.
Bio	" 🔥 (flame emoji) The brand ambassador of *Open Language*[12] 🔥 (flame emoji); English, music, drama king ☑️ (green check mark button emoji); Want to learn English? *Mryang277*[13]; Same ID name on *Vbo*[14] and *NetEase Cloud Music*[15] 👋 (waving hand emoji); Email 📫 (mailbox emoji) 908178063@qq.com."
Influencer Personal Verification	Yang's account verified as *creator of high-quality videos*
Starmap (influencer platform)	A link to the Starmap platform

In the context of *Douyin*, Yang's overall role, or in other words, his situated identity, is a

Douyin user. However, the information Yang provides on his profile page allows him to

construct various kinds of discourse and personal identities for self-branding and product

advertising purposes. The discourse identities that emerge from Yang's profile page information

include an online content creator, a microcelebrity, and an advertiser. Undoubtedly Yang's role

---

[12] *Open Language* is an English language learning app in China.
[13] A *WeChat* ID or username that people can search to add the account as a contact. *Wechat* is a Chinese messaging app.
[14] Vbo is a combination of the English letter "V" and the Chinese expression "bo". Vbo is used to refer to Weibo, a Chinese micro-blogging site, because "V" is phonetically similar to "Wei".
[15] *NetEase Cloud Music* is a Chinese music streaming platform.

as a content creator is indicated through the videos he posted, but this role is also further

highlighted through the Influencer Personal Verification badge he applied for and obtained from

Douyin, which recognizes him as *creator of high-quality videos*. In addition, to further

emphasize his microcelebrity status, Yang employs various multimodal resources including a

cover photo featuring him being surrounded by many fans, the Influencer Personal Verification

badge, the link to his Starmap platform he set up, as well as the emojis and text used to describe

his cooperation with the brand *Open Language* as its brand ambassador. The two flame emojis

(🔥) Yang puts at the beginning and the end of the message " 🔥 The brand ambassador of *Open*

Language 🔥 " in his bio are used to convey a sense of popularity as the Chinese expression of

火 (huo, meaning: fire or flame) can also be used to refer to popularity.

　　Another feature on Yang's profile page that helps construct Yang as a highly

accomplished microcelebrity is his profile picture. Small in size as it might be, the profile picture

is a head portrait of Yang holding a book of which the front cover is a photo of Yang's face. As

an observer of Yang's account and videos for over two years, I recognize that the book he is

holding in the photo is his autobiography which focuses on his English learning trajectory.

However, to those *Douyin* viewers who do not know what the book is about, the fact that Yang's

photo is on the front cover of the book still serves to indicate Yang's close relationship with this

book and his accomplishment as a microcelebrity involved in publishing. By actively promoting

his microcelebrity status through his cover photo, profile picture, Influencer Personal

Verification, Starmap platform, and bio, Yang engages in the activity of self-advertising in which

he is the advertiser. This discourse identity is also implied through other information Yang

includes in his bio including labelling himself as a humorous ("drama king") and versatile

individual who loves the English language and music. He also provides his other social media

accounts and business contact information to gain further attention from the public and from brands. In addition to Yang's discourse identities as an *Douyin* influencer and advertiser, several personal identities are established through both the linguistic and semiotic modes on the profile page. Specifically, in the cover photo, Yang's fans appear to be mostly young Chinese females. This visual information serves to indicate Yang's attractiveness to women. He also presents himself as an approachable individual through the cover photo that features him shaking hands with his followers. His profile image is a close-up shot of him smiling and having direct eye contact with the viewer, which conveys a sense of closeness. Moreover, in his bio, Yang indicates that he is versatile and talented by highlighting his strengths in English, music, and creating humor. Another important personal identity Yang establishes is a language expert identity. Although the manifestation of this type of identity is much more salient in his videos as will be demonstrated in the analysis of RQ3, the role of an expert is still emphasized on his profile page when Yang states at the very beginning of his bio that he is the brand ambassador of *Open Language*. As *Open Language* is an English language learning app in China, his collaboration with this online education platform not only reflects his popularity and success as a microcelebrity, but also endorses his expertise in English.

2) @大白外教英语 (Foreign Teacher Dabai's English)

The next account is @大白外教英语 (Foreign Teacher Dabai's English) with more than 4 million followers. It is the second most followed account among the 300 accounts collected. As the account name indicates, @大白外教英语 (Foreign Teacher Dabai's English) is run by a non-Chinese person Dabai, who self-identifies as a "foreign English teacher" to Chinese learners. The literal meaning of the Chinese name Dabai (大白) is "big white", which indexes the race of the content creator. Additionally, Dabai is also the translated name of Baymax, an inflatable

computerized robot in the Disney's popular movie *Big Hero 6*. The box office of Big Hero 6

achieved a tremendous success in China and Baymax has quickly become one of the favorite

Disney characters among young Chinese, especially Chinese girls. Many of them rated Baymax

as "the best boyfriend" and called it a 暖男 (*nuan nan*, meaning "warm-hearted boy"). Shortly

after the release of the movie in China, the expression, *Dabai* quickly went viral and was

frequently used by young women to praise a man who is considered a "*nuan nan*". Therefore, by

giving himself the Chinese name Dabai, the account owner of @大白外教英语 (Foreign teacher

Dabai's English) relates himself to Baymax, a superhero who is caring, gentle, protective, and

responsible. Moreover, Dabai is phonetically similar to Darby[16], the English name of the account

owner. Table 2 presents a translation of the information Dabai conveys through his *Douyin*

profile page for identity construction and product/service promotion.

The major identities Dabai constructs on the profile page (see Figure 4) to attract

followers include an authentic English teacher who happens to be Caucasian and a generous man

who is dedicated to helping others. Being Caucasian is Dabai's transportable identity, which is

stable and visible regardless of the changing contexts. The information about Dabai's race

conveyed through visual resources corresponds to Dabai's *Douyin* account name @大白外教英

语 (Foreign Teacher Dabai's English), which places an emphasis on Dabai being a foreigner.

The particular ethos Daibai establishes as an English teacher is also indicated through his bio in

which he claims his goal is to make English lessons taught by a foreign teacher affordable to

everyone. In addition to depicting himself as a generous person ready to help people in need,

what is also indicated through this message in his bio is that foreign English teachers, or native

---

[16] Based on my long-term observation of this account, the content creator indicated in several of his *Douyin* videos that his name is Darby, although this information is not explicitly stated on his profile page or the 6 videos sampled from his account for the present research.

English-speaking teachers, are relatively rare resources in China and their language courses are often pricey due to the perceived authenticity and effectiveness. Dabai's authenticity is further reinforced by his gesture, chin stroking, in the cover photo and profile page. The chin stroking gesture in China is commonly interpreted as a sign of evaluation from someone in power, implying that the authoritative figure who is making this gesture is going through the decision-making process.

Along with self-branding, Dabai also actively draws on multimodal resources to advertise his products and services on his profile page. The main services Dabai promotes are his live English classes on *Douyin*. In his cover photo, his stream content ("live English classes") and schedule ("every evening at 7p.m.") is highlighted though the layout/position of the texts (at the center of the image) as well as the large font sizes and the contrast between the color of orange and white. Similarly, Dabai directs viewers' attention to the information about his live English lesson in his bio by putting his stream schedule in the first line. In addition to Dabai's livestream, in the bio he also informs viewers about how to sign up for his English classes on *WeChat*. For business inquiries and potential collaborations with brands, Dabai indicates that he can be reached through *WeChat* and Starmap. Interestingly, whenever *WeChat* is mentioned in bio, Dabai uses "V ❤️ (red heart emoji)" to refer to it instead of using "WeChat" or "Weixin", the Chinese name of *WeChat*. In the expression of "V ❤️", the English letter "V" is phonetically similar to 微(Wei), and the red hear emoji "❤️" is employed to represent the Chinese word 心 (xin, meaning 'heart') , which is a homophone for 信(Xin). Therefore, "V" and "❤️" are combined to represent 微信(Weixin). The main reason for this communicative practice is that

Figure 4. A screenshot of @大白外教英语's (Foreign Teacher Dabai's English) *Douyin* profile page

Table 2. Information Dabai includes on his *Douyin* profile page

Information category	Content
Cover photo	The cover photo features a professional portrait of Dabai stroking his chin. To his right is a virtual whiteboard with Dabai's live stream/class schedule on it: "live English classes" "every evening at 7p.m."
Profile picture	A professional portrait of Dabai with a chin stroking gesture.
Bio	"Stream schedule: every day from 7p.m.; To sign up for my English classes, add 'mingshi 27149' [17]as a contact on V ❤ [18]; Make English affordable for everyone; Make English lessons taught by foreign teachers affordable for everyone; Business: 15510013016 (V ❤ [19])"
Starmap (influencer platform)	A link to the Starmap platform

Douyin blocks several other Chinese social media platform names, including *WeChat*, due to the growing competition among these sites. It is worth noting that this type of heteroglossic practice (Zhang, 2017), such as using English letters, emojis, Chinese pinyin initials, is frequently adopted to evade *Douyin* censorship by many content creators who intend to advertise the products and services they provide on other media sites. As non-Chinese person, Dabai's awareness of *Douyin* censorship and his strategic employment of local resources to avoid getting blocked demonstrate his familiarity with the Chinese socio-cultural context and online practices, which helps him establish group affiliation with the viewer.

---

[17] A *WeChat* ID or username that people can search to add the account as a contact.

[18] "V ❤" is used to refer to 微信(Wei Xin), the Chinese name for *WeChat*, because "v" is phonetically similar to 微 (Wei), and the heart emoji "❤" represents the Chinese word 心(xin, meaning 'heart'), which is used as a homophone for 信(Xin).

[19] See footnote 6

3) @晓莉英语-11:30 直播 (Xiaoli's English-Live Stream Starts at 11:30)

The account @晓莉英语-11:30 直播 (Xiaoli's English-Live Stream starts at 11:30) has accumulated nearly 4 million followers. As indicated in the account name, Xiaoli is the account owner and content creator. What is interesting about her account name is that it also includes her live streaming schedule. Informing viewers her stream schedule helps Xiaoli make a constant, regular presence to the audience. More importantly, I have observed that many microcelebrities of English language instruction, including Xiaoli provide live streams of English classes on *Douyin* using the live streaming feature embedded in the application. During the stream, they deliver demo language lessons, promote language learning products (e.g., books, video lessons, online tutoring services), encourage synchronous interactions with viewers, and receive fan donations.

In addition to her account name, Xiaoli's live broadcasting schedule is further specified in the first line of her bio: "Stream schedule: Monday-Friday, and Sunday, from 11:30". To provide viewers with different means to interact with her and sell her language courses, Xiaoli shares her WeChat ID in the bio "To learn English, add 'Lily100131' as a contact on v.x.. Again, v.x. is used to refer to "Weixin" (the Chinese name for *WeChat*) to evade *Douyin* censorship. The English letter "v" is used to represent "wei" due to their similar sounding, and "x" is the Chinese Pinyin initial of "xin". Additionally, she also provides links to her business contact information including her business address and phone number.

cover photo

profile picture

Influencer Personal Verification

bio

A link to the Starmap platform

A link to her business address, and a link to her business phone number.

total number of likes received, number of accounts the account owner follows, and the number of followers accumulated

Videos posted by the account

Figure 5. A screenshot of @晓莉英语-11:30 直播's (Xiaoli's English-Live Stream starts at 11:30) *Douyin* profile page

Table 3. Information Xiaoli includes on her *Douyin* profile page

Information category	Content
Cover photo	The profile picture is placed on the top of the cover photo. Under the cover photo a list of English classes and services Xiaoli offers, including "New Concept English lessons, business English lessons, spoken English lessons, pronunciation lessons, free assessments, free trial lessons, individual consultations".
Profile picture	A photo of Xiaoli smiling and holding her face with two hands. A filter is applied to the photo, making it look like a color pencil drawing.
Bio	"Stream schedule: Monday-Friday, and Sunday, from 11:30; We are located in Shenzhen and looking for qualified teachers to join our team; To learn English, add 'Lily100131' as a contact on v.x.[20]; Master's degree in Translation from the Chinese University of Hong Kong; 8 years of experience as a teacher I am proficient in English and serious about teaching."
Influencer Personal Verification	Xiaoli's account verified as "Educational Self-Media"
Starmap (influencer platform)	A link to the Starmap platform
Business Contact Information	A link to her business address and a link to her business phone number.

To construct her language teacher and expert identity, Xiaoli lists her professional credentials in the English language and English language teaching including her Master's degree in Translation from a high-ranking university in HongnKong ("Graduated from the Chinese University of Hong Kong with a Master's degree in Translation") and her extensive teaching experience ("8 years of experience as a teacher") in the bio. She also further promotes her expertise in English and teaching by asserting that she is "proficient in English and serious about

---

[20] "v.x." is used to refer to 微信(Wei Xin), the Chinese name for *WeChat*, because "v" is phonetically similar to 微 (Wei), and "x" is the Chinese Pinyin initial of 信(Xin).

teaching". Furthermore, her cover photo features a list of language courses and services she can provide for the learner, including "New Concept English lessons, business English lessons, spoken English lessons, pronunciation lessons, free assessments, free trial lessons, individual consultations". By specifying the types of English courses she can teach and offering free services, Xiaoli presents an image of a professional and responsible teacher who is ready to help. Additionally, the Influencer Personal Verification badge "Educational Self-Media" Xiaoli obtained from *Douyin* also speaks for her qualification in teaching. The term "self-media" is used on Chinese social media to refer to accounts that are run by individual users and produce original content. Acquiring the status of "educational self-media" not only indicates Xiaoli's competence in creating informal language lessons and learning materials, but also endorses her success as an influencer.

Another identity Xiaoli constructs on her profile page is an entrepreneur who leads a team of language teachers. Specifically, she states in her bio that "we are located in Shenzhen and looking for qualified teachers to join our team". The first-person plural pronoun "we" as well as the word "team" indicate that Xiaoli is running her English language training business and seeks to further expand her business. This information Xiaoli puts in her bio portrays her as a successful businessperson in China's ELT market. Consistent with the image of a high achiever, her profile image, which is a close-up selfie, elevates her persona through the visual effect applied to the photo. Xiaoli's face featured in the selfie loses detail through overexposure, leading to an ethereal brightness which indexes an ideational value (Kress & van Leeuwen, 2006).

4) @英语影视配音员 孙志立 (Zhili Sun the Voice Actor of English)

The account name @英语影视配音员 孙志立 (Zhili Sun the Voice Actor of English) consists of the name of the account owner, Zhili Sun, and his profession as a voice actor of English. As suggested by his account name, one identity he establishes on his profile page is a voice actor of English. For instance, the profile picture features Sun in a studio wearing a headset, which suggests that he is doing voice acting (see Figure 6). Another identity Sun constructs along with his voice actor identity is an English teacher. Both identities are communicated to the viewer through Sun's bio and Influencer Personal Verification badge. In Sun's bio, he self-identifies as a voice actor for English commercials and a teacher. Similarly, in Sun's Influencer Personal Verification badge, he is verified by *Douyin* as a "star English teacher" and a "voice actor of English". The cover photo features a cursive script in English. Although it does not directly contribute to Sun's professional identity, it nevertheless denotes Sun's close relationship with the English language.

Table 4. Information Sun includes on his *Douyin* profile page

Information category	Content
Cover photo	An image of English calligraphy on a piece of paper with two wooden clips and a pen.
Profile picture	A photo of Sun wearing a headset and standing in front of a microphone mounted on a mic stand in a recording studio, indicating that he is performing voice-overs.
Bio	"a voice actor of English commercial advertising, teacher" Personal information including gender (male), age (41), and location (Haidian District, Beijing) is provided.
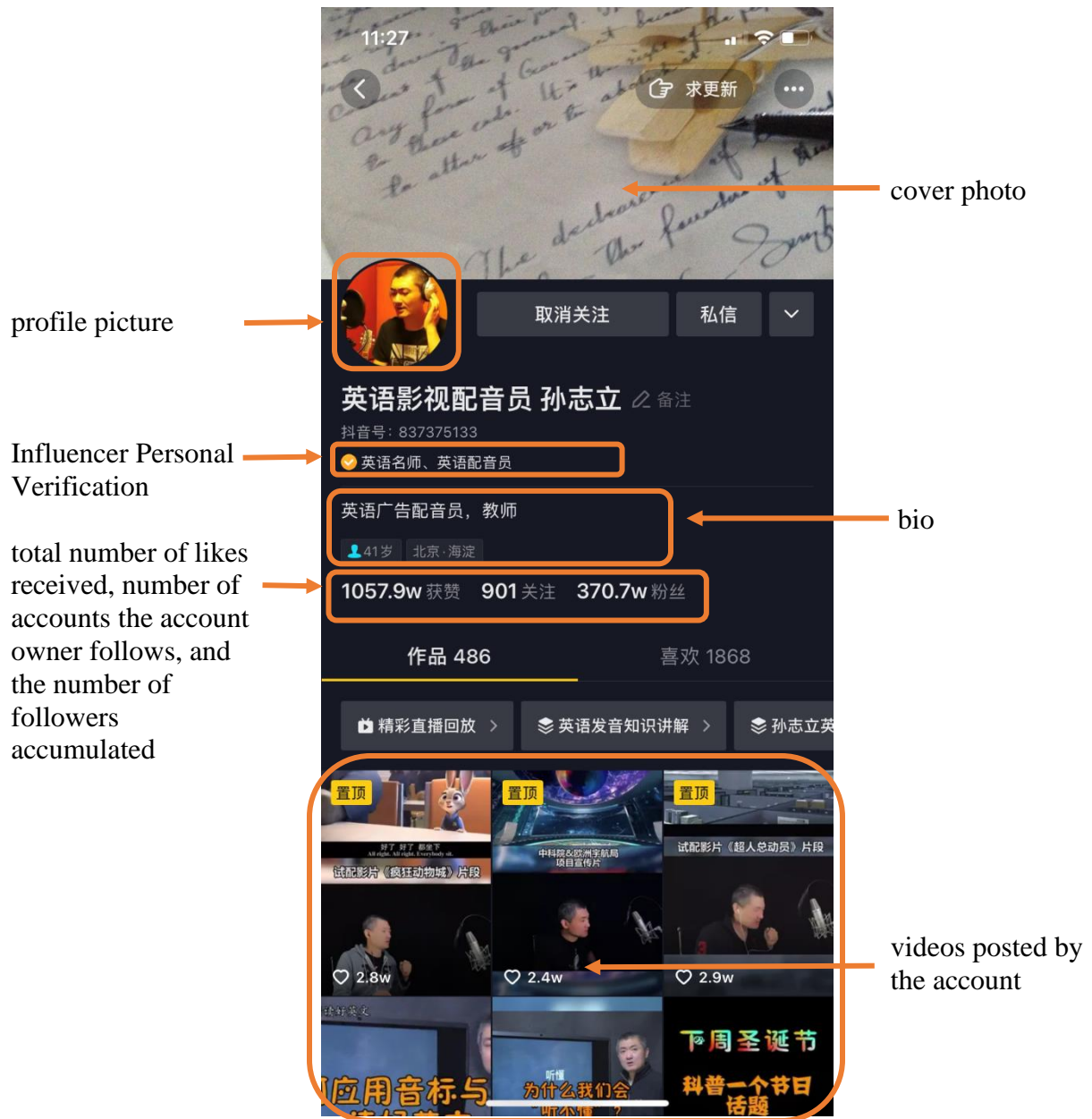
Figure 6. A screenshot of @英语影视配音员 孙志立's (Zhili Sun the Voice Actor of English) *Douyin* profile page.

5) @口语控 (Spoken English Fanatic)

Among the 14 selected accounts, @口语控 (Spoken English Fanatic) is the only account name that does not include the name of the account owner. In the Chinese online context, "控" (literally meaning "control") can be used to describe a person who is obsessed with and devoted to an interest, sharing a similar meaning with the English expression "fanatic". Although the account name "Spoken English Fanatic" invokes an image of an enthusiastic speaker of English, it is less personalized without explicitly addressing who the fanatic is. This impersonal feeling is also conveyed in the cover photo and profile image of this account. The cover photo features a low-angle long shot, in which the account owner is dressed in black and standing on piles of black tires at a dump site. The low-angle shot makes viewers "look up to" him, denoting an interactive relation in which the content creator has power over the viewer (Kress & van Leeuwen, 2006). The color black in the Chinese context can also symbolize power and authority. In addition, the account owner turns toward his side and lacks eye contact with the audience. As such, he demands the viewer to perceive him from a detached point of view. In a similar vein, the content creator in the profile image looks down to his side, which reduces the involvement of the audience. The chin-stroking gesture is associated with authority and evaluation in China, thus creating distance between him and the audience as well. In terms of color coordination, both the cover photo and the profile picture feature the account owner dressed in black. The color of his clothing, together with the piles of black tires in his cover photo, creates a visual rhyme that integrates his profile picture and cover photo into a coherent unit to construct him as an authority. What type of authority he belongs to is indicated by his account name "Spoken English Complex" and the Influencer Personal Verification badge he applied for and successfully

obtained from *Douyin*, in which he labeled himself as a "Knowledge Provider on *Douyin*". Thus, he establishes an image of an educator and an English language expert.

cover photo

profile picture

Influencer Personal
Verification

bio

A link to the
Starmap platform

total number of likes
received, number of
accounts the account
owner follows, and
the number of
followers
accumulated

videos posted by
the account
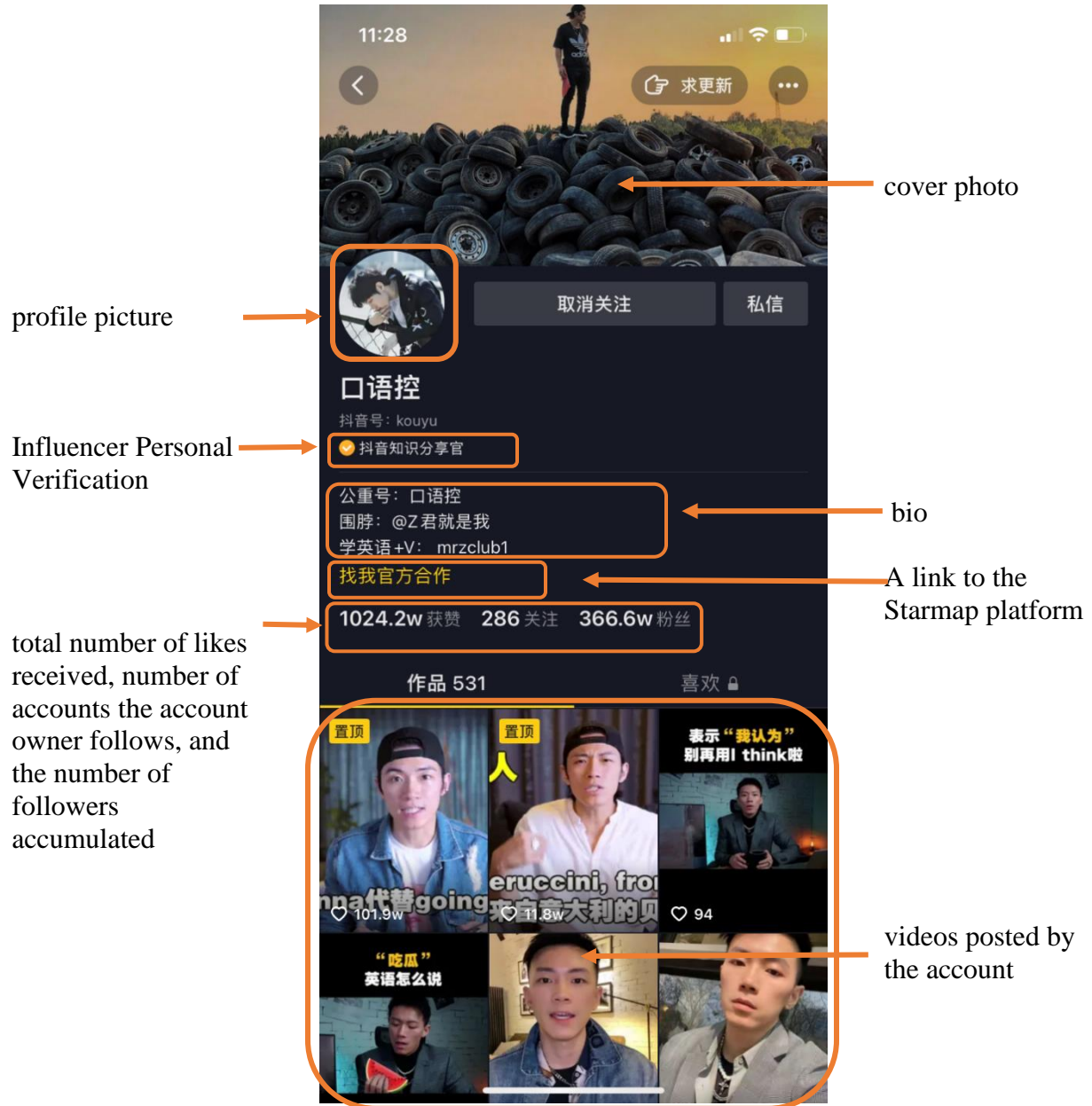
Figure 7. A screenshot of @口语控's (Spoken English Complex) *Douyin* profile page

Table 5. Information @口语控 (Spoken English Complex) includes on his *Douyin* profile page

Information category	Content
Cover photo	A low-angle long shot. The account owner is dressed in black, standing on piles of black tires at a dump site.
Profile picture	The account owner wearing a hip-hop streetwear black jacket, hip-hop fashion jewelry (rings and necklace). He is leaning against an iron fence, looking down to his side with a chin stroking gesture.
Bio	Official account[21]: Spoken English Complex Weibo[22]: @z 君就是我 (I'm Mr. Z) To learn English, +[23] 'mrzclub1' as a contact on V[24];
Influencer Personal Verification	The content creator is verified as "Knowledge Provider on *Douyin*"
Starmap (influencer platform)	A link to the Starmap platform

Apart from an authoritative image the content creator invokes, the content creator seeks to connect to the followers. In the bio, he provides his *WeChat* official account name, *Weibo* account name, as well as his personal *WeCha*t account ID, and he also encourages viewers to learn English with him on *WeChat*. The names of those social media sites, again, are represented through non-standard practices by using homophones (see footnotes 12 and 13), math symbols (see footnote 14), and English transliterations (see footnote 15) to avoid *Douyin* censorship. In addition, as a microcelebrity, he also includes a link to his Starmap platform to welcome cooperation with brands.

---

[21] 公众号 (official account) is a shorted form for "*WeChat* official account". The content creator uses 公重号 (gong zhong hao, literally translated as "official heavy account") to refer to 公众号(gong zhong hao, literally meaning "official public account") because these two expressions are homophonous in Chinese.

[22] In the bio, 围脖(weibo, meaning "neck warmer") is used as a homophone for 微博(Weibo).

[23] The symbol "+" is Chinese is referred to as "加", which is polysemous word that can also mean "add".

[24] "v" is used to refer to 微信(Wei Xin), the Chinese name for *WeChat*, because "v" is phonetically similar to 微 (Wei).

6) @Caroline 教英语 (Caroline Teaching English)

The account @Caroline 教英语 (Caroline Teaching English) is run by Caroline, as suggested by her *Douyin* account name. On her profile page, Caroline mainly focuses on establishing connections with the viewer. In her bio, she promotes her live stream class by advertising its effectiveness and providing a detailed, active live streaming schedule of English teaching ("Everyday livestream to help you improve English fast. I'll see you at 8:30a.m., 12p.m., and 8p.m."), through which her discourse identity as an English language expert and an advertiser are constructed. Similarly, her cover photo encourages interactions with the audience. The cover photo features the text "want to learn English" in Chinese in a large font size to attract the viewer's attention. To the right of the text is a big arrow pointing down at the "following" button under the cover photo. Combining the text and the visual element, a message that encourages following is constructed. The text added on the right side of the arrow, meaning "follow and send message", also serves to enhance the connection between the audience and the content creator. Furthermore, the profile picture features a close-up shot selfie of Caroline smiling and looking directly at the camera. As such, Caroline addresses fans visually through direct eye contact. Together with her smile and the intimacy invoked by a close-up shot, she builds rapport with the audience. In addition, Caroline's race as Caucasian, which is one aspect of Caroline's transportable identity, is evident in her profile page. Different from other identity types that need to be constructed, transportable identity is given and stable regardless of social context. Although Caroline's profile image communicates her racial identity, she is not engaged in identity work to actively establishing this identity type.

cover photo

profile picture

总 number of likes received, number of accounts the account owner follows, and the number of followers accumulated

bio

videos posted by the account

Figure 8. A Screenshot of @Caroline 教英语's (Caroline Teaching English) *Douyin* profile page

Table 6. Information Caroline includes on her *Douyin* profile page

Information category	Content
Cover photo	An image with text "want to learn English" in a large font size, to the right of the text is an big arrow pointing down at the following button under the cover photo. On the right side of the arrow is the text "follow + send message".
Profile picture	A selfie of Caroline standing on a busy street and smiling at the camera.
Bio	Everyday livestream to help you improve English fast. See you at 8:30a.m., 12p.m., and 8p.m.

7) @Scott 学口语 (Learning Spoken English with Scott)

The next *Douyin* microcelebrity is Scott, who runs the account @ Scott 学口语's

(Learning Spoken English with Scott). Like most other influencers, Scott seeks to reach out to as

many viewers as possible through different Chinese social media platforms. To introduce his

Weibo account, Scott uses a screenshot of his Weibo homepage with his Weibo account name

appearing in the center as his *Douyin* cover photo. Moreover, in his *Douyin* bio, he presents his

live streaming schedule of English teaching, his WeChat official account, as well as his personal

WeChat account (see Table 7). To advertise himself to brands that are looking for influencers for

business cooperation, he includes a link to his Starmap platform as well. Scott's *Douyin* profile

picture is a close-up shot of him smiling and looking directly at the camera to address the viewer

in order to convey a sense of affinity. His gesture of pointing to his head with one finger signifies

the action of thinking in the Chinese culture, which portrays him as an inquisitive person.

Figure 9. A Screenshot of @ Scott 学口语's (Learning Spoken English with Scott) *Douyin* profile page

Table 7. Information Scott Includes on his *Douyin* profile page

Information category	Content
Cover photo	An image of the introduction area of Scott's Weibo account homepage, on which Scott's Weibo profile picture (a headshot of Scott) and account name (Scott 学口语) are displayed. It also show Scott's Weibo account has been verified as an account of "Teacher Blogger".
Profile picture	A headshot of Scott in business suit smiling and pointing to his head with one finger.
Bio	English teaching livestream starts at 21:30 Find more content in the official account[25]: Scott 学口语 (Scott learning spoken English) To learn English, add "kouyu5" as a contact on v ❤ [26] Personal information including gender (male), and location (Chengdu, Sichuan Province) is also provided.
Influencer Personal Verification	Scott is verified as "Knowledge Provider on *Douyin*"
Starmap (influencer platform)	A link to the Starmap platform

In addition to depicting himself as a readily available and approachable person, Scott also constructs an English expert identity. His cover photo, which is a screenshot of his Weibo homepage shows that Scott's Weibo account is verified as a "Teacher Blogger". Similarly, the Influencer Personal Verification badge Scott obtained recognizes him as a "Knowledge Provider on *Douyin*", which is the same badge that @口语控 (Spoken English Complex) has. The

[25] A shortened form for "*WeChat* official account". The platform name "WeChat" is not mentioned because *Douyin* blocks it due to the growing competition between *WeChat* and *Douyin*. Scott uses 公粽浩 (gong zong hao) to refer to 公众号 (gong zhong hao, meaning "official account") because these two expressions are phonetically similar in Chinese.

[26] "v ❤" is used to refer to 微信 (Wei Xin), the Chinese name for *WeChat*, because "v" is phonetically similar to 微 (Wei), and the heart emoji "❤" represents the Chinese word 心 (xin, meaning 'heart'), which is used as a homophone for 信 (Xin).

verification from two popular Chinese social media platforms helps validate his English teacher identity.

As will be discussed in Chapter 6, the location information (Chengdu, Sichuan Province) Scott includes in his bio also indexes a humorous identity he constructs in his video.

8) @EMY

The account @EMY is owned by a young Chinese woman who refers to herself as Emy. Different from most other microcelebrities in my data, she does not present a language teacher or expert image on her profile page. Instead, she positions herself as an enthusiastic and persistent learner of English as she states in her bio "like learning English, keep practicing". Based on my long-term observation of Emy's account and her videos, in contrast to the learner identity she claims, she is actually a proficient English speaker and holds a Master's degree in Interpretation from Middlebury College. For viewers who are familiar Emy and her English performance, the learner identity Emy assigns to herself helps establish a positive image of her as a humble individual who is always willing to learn. This also meets the expectation of humility in Chinese culture.

In addition to her learner identity, she also demonstrates a lively, positive, and amiable personality. Her profile picture features her at Disney World, wearing a Mickey Mouse hairbow and holding a popsicle in the shape of the Mickey Mouse head, though which she depicts herself as fun, lively, and passionate about life. The medium shot of the photo and Emy's direct looking at the camera encourages the viewer's involvement in the world Emy is in, resulting in an increased connectedness. In addition, as Emy indicates in the bio that one of her reasons for making *Douyin* videos is to share life, she expresses her passion for life. Her lively and positive
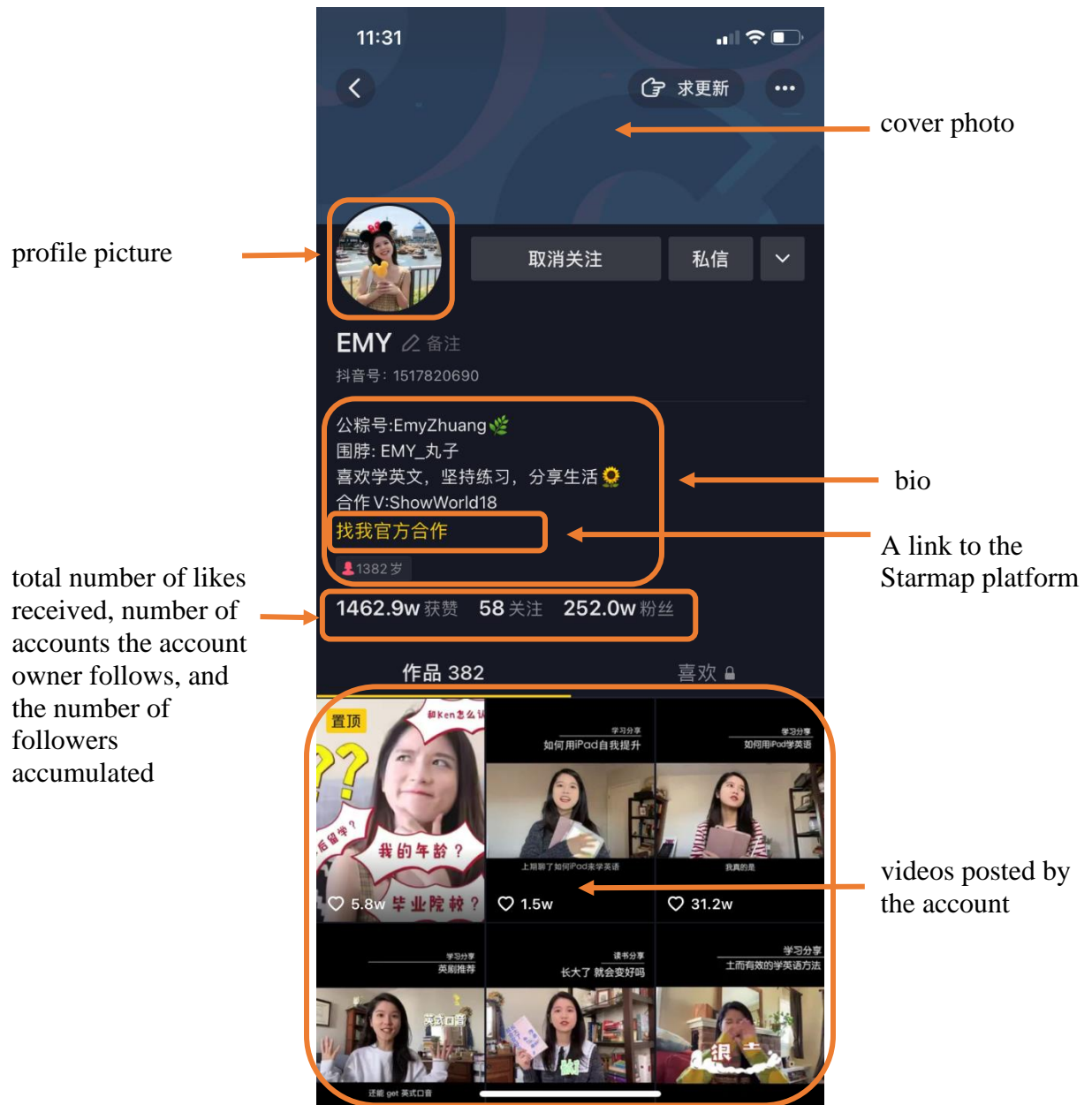
Figure 10. A Screenshot of @ EMY's *Douyin* profile page

Table 8. Information Emy includes on her *Douyin* profile page

Information category	Content
Cover photo	The default *Douyin* cover photo
Profile picture	A photo of Emy in a Walt Disney World theme park smiling at the camera. She is wearing a strap top and a Minnie mouse headband and holding a Mickey Mouse shaped popsicle.
Bio	"Official account[27]: EmyZhuang 🌿 (a herb emoji)[28] Weibo[29]: EMY_丸子 Like learning English, keep practicing, share life 🌻 (a sunflower emoji) V[30] for business inquiries: ShowWorld18" Emy also includes her gender (female) and age (1382) in her bio, although the latter one cannot be true.
Starmap (influencer platform)	A link to the Starmap platform

persona is also indexed in the two emojis she includes in her bio. The first one is a green herb emoji (🌿). Green may be associated with health, nature, and youth (Lirola & Chovanec, 2012), and herb as a plant also has similar indexical meanings. The other emoji, which Emy adds after she expresses her purpose of sharing her life through short videos, is a sunflower emoji (🌻), which is an emblem of positivity and vibrance in Chinese culture.

As a *Douyin* microcelebrity, Emy also increases her presence by sharing her account information on other popular Chinese social media sites including *WeChat* and *Weibo*. Furthermore, she indicates her willingness to cooperate with brands by providing two ways

---

[27] A shortened form for "*WeChat* official account". Emy uses 公粽浩 (gong zong hao) to refer to 公众号(gong zhong hao, meaning "official account") because these two expressions are phonetically similar in Chinese.
[28] I searched Emy's official account on WeChat and found that the herb emoji is not part of Emy's WeChat official account name.
[29] In the bio, 围脖(weibo, meaning "neck warmer") is used as a homophone for 微博(*Weibo*).
[30] "V" is used to refer to *WeChat*

(*WeChat* and Starmap) for business inquiries. Perhaps another interesting thing in Emy's bio is the age information. Emy puts an unrealistic age of 1382, which might serve as a playful response to those who are curious about her age. It might also suggest her unwillingness to disclose her age.

9) @Chris 带你猛撕英语 (Ripping through English Classes with Chris)

The next account is @Chris 带你猛撕英语 (Ripping through English Classes with Chris). In this account name, "Chris" is the name of the account owner. The expression 猛 (violently)撕 (tear) 英语 (English) is also a Chinese online slang term to refer to having a fierce fight with someone and winning it. In the account name, Chris combines 猛撕 with 英语 (English), through which English is personified. As such, the expression "猛撕英语" creates an image of battling with English fiercely and defeating it, and therefore can be translated loosely into "ripping through English classes". The adoption of a Chinese online expression and the rhetorical device of personification in the account name creates playfulness and contributes to a humorous persona. In addition, Chris also establishes a humorous identity in his bio. Specifically, in the first line of the bio, Chris puts "A witty talker and a fitness enthusiast make a good English teacher. 🏋 (a man lifting weights emoji)". In so doing Chris constructs a strong and competitive image from the aspect of personality, personal interest, and profession simultaneously.

To establish connections with the audience, Chris uses his cover photo to advertise a fan group established for his followers who are interested in learning English with him. The cover photo features the text "Learn English with fun! Here you can find the group for Chris's fans to discuss English learning. Come and join!". The fun learning approach Chris indicates in the

message is also associated with the humorous personality he projects. In addition, the profile

picture features Chris boating with others, which constructs him as a gregarious, outdoor person.

Figure 11. A screenshot of @ Chris 带你猛撕英语's (Ripping through English Classes with Chris) *Douyin* profile page

Table 9. Information Chris includes on his *Douyin* profile page

Information category	Content
Cover photo	A screenshot of the information area on the homepage of Chris's *WeChat* official account, in which Chris's *WeChat* account name, profile picture, account information ("Learn English with fun! Here you can find the group for Chris's fans to discuss English learning. Come and join!)
Profile picture	A photo of Chris boating with several people
Bio	"A witty talker and a fitness enthusiastic make a good English teacher. (a man lifting weights emoji); To learn English, add 'mengsiyy04' as a contact on V[31] (for individual learners) For business inquiries, add 'holywack' as a contact on V (please include the reason why you are reaching out)." Chris also includes his gender (male) and age (103) in his bio, although the latter one cannot be true.

In the photo, although Chris is angled away from the plane of the viewer, his head and gaze are still turned towards the camera. The visual representation conveys the message that while Chris is aware he is not actually involved in the audience's world, he nevertheless reaches out to them to establish connections. Moreover, he lists two different *WeChat* accounts in the bio and specifies what each account is for: one is for individuals who want to learn English with him, and the other account is for business inquiries from brands. As for the age information Chris puts in bio, similar to Emy, the age of 103 Chris claims cannot be true but is used as a playful response to indicate his intention to keep his age as private information.

---

[31] 'V' refers to *WeChat*.

10) @Peter 教口语 (Peter Teaching Spoken English)

The account @ Peter 教口语's (Peter Teaching Spoken English) is operated by a *Douyin* microcelebrity who refers to himself as Peter. The main identity Peter constructs on his profile page is a very professionally looking English language expert and teacher. In his bio, he introduces himself first as a CEO of an English language training center, which indicates his teacher and entrepreneur identities. To convince the viewer that he is a qualified English teacher, he then mentions he scored an IELTS Speaking band score 8. Leaners of English who are familiar with the IELTS test know that the highest score one can get in IELTS Speaking is 9. Scoring an 8 in IELTS Speaking is extremely challenging for most Chinese test takers of IELTS. Therefore, Peter's IELTS score serves to qualify him as a proficient speaker of English. In addition, Peter claims he is a bilingual host, which futher speaks for his English proficiency. To enhance the credibility of this information, Peter chooses a photo featuring him hosting an event as his cover photo. Furthermore, the Influencer Personal Verification badge on the profile page shows Peter's *Douyin* account is verified as an "Educational Self-Media", contributing to his teacher identity as well.

The profile picture features a professional portrait of Peter, which aligns with the teacher/expert and entrepreneur personae he constructs in the bio and cover photo. In the profile picture, Peter looks directly at the audience and a smile, through which he invites viewers into an imaginary relation with him and conveys a sense of connectedness. To provide learners with different ways to reach out to him and advertises his language courses, he also encourages viewers to sign up for his demo classes for beginners on *WeChat*.
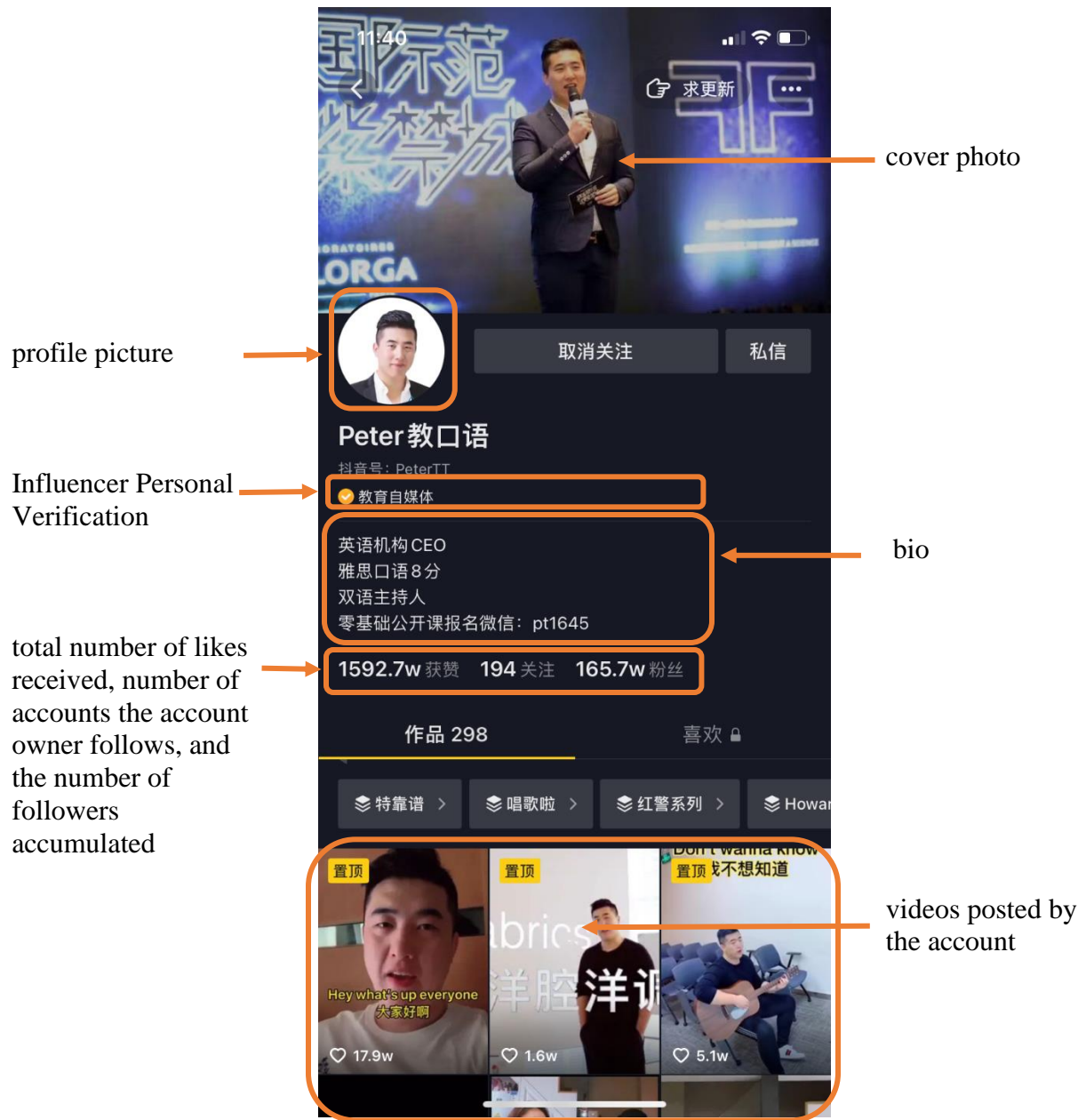
Figure 12. A screenshot of @ Peter 教口语's (Peter Teaching Spoken English) *Douyin* profile page

Table 10. Information Peter includes on his *Douyin* profile page

Information category	Content
Cover photo	Peter in a business suit appears to be giving a professional presentation at a business event. The words appearing on the backdrop banner include "international style" and "the Forbidden City".
Profile picture	A professional headshot of Peter.
Bio	"A CEO of an English language training center; Scored an IELTS Speaking band score 8; A bilingual host; To sign up for demo classes for beginners, add "pt1645"as a contact on *WeChat*."

11) @嘉德仓自学 English 早 7:50 播 (Jiadecang Self-learning English, Stream Starts at 7:50 a.m.)

So far, I have analyzed the self-branding and advertising practices on the profile pages of the Top 10 most followed "personal" accounts among the 300 accounts of informal English language teaching I collected for the present study. In addition to these 10 accounts, I also discuss the identity construction and advertising strategies of another 4 accounts selected for some unique features they demonstrate. The first account of these four is @嘉德仓自学 English 早 7:50 播 (Jiadecang Self-learning English, Stream Starts at 7:50 a.m.). The account owner is Jiadecang, as indicated in his account. Jiadecang is a transliterated name from Tibetan to Mandarin Chinese. At the time Jiadecang's *Douyin* profile page was selected for the present study, he had accumulated about 381,000 followers. Compared to the top 10 accounts of which the number of followers range from 1.6 million to 11.7 million, Jiadecang is far less popular. However, he still acquired a minor microcelebrity status with the fanbase he has established.

One of the main identities Jiadecang constructs on his profile page is his Tibetan ethnic identity. This identity type is not only displayed as his transportable identity through his appearance as shown in his profile picture and his videos, but also established and indexed through his use of linguistic and other semiotic resources. Linguistically, his transliterated Chinese name from Tibetan is a recognizable Tibetan sounding name to many Chinese users. As for the textual information exclusively presented on his profile page, Jiadecang also highlights his Tibetan ethnicity in his bio by stating "我们自己的藏区农牧好产品推荐者" (I promote fine farm produce from our own Tibetan region). The first-person plural possessive adjective "我们" (our) is used to refer to the belonging of the Tibetan farm produce to Jiadecang and his Tibetan followers in the same region. The expression "自己" (own) is also used with the possessive adjective "我们" (our) to emphasize the association between him and Tibetan products. Additionally, Jiadecang shares his location information, which is Lhasa[32], Tibet, in his bio.

Apart from the main identity of a Tibetan person he projects, Jiadecang also displays an image of hegemonic masculinity through his cover photo and profile picture, which feature himself wearing a sports tank top, working out in a gym, and showing his muscles to present his physical attractiveness. Moreover, the content creator aligns himself with an English learner identity by stating in his bio that he enjoys learning English. He also shares with the audience his other hobbies, including working out, playing basketball, and swimming. In disclosing these personal aspects of him, Jiadecang creates a sense of intimacy with the viewers. In addition, he provides his university information in the bio, suggesting that he holds a post-secondary degree and is an educated person.

---

[32] Lhasa is the capital of the Tibet Autonomous Region in China.

Furthermore, the accessibility, connectedness, and intimacy Jiadecang signals through his account name and other information on his profile page serves to improve viewership. Similar to @晓莉英语-11:30 直播 (Xiaoli's English-Live Stream starts at 11:30), the account name @嘉德仓自学 English 早 7:50 播 *(Jiadecang Self-learning English, Stream Starts at 7:50 a.m.)* includes the content creator's stream schedule. Both vloggers also post their streaming schedule in the first line of their bio. The streaming schedule Jiadecang lists in his bio is "every day starting from 7:30a.m.", which has a slightly different starting time from the schedule indicated in his account name (7:50a.m.). This is probably because Jiadecang updated his streaming information in his account name to reflect his most recent schedule. In fact, I have observed that several content creators of the 14 selected accounts occasionally modify their account names to remind the followers of their up-to-date streaming schedule or daily broadcasting sessions. For example, Yang, who runs the most followed account in my dataset, @MrYang 杨家成 (Mr. Yang Jiacheng Yang), sometimes adds the starting time of his stream to his account name. As *Douyin* allows users to change their account name up to 4 times within a 30-day period, it is relatively flexible for streamers to uses their account name to reflect the immediate update on their streaming activity. In addition to live streaming schedule, Jiadecang further increases his accessibility by providing his *WeChat* information in his bio to encourage followers to join his *WeChat* fan group as well as to welcome business inquires. There is also a *Douyin* "fan group" feature on his profile page under the bio. Users can tap on this bottom to find the way to join Jiadecang's fan group on *Douyin*. To encourage viewers' interactions with him through text, Jiadecang refers to his followers as friends and addresses them directly in his bio by asking "I like to make friends. Would you like to be my friend?".

To the left of Jiadecang's *Douyin* "fan group" feature is his "personal store", an e-commerce feature where microcelebrities can sell their own products and services or commodities offered by third-parties. The *Douyin* store on Jiadecang's profile page shows that he was selling 20 products, although I observed that he added new products and took down some existing commodities from time to time. Some of the products he has listed on the store are Tibetan local specialties, such as yak meat[33] jerky, high land barley[34] snacks, caterpillar fungi[35], and so forth, to which his bio information ("I promote fine farm produce from our own Tibetan region") has reference. There are also commodities from third-party sellers available on his e-commerce platform through which Jiadecang earns revenue via commissions upon each purchase of the product. These products are mainly household essentials and are irrelevant to English language learning. As will be presented in the analysis of one of Jiadecang's English learning vlogs in Chapter 6, Jiadecang's video content features an English language teaching component. However, the commodities he promotes through his profile page and *Douyin* store suggest that this particular content producer utilizes the English language performances in his videos to contribute to a public image that helps him market products that are largely irrelevant to English. This unique public character is a result of combinations of multiple identities as will be explained further in Chapter 7.

---

[33] Yak meat is the most popularly eaten meat in Tibet.
[34] High land barley, or Tibetan barley, is the principal food grain on the Tibetan Plateau.
[35] Caterpillar fungus is mainly found on the Tibetan Plateau and has a long history of being used in Tibetan and Chinese medicine. It is one of the most expensive fungi in the world.
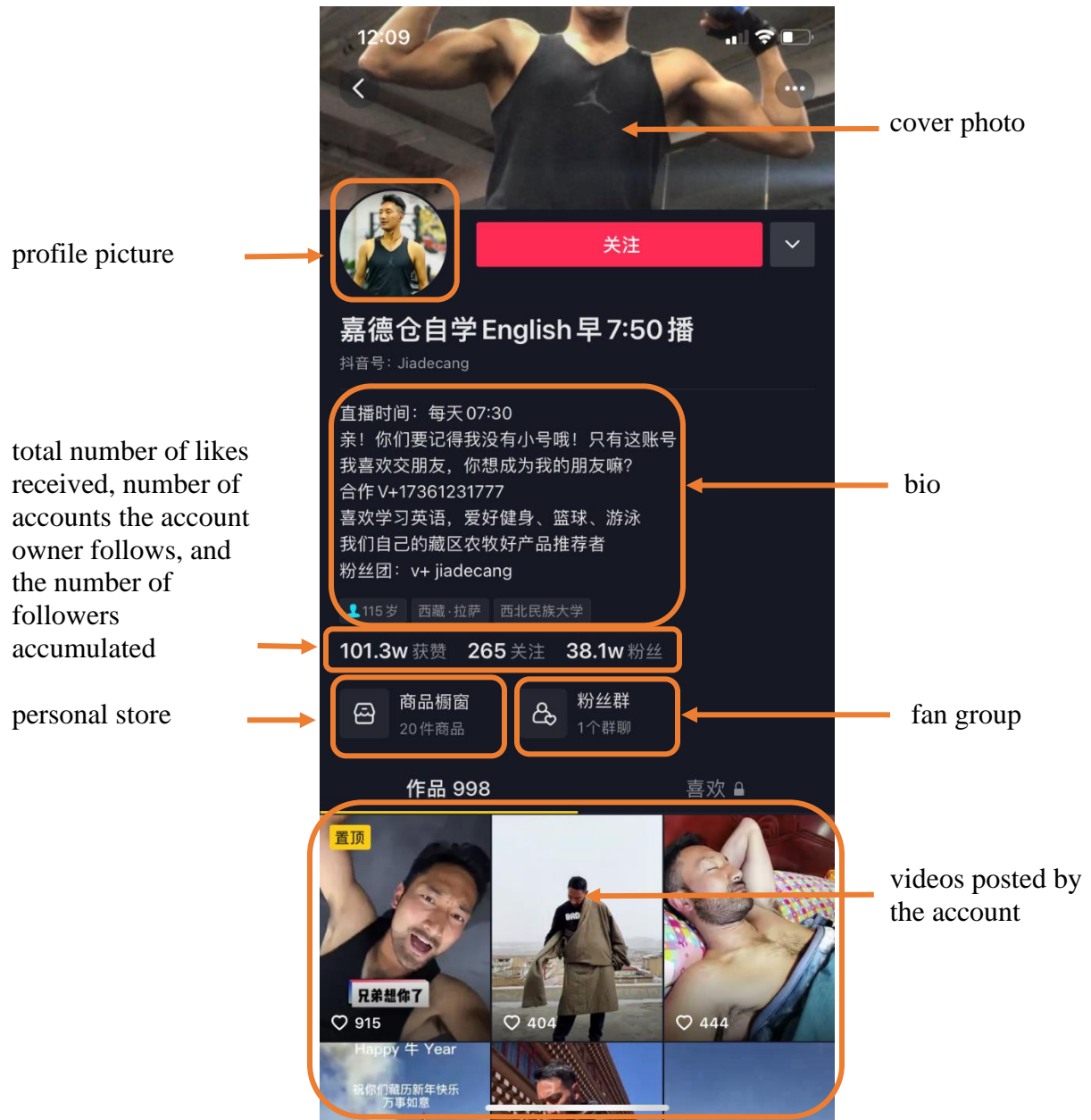
Figure 13. A screenshot of @嘉德仓自学 English 早 7:50 播's *Douyin* profile page

Table 11. Information Jiadecang includes on his *Douyin* profile page

Information category	Content
Cover photo	Jiadecang in a Jordan tank top showing his muscles
Profile picture	A photo of Jiadecang in a gym, wearing a Jordan tank top and earphones
Bio	"Stream schedule: everyday starting from 7:30a.m. Friends! Remember that I don't have other *Douyin* accounts other than this one. I like to make friends. Would you like to be my friend? For business inquiries, add "17361231777" as a contact on V[36] I like learning English, working out, basketball, and swimming. I promote fine farm produce from our own Tibetan region. Add "jiadecang" on V[37] to join my fan group Male, 115 years old; Lhasa, Tibet; Northwest Minzu University"
Personal store	The personal store feature shows Jiadecang is selling 20 products in his *Douyin* personal store.
Fan group	The fan group feature shows there is one chat group for Jiadecang's fans.

12) @一森教英语 (Yiseng Teaching English)

The next account I selected in addition to the top 10 most followed accounts is @一森教英语 (Yiseng Teaching English). Yiseng in the account name is the Chinese transliteration of the English name Ethan, which is the name of the young boy featured in the informal English language teaching videos of this account. Although Yiseng appears to be giving language instructions independently in these videos, there is a female voice in the background at times,

---

[36] Referring to *WeChat*.
[37] Again, "V" refers to *WeChat*.

supposedly Yiseng's mother who is shooting the video. Typically, Yiseng's mother's speech is included in the video in order to ask an English language question, and Yiseng answers it while looking at the camera Yiseng's mother is holding. In this sense, Yiseng's mother is the co-producer of the video content.

This co-production is also indicated through the cover and profile picture of this account, both of which feature Yiseng and Yiseng's mother (see Figure 14). However, compared to the profile picture showing a recent photo of the two people, the cover photo is a picture of Yiseng's mother holding baby Yiseng, through which Yiseng's mother's social role as a parent and a caregiver is more evident.

What is interesting about Yiseng's bio is that the information in it is predominantly about Yiseng's mother, despite the account name indicating Yiseng as the account owner. As a young boy, Yiseng is probably not capable of operating a *Douyin* account on his own. What Yiseng really does is to let his mother make videos of him giving informal English instruction while the mother takes care of everything else including selecting topics for the videos, postproduction editing, interacting with followers, and advertising products. Therefore, the actual account owner of "Yiseng Teaching English" is Yiseng's mother, which also answers why the profile page of this account centers on the identity construction of Yiseng's mother, not Yiseng.

In the bio, Yiseng's mother first provides information about her educational background, profession, and professional credentials (see Table 12), through which she indicates her middle to high socioeconomic status and high education level. Next, she lists two Chinese reality and variety shows she and her son participated. Both shows have a focus on parenting. In fact, Yiseng's mother posted a short video clip of one of the shows on *Douyin* featuring her and Yiseng singing an English song. This video is pinned and was therefore collected as one of the

84 videos in my dataset. Other information Yiseng's mother shares in the bio includes a short

description of Yiseng as a mixed-race boy of half Chinese and half British, and a *WeChat*

Figure 14. A screenshot of @一森教英语's (Yiseng Teaching English) *Douyin* profile page

Table 12. Information Yiseng's mother includes on her *Douyin* profile page

Information category	Content
Cover photo	Yiseng's mother is holding the baby boy Yiseng, smiling at the camera.
Profile picture	A photo of Yiseng's mother and baby Yiseng smiling at the camera.
Bio	" 💛 (yellow heart emoji) Yiseng's mom: holding two Master's degrees: graduated from the Harbin Institute of Technology and the University of Warwick in the U.K.; Aerospace Engineer in China; Artificial Intelligence Engineer in the U.K. (Coventry); Lecturer of Lean Production in the U.K. (Birminghan); 💙 (blue heart emoji) Reality show: performers of contestants in *Super Diva*[38] Season 7, Episode 1; Performer in *Tian Na Ni Zhen Gao*[39]. Baoshi Dong[40] was the babysitter 😂 (face with tears emoji); ❤️ (red heart emoji) Ethan: a boy of half Chinese and half British 💚 (green heart emoji) Business inquiries: yoyoquuk[41]"
Starmap (Influencer platform)	A link to the Starmap platform
Personal store	The personal store feature shows Yiseng's mother is selling 10 products on her *Douyin* personal store.

account for business inquiries. Another interesting aspect about the bio is the way Yiseng's

mother visually presents the information. As shown in Figure 14 and Table 12, four heart emojis

of different colors are used to function like bullet points, leading the four sections (i.e., Yiseng's

mom, TV show, Yiseng, and business inquiries) of the bio information Yiseng's mother

categorized. The detailed description in each section is indented. In so doing the bio information

---

[38] A Chinese reality show on which mothers demonstrate their talents and share stories about parenthood.

[39] It seems this variety show does not have an official English name. Therefore, I use the Pinyin of the Chinese name of this show "Tian Na Ni Zhen Gao" to refer to it, which can be roughly translated as "My god you are amazing".

[40] A permanent guest of the variety show "Tian Na Ni Zhen Gao" and is known as a Chinese rapper.

[41] I have searched this on *WeChat* and confirmed that it is Yiseng's mother's *WeChat* ID.

becomes more salient to attract the viewer's attention and its clear formatting provides audiences with a better viewing experience. Furthermore, the four hearts emojis are used to convey affection and establish intimacy with the followers.

Additionally, Yiseng's mother includes the Starmap feature on the profile page to welcome business cooperation. The *Douyin* store feature is embedded in the profile page as well. There are 10 products listed in the store, including English picture books, books of English reading for children in primary school, dictionary pens, and some other commodities (such as household essentials) not relevant to English language learning.

13) @英国小克里 (British Xiaokeli)

The next account I focus on is @英国小克里's (British Xiaokeli) run by a British named Chris. He also refers to himself as 小克里(Xiaokeli) in *Douyin*. The Chinese expression "xiao" (小) literally means "little". It is also often used to address a young acquaintance by combining it with a person's last or first name. "Keli" is the transliteration of the English name "Chris". Therefore, the expression "Xiaokeli" creates a sense of closeness.

There are two main identities Xiaokeli constructs on his Douyin profile page. The first one is a bilingual expert. In his bio, Xiaokeli first demonstrates his language competence in Chinese by providing his educational background as a graduate of the Chinese Studies program at the University of Cambridge and specifying that he has been learning Chinese for 10 years. To prove the information he provides about his educational background is real, he uses a photo featuring him kneeling and clasping before the Vice Chancellor, which is a ritual for transforming students into graduates during the graduation ceremony at the University of Cambridge (see Figure 15). As such, another layer of his identity emerges. As a graduate of one of the most prestigious universities in the world, Xiaoleli projects an image of an elite student,

which contributes to his construct of an expert identity. He also claims that he is a former student of Kevin Lin, who is known as "the royal interpreter" to Queen Elizabeth II, which indicates his expertise in English and Chinese interpretation. Additionally, the profile image features a business portrait of Xiaokeli in formal attire, denoting a sense of professionalism.

The second identity Xiaokeli establishes is someone who knows well about China and affiliates with the Chinese people. Specifically, he states in the bio that he worked and lived in China for 6 years, married a Chinese woman from Shangdong province, and has a daughter. To encourage interactions with viewers and further establish affiliations, he also invites audiences to discuss Chinese and British culture, as well as the English language. His familiarity with both cultures also indicates his cross-cultural expertise. Another aspect of information Xiaokeli adds to his bio is his age. Different from most other microcelebrities in my dataset who prefer not to disclose their private information, Xiaokeli shares his personal information in order to create sense of intimacy with the audience. This decision is perhaps motivated by his non-Chinese identity. As someone who tends to be considered as "other" by Chinese users, Xiaokeli strives to build rapport and affiliation with the audience.

To promote his language products/services, he included his *WeChat* account in the bio description. There is also a *Douyin* "personal store" feature on his profile page, but no products are listed in the store. This is probably because the time I sampled his profile page happened to be when he was updating items in his store, and therefore nothing was listed at that moment. However, I have observed Xiaokeli has continued using *Douyin* store to sell his language video lessons.

Figure 15. A screenshot of @英国小克里's (British Xiaokeli) *Douyin* profile page

Table 13. Information Xiaokeli (Chris) includes on his *Douyin* profile page

Information category	Content
Cover photo	Chris in his graduation ceremony at the University of Cambridge going through the rituals for transforming into graduates including kneeling and clasping before the Vice Chancellor.
Profile picture	A business portrait of Chris
Bio	"Graduated from the Chinese Studies program at the University of Cambridge, I have been learning Chinese for over 10 years; Worked and lived in China for 6 years; Participated in the exchange programs at the Ocean University of China and Zhejiang University; worked as a teacher in Suzhou, Former student of Kevin Lin, "the royal interpreter" to Queen Elizabeth II Married a girl from Shandong Province, have a 2-year-old daughter; I talk about Chinese and British culture, and the English language with people on *Douyin*; Thank you everyone for your support; To learn English, add 'xiaokeli111' as a contact on V[42]; Male, 33 years old, Harbin Cambridge University"
Personal store	Although the personal store feature is displayed on Chris profile page, there is no product listed in his personal store.

14) @ 🇺🇸Allie 爱丽 💕 (<the US flag emoji>, Allie Aili, <two hearts emoji>)

The last profile page I discuss is from the account @ 🇺🇸 Allie 爱丽 💕 's (<the US flag emoji>, Allie Aili, <two hearts emoji>) that has about 550 thousand followers. The account owner is Allie, and "Aili" in her account name is transliteration of her English name. Different from the previous 13 accounts, Allie incorporates emojis into her account name. This first emoji is the US flag, a semiotic resource that indexes Allie's nationality, and the second emoji of two hearts is used to convey affection to the viewer.

---

[42] Refering to *WeChat*.

In addition to the US flag emoji, Allie's cover photo features Golden Gate Bridge in San Francisco, which is also a visual index of her nationality. Furthermore, she also adds a location

cover photo

profile picture

total number of likes received, number of accounts the account owner follows, and the number of followers accumulated

bio

videos posted by the account

Figure 16. A Screenshot of @ 🇺🇸 Allie 爱丽 💕 's (<the US flag emoji>, Allie Aili, <two hearts emoji>) *Douyin* profile page

Table 14. Information Allie includes on her *Douyin* profile page

Information category	Content
Cover photo	A photo of Golden Gate Bridge in San Francisco.
Profile picture	A photo of Allie looking back and smiling at the camera
Bio	"I'm Allie. I'm in ninth grade. I share authentic everyday English and create real-life, vivid English learning context to help you improve speaking and listening fast. English is my native language. I'm learning Chinese and Spanish. I will do my best to post new videos on Wednesdays and Saturdays. I am very busy with schoolwork and do not have time to respond to messages as there are too many. I have turned off direct messages. Business email: AlliesEnglish@yahoo.com 15 years old, United States"

label of United States to her bio. Same as Xiaokeli, Allie shares her age information and tells the audiences that she is a ninth-grader. However, Allie does not seek to establish intimacy and affiliation as much as Xiaokeli. She explains in her bio that her schoolwork is intense, and she does not have time to respond to messages from the viewers. She also informs the viewer that she has turned off direct message. Nevertheless, Allie still creates a sense of connectedness to the audience by providing the schedule she tries to follow to post new videos. The profile page that features Allie smiling and looking at the camera invites viewers build connections with her too.

To promote her account, Allie states in her bio that she teaches authentic everyday English and creates real-life, vivid English learning context to help leaners improve speaking and listening fast. In addition, she explicitly states that English is her native language, and she is also learning Chinese and Spanish. The information of her language background Aili shares constructs her English native speaker identity and projects an image of a multilingual user. Moreover, she includes her email address in the bio to welcome business inquiries.

Cross-case Analysis of RQ1

So far, I have examined how each of the 14 selected *Douyin* microcelebrities of informal English instructional videos harnesses various multimodal resources on their profile pages to create a public image and promote their products/services. It is found that many of those content producers construct multiple identity roles, including an English language teacher/expert, a multilingual, a successful entrepreneur, a masculine, attractive, and powerful man, a dedicated parent, a talented, humorous individual, an approachable person, and so forth. Each identity serves a specific communicative goal. For example, an English language teacher/expert identity helps the content creators to establish credibility in English teaching. A successful entrepreneur image serves to speak for one's competence and to establish authority. The multiple identities they construct work together to project a unique public image that differentiates them from other content creators of English language instruction videos, which therefore enhances their competitiveness.

In addition, the identity construction of *Douyin* microcelebrities is found to emphasize the relationship of vloggers with their audiences. The relationship is heavily dependent on credibility, connectedness, and accessibility content creators establish with their followers. To communicate a sense of closeness and encourage interactions, many content creators share their other social media accounts, list their live stream schedule, and encourage followers to join their fan groups.

Furthermore, it is observed that Chinese vloggers and non-Chinese content creators differ in the ways they adopt to establish connections with followers. While Chinese microcelebrities perform connectedness and accessibility by providing followers with different ways to interact with them, they in general do not reveal as much private information as those non-Chinese

124

vloggers. For example, most of the Chinese microcelebrities in my sample do not disclose their

real age on their profile page, and two of them even marked their age as 1382 and 103 years old

as a playful way to indicate their unwillingness to question age-related questions. Unlike in many

Western countries, talking about one's age is not a cultural taboo in China. In fact, it is common

practice among Chinese people, especially older generations. However, a majority of Chinese

microcelebrities in my sample do not provide their real age information. Instead, they tend to

prioritize the professional over the personal aspects of identity, focusing more on projecting an

English language expert/teacher identity. On the other hand, Chris from the U.K., and Allie from

the U.S., provide their (presumably real) age information in order to conform to the Chinese

cultural norms and perform affiliation. This distinction between Chinese vloggers and non-

Chinese content creators in their identity construction is shaped by the different elements they

rely on to establish authority in ELT. To construct an English language teacher/expert identity,

many Chinese content creators share their educational background, work experience,

professional credentials, and so on in order to prove that they are proficient in English. On the

other hand, three out of the four Caucasian vloggers in my sample (except Xiaokeli) do not

provide any information about their professional training relevant to English. Their native

speaker identity and authority in ELT are indexed in their race.

Apart from their identity construction practices, the 14 microcelebrities of informal

English pedagogical videos utilized various features and multimodal resources on their *Douyin*

profile pages to monetize their following. A majority of them use *WeChat* as one of the main

social media platforms for English learning and/or business inquiries. Although *Douyin* has its

own messaging feature, *WeChat* is the most dominant messaging mobile app and also used as a

popular mobile payment tool by Chinese consumers. As such, many transactions between

Douyin microcelebrities and customers take place on *WeChat*, which leads to *Douyin*'s censoring of it. However, the microcelebrities in the present study managed to evade the commercial censorship by adopting non-standard written practices such as using emojis, homophones, and English letters. In addition to *WeChat*, several content producers also exploit *Douyin*'s live streaming and e-commerce features to sell their products and services.

CHPATER FIVE: RQ2 FINDINGS

Before investigating how content creators of selected *Douyin* accounts of informal English instructions marshal multimodal resources to teach English in their videos, it is first necessary to find out what exactly they teach about the English language. The relevant findings have potential to inform online language teachers' choices of topics, language focus, and teaching methods with the consideration of digital affordances. Therefore, this chapter is devoted to addressing this aspect, as the study's second research question:

RQ2: What kinds of English language topics or other topics are addressed in the videos and in what ways these topics are addressed?

To better capture any trends in the selection of topics, language components (grammar and vocabulary) and skills (e.g., listening, speaking, reading, writing, pragmatics), as well as teaching styles and methods that content creators adopt, I break down this research question into the following three sub-questions:

a. What content topics are discussed?

b. What English language components and skills are being taught?

c. What are the teaching styles and methods used?

Some videos in my dataset appear to be more focused on one specific language component or skill, while other videos tend to cover several aspects in the content. However, these components tend to be addressed in various degrees and there is usually one language aspect or skill that is the main focus of the video. Therefore, when categorizing the language

components and skills taught in these videos, I only identified the most salient element and assigned one primary category to each video.

Information of the Selected Videos from the 14 Sampled Accounts

Table 15 presents information of the 6 selected videos from the account @MrYang 杨家成 (Mr. Yang Jiacheng Yang), the owner of which is Jiacheng Yang, who projects himself as an English language teacher, an entertainer, and an attractive and versatile man on his profile page.

Table 15. Information of Yang's six selected videos

	Topic	Language skills	Teaching styles/methods
Video 1 (pinned)	Airport English: how to go through customs	Speaking (pragmatics)	Role play; lecture
Video 2 (pinned)	How fast can Yang speak English	N/A	N/A
Video 3 (pinned)	Why is it a bad idea to sing an English song with lyrics directly translated from English into Chinese	N/A	N/A
Video 4	Yang dubbing for a movie in English	N/A	N/A
Video 5	Why is it a bad idea to sing an English song with lyrics directly translated from English into Chinese	N/A	N/A
Video 6	Introducing himself as an English teacher through singing a song with Yang's self-made lyrics	N/A	N/A

Among the 6 selected videos from Yang, only Video 1 is clearly pedagogical. The video includes a series of role play episodes featuring conversations between Chinese tourists and a U.S. immigration officer (all roles played by Yang) at an international airport in the English language. In addition to the role play, Yang as a lecturer explains why some responses to the immigration officer's questions exemplified in the role play are considered inappropriate and

offers suggestions on answering the custom questions. The main language skill Yang focuses on teaching is speaking, as the video centers on conversations between a tourist and an immigration officer. More specifically, the pragmatic aspect of speaking is mostly addressed. Rather than explicitly discussing vocabulary or pronunciation, what Yang highlights during the lecture part of the video is on how to formulate appropriate responses as a Chinese tourist with consideration of cultural and bureaucratic elements.

Different from Video 1, Yang's second video is not primarily instructional. Although Yang does speak English in this video, teaching does not seem to be one of his main purposes. Rather, Yang uses this video to demonstrate his English proficiency. The video was made in response to a hashtag challenge on *Douyin*, which could be roughly translated as "#English fluency challenge". On both *Douyin* and *TikTok*, a hashtag challenge is a trending activity that involves creating a catchy hashtag along with a video that exhibits a specific action and invites others to complete the same challenge. I have observed that at the time when Yang initially posted the video, which is January 2nd, 2021, the "#English fluency challenge" was a very popular hashtag and had been used as the theme of numerous videos by various Douyiners. The challenge originated from several viral videos made by a *Douyin* content creator of English language teaching, who in those videos attempted to showcase her fluency in spoken English by reading a list of English sentences with as little time as possible in one breath. However, her action of reading nearly 10 sentences in 5 to 7 seconds was in general, not well received by other *Douyin* users, especially other content creators who also made videos of English instruction videos. Many criticized her way of teaching fluency as ineffective and misleading, and her fast-paced speech was at the expense of pronunciation accuracy, which makes her utterance incomprehensible. Nevertheless, a number of content creators of English language teaching,

129

including Yang, posted videos in response to this challenge, but mostly in a parodic way. In Yang's Video 2, he first explicitly refers to the "#English fluency challenge" and one of the viral videos by embedding part of the video into the beginning of his one, which is also known as the "Stitch" feature on *TikTok*, and pretends he is ready to read the 8 sentences in the original video. Then he quickly makes a few random sounds, indicating that he is talking gibberish, and the timer he holds shows that he finished his utterance in 0.45 seconds. However, right after his parodic performance, Yang actually takes the challenge this time to introduce himself in English as a "professional English teacher" in 6 sentences within a few seconds. Despite the fast pace, Yang's utterance is clear and articulate. By doing so, he demonstrates his English expertise, indicates the importance of both fluency and accuracy in speaking, as well as entertains his viewers.

Video 3 and Video 5 are similar in content. Both videos focus on showing why it is a bad idea to sing an English song with lyrics directly translated from English into Chinese. Each video features Yang singing a song, one from Maroon 5, an American pop rock band, and one from Chris Wu, a Chinese Canadian rapper. However, Yang sings both songs in Mandarin Chinese by directly translating the English lyrics into Chinese, during which the rhymes in the English lyrics are lost. In addition, the colloquial style of the original lyrics becomes more noticeable to a larger audience when it is translated into Chinese. The colloquial expressions of the lyrics in Chinese make the songs further amusing since Chinese lyrics are, in general, expected to be more poetic and formal. Yang's performance in these two videos is not primarily pedagogical oriented. Instead, he entertains the viewers through his good singing skills and awkward lyrics.

Video 4 features Yang dubbing for a short video clip from a Hong Kong movie in the English language using "Mock Chinese", which refers to parodically revoicing a

130

Chinese/Cantonese speaker of English (see more detailed discussion in Chapter 6). Again, although Yang does speak English in this video, his main purpose is to generate humor and showcase his English language ability.

Similar to Video 3 and 5, Video 6 includes Yang singing an English song (from Akon, a Senegalese American singer) with Chinese lyrics. However, Yang does not directly translate the English lyrics into Chinese this time. Instead, he uses self-composed Mandarin and Cantonese lyrics in which he introduces himself as a respected English teacher and describes his personality. The lyrics rhyme well, which makes the song pleasant to listen to. Through singing and the lyrics, Yang again projects himself as an English language expert and an entertainer with talent in singing.

Although only one video clearly focuses on English language teaching, all of the six selected videos serve to present Yang as an English language expert in various ways, ranging from demonstrating his expertise in teaching and using English, as well as translating between English and Chinese, to directly advertising himself as a popular English teacher on *Douyin*. Furthermore, his entertainer identity is consistently revealed in all the six videos through his humorous performance in role play (see Case Study 1 in Chapter 6) and singing, and intertextual references to Chinese online culture and Western popular culture. The identities Yang performs in those videos appear to align with the image he presents on his profile page.

Table 16 presents information of the 6 selected videos from the account @大白外教英语 (Foreign Teacher Dabai's English). The account owner, Dabai, constructs a public image as a native speaker teacher, a Sinophile, and a generous individual on his profile page.

Table 16. Information of Dabai's six selected videos

	Topic	Language skills	Teaching styles/methods
Video 1 (pinned)	The story of how Dabai's impression of China has changed over time	N/A	N/A
Video 2 (pinned)	Advertising his live class of English on *Douyin*	N/A	N/A
Video 3 (pinned)	Provide suggestions for people suffering from depression speaking from his own experience	N/A	N/A
Video 4	Adding a word to a sentence might lead to a completely different meaning and cause misunderstanding.	Grammar	Role play
Video 5	An inspirational quote about life	Speaking	Reading aloud
Video 6	The meaning difference in words with similar spellings	Vocabulary	Reading aloud

The first three videos of Dabai are pinned ones, but none of them are essentially pedagogical. In Video 1, Dabai talks about how his impression of China has changed over time. Specifically, he mentions how his previous opinion about China was swayed by Western media's negative coverage, and how his unfavorable view has changed after he started to know about China by himself. The story Dabai shares conveys a sense of affinity with the viewers as a non-Chinese person. Moreover, his speech is in Mandarin Chinese the entire time, which also serves to enhance the connection between the content creator and the audience.

The second pinned video features Dabai using Mandarin Chinese to advertise his free live class of English on *Douyin*. In the video, Dabai emphasizes that although many other people are

also streaming English classes on *Douyin*, his live class stands out because as a native speaker, he teaches authentic English only. In addition, he frames his English teaching activity on *Douyin* as a way to exercise charity, reducing the disparity in English resource distribution between urban and rural areas.

In Video 3, Dabai discloses his suffering from depression and encourages people with depression to talk about their illness and get help. Different from the first two pinned videos, Dabai talks in English in Video 3 with both English subtitles and translated captions of Chinese superimposed on the video, which makes the video a potential English learning resource. However, the main purpose of this video is to raise people's awareness of depression in the Chinese context where mental illness is often stigmatized and ignored. By revealing his medical conditions to the public and helping others with depression, Dabai arouses viewers' sympathy and establishes an image of an ordinary person and a friend who empathizes.

The next three videos are unpinned, yet all are pedagogically oriented. In Video 4, Dabai uses two role-play conversations to perform how collocations that differ in one word alter the entire meaning of a sentence. Metalinguistic explanations are also provided during the role play to help viewers understand the grammar points covered in the video. Video 5 features Dabai reading aloud an inspirational English quote about life, with both English and Chinese captions added to the video. Similar to Video 3, Dabai also speaks English in Video 5 with bilingual subtitles. However, I identify Video 5 as clearly instructional and rule out Video 3 because Video 5 is much shorter and easier for language learners to follow. On the other hand, although Video 3 can be utilized as a potential learning resource, its main focus is still on raising people's awareness of depression. In addition, the larger amount of information Video 3 includes might

also make it less likely to draw learners' attention to the English expressions, especially considering that Dabai in this video does not highlight any particular language use in any means.

Video 6 is one of the videos I selected for a detailed multimodal analysis (see Case Study 2 in Chapter 6). This selected video has a clear focus on vocabulary. Specifically, Dabai compares the meaning of several pairs of English words that are similar in spelling through reading aloud the words and their Chinese meanings.

As previously mentioned, one clear distinction between Dabai's pinned videos and unpinned ones is that none of the three pinned videos are primarily instructional, while all the three unpinned videos center on English language teaching. In the three pedagogical videos, the main identity aspects Dabai displays are an English language expert/teacher through teaching English and a non-Chinese who loves China through his use of Mandarin Chinese. These two identity aspects are also constructed in Dabai's three pinned videos. However, the image Dabai projects in pinned videos is also more multifaceted, including a generous person, an intimate friend, and a native speaker of English as well. It is also worth noting that the identities exhibited through Dabai's pinned videos are consistent with the public image established on his profile page. In addition, the identity representation in pinned videos is often realized in a more explicit way through direct advertising. For example, to promote his English language expert identity in Video 2, Dabai verbally addresses his native speaker identity and associates it with authentic English. While in his unpinned videos that are geared toward teaching, his native speaker identity is not specifically constructed, but indicated through his race as part of his transportable identity.

Table 17 summarizes information of the 6 selected videos from the account @晓莉英语-11:30 直播 (Xiaoli's English-Live Stream starts at 11:30). The account owner, Xiaoli, displays a public self as an English language teacher and an entrepreneur on her profile page.

Table 17. Information of Xiaoli's six selected videos

	Topic	Language skills	Teaching styles/methods
Video 1 (pinned)	Sharing her English learning experience and highlighting what she has accomplished, and promote her live English classes	N/A	N/A
Video 2 (pinned)	Should women prioritize their family or themselves?	Speaking	Role play
Video 3 (pinned)	Sharing her story of learning English and addressing what she has accomplished, and promote her live English classes	N/A	N/A
Video 4	Words in which letter "b" is silent	Speaking (Pronunciation)	Reading aloud
Video 5	How to ask people to repeat what they said	Speaking (pragmatics)	Reading aloud
Video 6	Vocabulary related to surrogacy and why Xiaoli strongly advocates against it	Vocabulary	Reading aloud

Among the six selected videos from Xiaoli, the first three are pinned and stay on the top of the video list. In terms of topic and content, both Video 1 and 3 feature Xiaoli sharing her own English learning experience, encouraging learners, displaying her professional credentials and accomplishments, and advertising her live English classes. She relates to the learners by talking about her struggles with English learning and self-identifying as a "daughter of a farmer" to

indicate her economically disadvantaged background. Then she reassures learners they can be as successful as she is in learning English, and establishes her trustworthiness and English language expertise by listing her awards and accomplishments.

Video 2 is the only instructional video among the three pinned ones. This pedagogical video is structured as a role play, in which Xiaoli and a man perform as a couple, and a Caucasian lady as a fitness center salesperson approaches the couple to sell gym membership. In order to convince Xiaoli of the necessity of gym membership, the salesperson tries to make Xiaoli believe that she will end up having a cheating husband who uses up her money on his lover once she loses her shape. Although Xiaoli in the role play initially fully trusts her husband, she eventually agrees to purchase a gym membership after the Caucasian lady proves to her that her husband is a real cheater. The role play conversation is done in the English language mainly, with both English subtitles and translated captions of Chinese superimposed on the screen. Keywords and phrases are highlighted with a different font color, which also helps frame the video as instructional. One of the main discourse identities Xiaoli establishes in the role play is a feminist who defends her own rights. However, it is also worth noting that her construction of feminism is also based on the depiction of men as faithless and unreliable, which might lead to gender antagonism. In addition, the importance of maintaining physical attractiveness to a man indicated in the role play might also reinforce gender inequality.

The following three videos, which are Video 4, 5, and 6, are all pedagogically oriented. Video 4 showcases Xiaoli reading several English words aloud and directing viewers' attention to the silent letter "b" in those words. Thus, the main skill addressed in this video is speaking, or pronunciation in particular. Video 5 is framed as a lecture, during which Xiaoli shares a few English expressions that can be used to politely ask people to repeat what they said. Near the end

136

of both Video 4 and 5, Xiaoli promotes her live English classes on Douyin, emphasizing that viewers will have the opportunity to learn with her for free.

Similarly, Video 6 is instructional and introduces a few English words and expressions to the viewers. Interestingly, those expressions are all related to surrogacy (e.g., "surrogacy," "surrogate mother") and infrequently used in everyday communication. However, Xiaoli chooses those words because, at that time, Zheng Shuang, a Chinese actress, was accused of abandoning her two babies born to surrogate mothers in the U.S. The scandal triggered heated discussions on Chinese social media and the actress received heavy criticism. As surrogacy is illegal in China, Zheng was then quickly blacklisted by the Chinese state regulators. Thus, Video 6 was made to address a trending topic to attract attention with an instructional purpose. Furthermore, Xiaoli also expresses her reason against surrogacy from a feminist perspective: the potential for exploitation of Chinese women from lower social classes who are forced and misled to become a surrogate.

It seems apparent that in addition to an English language expert identity, another main identity Xiaoli utilizes to attract viewers is gender related. Through promoting herself as a hardworking, independent, and successful woman, she gains the favor of viewers who are also concerned about gender equality in China.

Table 18 summarizes information of the 6 selected videos from the account @英语影视配音员 孙志立 (Zhili Sun the Voice Actor of English). The account owner, Zhili Sun, presents a public image as an English language teacher and a voice actor of English on his profile page.

Table 18. Information of Sun's six selected videos

	Topic	Language skills	Teaching styles/methods
Video 1 (pinned)	Voicing the Disney character Chief Bogo in the movie *Zootopia*	N/A	N/A
Video 2 (pinned)	Voicing a promotional video for a joint mission of the Chinese Academy of Science and the European Space Agency	N/A	N/A
Video 3 (pinned)	Voicing the Pixar character Mr. Incredible in the movie *The Incredibles*	N/A	N/A
Video 4	The importance of associating English phonetic alphabets with the specific articulatory features of the sound that each phonetic symbol represents.	Speaking (pronunciation)	Lecture
Video 5	Improving listening skills by training ourselves to recognize sounds before looking at scripts	Listening	Lecture
Video 6	The difference between the expression of "Happy Christmas" and "Merry Christmas"	Sociolinguistics	Lecture

To prove that he is a real voice actor and exhibits his English language proficiency, all the three pinned videos of Sun feature him doing voice acting. None of the pinned videos are specifically designed to be instructional. However, the three unpinned videos all have a teaching focus. In Video 4 and 5, Sun stands next to a screen projecting his lecture slides. Specifically, Video 4 trains learners to associate articulatory features (e.g., lip rounding, place of articulation) with phonetic alphabets, and Video 5 addresses the importance of recognizing sounds before looking at scripts in listening practice. Still in the form of a lecture but in a less formal setting without the lecture slides on the screen, Sun explains the difference between the expression of

138

"Happy Christmas" and "Merry Christmas" from a sociolinguistic perspective. Similar to Dabai's six selected videos, Sun's pinned videos center on constructing himself as a professional voice actor and an English expert and are not specifically tailored to teach English. The unpinned videos, on the other hand, are all pedagogical and highlight Sun's English language teacher identity.

Table 19 provides information of the 6 selected videos from the account @口语控 (Spoken English Fanatic). The account owner identifies himself as an authority in the English language and an educator on his profile page. There are only two pinned videos, one of which teaches viewers common English contractions while the other features the content creator performing different accents around the world. Clearly, the first one is an informal instruction video. Although the second video can be used to familiarize learners with various English accents, it is essentially created to demonstrate the content creator's expertise in English and entertain the audience through an exaggerated and parodic performance. The rest of the four videos are unpinned. Videos 3 and 4 are instructional, as the content creator lectures on alternative ways to express "I think" and the meaning of "tea" as an internet slang term. Videos 5 and 6 are not categorized as pedagogical. Video 5 was posted on Chinese New Year's Eve, 2021, and features the content creator sending new year wishes and positive vibes. Although the speech is in English, it is not accompanied by any subtitles like other instructional videos from this account, making it less pedagogically focused. The content creator speaks English in Video 6 as well, which is a vlog of him having his breakfast in a café. It is not identified as an English teaching video for the same reason—the absence of subtitles and instructional language. Rather, Video 5 and 6 serve to construct the content creator as a proficient English user and establish connections with viewers.

Table 19. Information of @口语控's (Spoken English Complex) six selected videos

	Topic	Language skills	Teaching styles/methods
Video 1 (pinned)	Speak more like a native by using English contractions	Speaking (pronunciation)	Lecture
Video 2 (pinned)	Performing different accents around the world	N/A	N/A
Video 3	Alternative ways to say "I think" in English	Speaking (pragmatics)	Lecture
Video 4	The meaning of "tea" as slang for "gossip"	Vocabulary	Lecture
Video 5	Sending new year wishes	N/A	N/A
Video 6	A vlog of the content creator having his breakfast in a café	N/A	N/A

Table 20 provides information of the 6 selected videos from the account @Caroline 教英语 (Caroline Teaching English). The account owner presents a self-image of a native English teacher on her profile page. The six selected videos from Caroline consist of three pinned videos and three unpinned ones. Two pinned videos are pedagogically focused, which are Video 1 and 3. In Video 1, Caroline reads several short English sentences/expressions that contain only one or two words (e.g., "How come?", "Whatever."). In Video 3, she uses minimal pairs to explain the differences between the high front tense unrounded vowel [i] and the high front lax unrounded vowel [ɪ]. Video 2 is the only non-instructional video that is pinned. This video was created to promote Caroline's English audio books that were listed in her *Douyin* store. The video features Caroline describing the content of her two audio books, and advertising them as English learning materials.

The following three videos are unpinned and instructional. In Video 4, Caroline reads aloud a conversation about ordering a birthday cake at a slow pace for viewers to follow. However, the video does not include any explicit instructions on speaking, and none of the words in the conversation is highlighted through either linguistic or other semiotic resources. Therefore, I identified the focus of this video as speaking skills in general. Video 5 features her reading English month names from "January" to "June". She reads aloud each of the month names twice in a slow and clear manner, directing learners' attention to the pronunciations. Again, in Video 6, Caroline reads aloud a few expressions that can be used to give compliments in spoken English. As giving compliments is a speech act, I coded the language skill the video centers on as pragmatics under the category of speaking.

Table 20. Information of Caroline's six selected videos

	Topic	Language skills	Teaching styles/methods
Video 1 (pinned)	Examples of short English sentences used in daily life	Speaking	Reading aloud
Video 2 (pinned)	Promoting her English audio books to language learners	N/A	N/A
Video 3 (pinned)	The differences between the high front tense unrounded vowel [i] and the high front lax unrounded vowel [ɪ]	Speaking (pronunciation)	Lecture
Video 4	English for ordering a birthday cake	Speaking	Reading aloud
Video 5	Months in English	Speaking (pronunciation)	Reading aloud
Video 6	Giving compliments in English	Speaking (pragmatics)	Reading aloud

One interesting pattern across the selected videos from Caroline is that most of her videos are with a clear teaching purpose, and she focuses on teaching speaking skills in all of the five pedagogical videos. Reading aloud is her main teaching style, and there are always English and Chinese subtitles on her videos.

Table 21 showcases information of the 6 selected videos from the account @ Scott 学口语's (Learning Spoken English with Scott). The account owner establishes an English language teacher identity on his profile page.

As shown in Table 21, five of the six selected videos are created with an English language teaching purpose. The only exception is Video 3, a pinned video in which Scott explains the meaning of the expression "call me señorita" in a popular English song *Señorita* and "ita" as a morpheme in Spanish. The video is in fact instructional as well. However, since my dissertation project focuses on English instruction videos, and Scott is primarily a content creator of English language teaching, I chose not to code the video in terms of language skills and teaching styles/methods. To learners of Spanish, this video is clearly pedagogical. However, I argue that the main purpose of this video is to highlight Scott's multilingual competence, which helps him construct a language expert identity.

For the five English instructional videos, Scott adopts the same teaching style, which is an informal lecture. In Video 1, Scott details the process of how "I am going to" is finally simplified as "Imma" in spoken English. Video 2 features him elaborating on the selection between two English prepositions "in" and "on" when they are followed by a transportation vehicle, which is a grammar point that many Chinese learners of English are constantly struggling with. Video 4 presents another short grammar lesson on adjective order in English. Video 5 has a dual focus on teaching and advertising. In the first part, Scott reminds the viewers

142

that not all online English language learning resources are reliable. Some materials might also be purposefully misleading in order to attract attention. One example he provides is that many online resources problematize the expression "How are you" as an overused way of greeting, which leads to increased learner anxiety in speaking. Scott's critiques of those online learning resources serve to distinguish him from other content creators, projecting a sincere and professional image and establishing authority. These identity aspects also help him advertise Duolingo English Test featured in the second half of the video as a valid and reliable assessment

Table 21. Information of Scott's six selected videos

	Topic	Language skills	Teaching styles/methods
Video 1 (pinned)	The meaning of "Imma" and how it is simplified from "I am going to".	Speaking (pronunciation)	Lecture
Video 2 (pinned)	The difference between "in" and "on" when these two prepositions are followed by a transportation vehicle.	Grammar	Lecture
Video 3 (pinned)	The meaning of the Spanish expression "call me señorita"	N/A	N/A
Video 4	Adjective order in English	Grammar	Lecture
Video 5	Online information/resources of English language learning might be unreliable and misleading & advertising Duolingo English Test	Speaking	Lecture
Video 6	Difference between the meaning in "chicken legs" and "chicken drumsticks"	Vocabulary	Lecture

of one's language proficiency. The last video is about vocabulary, specifically, the meaning difference between "chicken legs" and "chicken drumsticks". As a former English language

learner and teacher in China, I have observed that many of the instruction videos Scott creates address common confusions and mistakes that Chinese learners have. The selection of topics of this kind not only indicates Scott's English language proficiency and teaching expertise, but also helps increase views and attract followers.

Table 22 summarizes information of the 6 selected videos from the account @EMY. The account owner projects a positive and humble persona and an English learner identity on her profile page.

Emy has only one pinned video, in which she shares random facts about herself based on the top questions raised by followers through written comments. The questions range from Emy's personal information such as age and relationship to educational background and career path. She also gives advice on learning English and studying abroad preparations. This interaction between Emy and her followers insinuates a close relationship in which Emy is an intimate friend who shares her private life and a mentor who offers suggestions on personal advancement. In addition, Emy answers most questions in English, which speaks for her competence in using English. Nevertheless, this video is not primarily designed to be pedagogical.

Another four videos (Video 2, 3, 5, and 6) are not instructional either, although they all have something to do with self-studying. In Video 2, Emy shares how she uses iPad for self-improvement, including reading e-books, practicing handwriting, and increasing productivity. Similarly, Emy in Video 3 introduces how she utilizes technological features on iPad to help her learn English. Video 5 features Emy advertises a Chinese book of life lessons listed on her *Douyin* store. Another merchant Emy was selling in her store was her pronunciation lessons, which she advertises in Video 6 along with her tip on how to learn English effectively.

The only instructional video is Video 4, during which Emy recommends several British TV shows for learning English. In addition to that, she also selects a few expressions from one of the shows and reads them aloud. While she is reading aloud, both English and Chinese captions are provided with English words visually highlighted through color, framing this video as pedagogical.

Unlike an English language learner Emy self-identifies as on her profile page, she is projected more as an English language expert and teacher in her videos. Meanwhile, her learner identity is also indicated through the videos as a means to construct her as a humble and motivated individual who is enthusiastic about self-improvement. Emy hybrid identity of amateur-expert appears to be consistently indexed in the content of her videos.

Table 22. Information of Emy's six selected videos

	Topic	Language skills	Teaching styles/methods
Video 1 (pinned)	Random facts about Emy	N/A	N/A
Video 2	The ways of using iPad for self-improvement	N/A	N/A
Video 3	The ways of using iPad for English learning	N/A	N/A
Video 4	Recommending British TV shows for learning English	Speaking	Lecture
Video 5	Advertising a book of life lessons	N/A	N/A
Video 6	Ways to learn English effectively and advertising her pronunciation lessons	N/A	N/A

Table 23 shows information of the 6 selected videos from the account @Chris 带你猛撕 英语 (Ripping through English Classes with Chris). The account owner self-identifies as a humorous person, a body builder, and an English teacher on his profile page.

The first three videos from Chris are pinned and non-instructional. Video 1 presents a self-introduction in which Chris self-identifies as an English language teacher. In addition, he claims that he speaks six languages and has traveled to 148 countries. In video 2, Chris narrates a story of him and his friend in the U.S. encountering someone on the street who tried to rob them. The robber ended up running away after Chris's friend faked a Russian accent. Additionally, Chris comments that Russians are the most fierce and tough people in the world, and therefore are the ones Americans are scared about the most. He also makes a reference to another stereotype that Russians keep bears as pets. In video 3, Chris shares another story of him getting pulled over in Dubai for speeding and finding an excuse to get out a ticket. Although part of the narration is done in English, the primary focus of the video is storytelling and therefore was not coded instructional.

The following three videos are unpinned and are pedagogically oriented. In Video 4, Chris performs how African Americans, British people, and Russians react to jokes differently, through which he informs the learners of some pragmatic differences. Likewise, Video 5 features Chris performing different reactions from White American parents, African American parents, and Russian parents to their children's school grades. These three different discourse identities in the role play are also realized through Chris' adoption of Mainstream American English, African American Vernacular, and a Russian accent respectively. Finally, Video 6 is a vocabulary lesson that introduces the distinct meaning of three English words in the context of the U.K.. To highlight the significance of knowing the meaning of the three words when traveling to the U.K.,

Chris also adds a few role-play episodes through which he showcases the possible consequences brought by the misunderstanding of those key vocabulary words.

One interesting pattern across Chris's videos is his frequent references to various kinds of national, cultural, and racial stereotypes that are widely circulating in the Chinese context in order to create a sense of playfulness and attract viewers' attention. These stereotypes are indexed in Chris exaggerated and parodic performances of various accents. Chris use of all these accents also suggest that he is very linguistically skilled. Moreover, to enhance the credibility of his depictions of those stereotypical images, Chris indicates that he is well-traveled person speaks from his personal experiences.

Table 23. Information of Chris's six selected videos

	Topic	Language skills	Teaching styles/methods
Video 1 (pinned)	Self-introduction	N/A	N/A
Video 2 (pinned)	Fake a Russian accent to scare Americans when in danger	N/A	N/A
Video 3 (pinned)	The story of Chris was pulled over by police in Dubai for speeding	N/A	N/A
Video 4	How do African Americans, British, and Russians react to jokes differently?	Speaking (Pragmatics)	Role play
Video 5	Different reactions from White American parents, African American parents, and Russian parents to their children's school grades	Speaking (Pragmatics)	Role play
Video 6	Three British words that might confuse people who didn't grow up in the U.K.	Vocabulary	Lecture and role play

Table 24 summarizes information of the 6 selected videos from the account@ Peter 教口语's (Peter Teaching Spoken English). The account owner focuses on presenting himself as an English language teacher and expert on his profile page.

Like many other content creators in my dataset, Peter has three pinned videos. Among these three videos, Video 1 is designed as an advertisement to encourage viewers to attend Peter's demo classes. To promote his lessons, Peter asserts that fully acquiring the English language requires learners to be able to automatize what they learned and think in English. Then he states that this goal cannot be achieved without systematic training, through which he attempts to convince the followers of the need to attend his classes. At the end of the video, Peter explicitly advertises his lessons by claiming that he uses the simplest and straightforward way to instruct learners to speak fluent English. The next pinned video, Video 2, is pedagogical and centers on pronunciation. It features Peter giving suggestions on how to improve pronunciation, which includes prioritizing accuracy over fluency and speaking slowly and clearly. Another pinned video, Video 3, is not instructional yet still relevant to English. In this video, Peter is playing guitar and singing the English song *We Don't Talk Anymore*, through which he projects an image of a proficient English user and a versatile individual.

Video 4-6 are unpinned and none of them are instructional. Similar to Video 3, Video 4 and 5 also feature Peter playing guitar and singing popular English songs. However, in each of these two videos, he is singing with another content creator of English language teaching videos together. Inviting other influencers of the same kind to be featured in one's videos is a common practice that many *Douyin* microcelebrities engage in to gain followers. The interaction between content creators and the co-creation of videos helps them grow their fanbase through receiving and offering the same audience to fellow influencers. Furthermore, the connections and

148

collaborations between content creators exhibited in videos can also serve to endorse each other's expertise and popularity. The last video of Peter (Video 6) in my dataset has Peter doing a Lord Voldemort impression. Specifically, Peter imitates the way of the character Lord Voldemort from the *Harry Potter* movies announcing that Harry Potter is dead and casting the killing curse *Avada Kedavra*.

Table 24. Information of Peter's six selected videos

	Topic	Language skills	Teaching styles/methods
Video 1 (pinned)	Promoting his English demo classes	N/A	N/A
Video 2 (pinned)	Prioritizing pronunciation accuracy over fluency and intonation	Speaking (pronunciation)	Lecture
Video 3 (pinned)	Peter singing the English song *We Don't Talk Anymore* and playing guitar	N/A	N/A
Video 4	Peter playing guitar and singing the English song *Apologize* with another content creator of English language teaching videos	N/A	N/A
Video 5	Peter playing guitar and singing the English song *Counting Stars* with another content creator of English language teaching videos	N/A	N/A
Video 6	Peter doing a Lord Voldemort impression	N/A	N/A

Same as Yang, the most followed microcelebrity in my dataset, there is only one instructional video among the selected six from Peter. The remaining videos often feature the content creator exhibiting his English expertise and entertaining the viewers through singing English songs, playing guitar, as well as imitating famous English movie characters. Another

interesting observation is Peter's interactions with fellow influencers in his video, which shows that self-branding on *Douyin* is a collaborative effort that involves socialization and engagement with not just viewers but also other content creators.

Table 25 lists information of the 6 selected videos from the account @嘉德仓自学 English 早 7:50 播 (Jiadecang Self-learning English, Stream Starts at 7:50 a.m.). The account owner constructs a public image as a Tibetan, a masculine man, and an outdoor, sociable person on his profile page.

Video 1 from Jiadecang's six selected videos is pinned and non-instructional. In the video, Jiadecang appears to be in a gym lip-syncing a Chinese song that was trending on *Douyin* during the time of data collection. The lyrics of the song describe a man missing his close friends who he addresses as "brothers". Thus, the song is used by the content creator to establish connections with his viewers.

The following five videos as shown in Table 25 are unpinned, and four of them (Video 2, 3, 4, 6) are pedagogical oriented. All these four videos are framed as English learning vlogs, which indexes Jiadecang's learner role. The learning vlogs feature him practicing spoken English. The ways he practices speaking include reading aloud English expressions, describing what he is doing, and having an English conversation with himself. However, he is more than just a learner. As will be detailed in Case Study 4, Chapter 6, Jiedecang deploys various linguistic and other semiotic resources to frame those videos as instructional. Another interesting aspect about his English learning vlogs is that Jiadecang frequently integrates multimodal resources (e.g., Tibetan style architecture, Tibetan clothing, Tibetan food) into those videos to construct his ethnic identity. Finally, Video 5 is about the content creator sending Tibetan New Year (Losar) wishes to the audience in Mandarin Chinese, and therefore is non-instructional.

150

However, consistent with his English learning vlogs, Jiadecang's Tibetan identity is indicated in this video as well.

Table 25. Information of Jiadecang's six selected videos

	Topic	Language skills	Teaching styles/methods
Video 1 (pinned)	Lip-syncing a song in a gym	N/A	N/A
Video 2	English learning vlog	Speaking	Reading aloud
Video 3	English learning vlog	Speaking	Role play
Video 4	English learning vlog	Speaking	Role play
Video 5	Sending Tibetan New Year greetings	N/A	N/A
Video 6	English learning vlog	Speaking	Role play

Table 26 summarizes information of the 6 selected videos from the account @一森教英语 (Yiseng Teaching English). The account owner projects an image of a social elite and a dedicated mother.

Among the six selected videos from this account, Video 1 and Video 2 are pinned, while the other four are unpinned. Video 1 is an instructional video that teaches English vocabulary, specifically the vocabulary differences between American English and British English. One unique feature about this video is that Dabai, the second most followed influencer in my dataset, is co-teaching vocabulary with Yiseng, a young boy of half British and half Chinese. The instruction is delivered through Dabai and Yiseng taking turns to read aloud the different vocabulary words in American English and British English that are used to refer to the same

things. Dabai as Canadian reads the American words and Yiseng as half British reads the British expressions. By doing so, the English variety they speak communicates their cultural and ethnic background. As discussed earlier, again the content co-production points to the need for socialization and collaboration with fellow influencers as a way to share a fanbase and gain approval for a self-branding purpose.

Video 2 is a Chinese reality show clip that features Yiseng and his mom singing the English song *Tough Love* from the Disney musical *The Villains Lair*. The main purpose of this pinned video is to demonstrate the English proficiency of both Yiseng and Yiseng's mother, as well as to project Ethan's mother as a model parent who raises her son to be successful following the value of Chinese middle-class parents.

Video 3, 4, and 6 are all pedagogically oriented and have a focus on teaching vocabulary. In addition to reading the English expressions aloud in those videos, Yiseng also repeatedly "complains" that English is more complicated than Chinese in its pronunciation for having multiple syllables in one word. As a mixed-race boy of half British and half Chinese, his Caucasian look as part of his transportable identity serves to associate him with a native speaker of English, which establishes his expertise and authority in using English. However, the native speaker's image indexed by his race also leads to a disconnection with his Chinese identity. In order to establish affiliation and membership with *Douyin* audiences, Yiseng distances himself from his British identity by using "foreigners" to refer to people who speak English as their first language and explicitly expresses his preference for Chinese over English.

Video 5 is a non-instructional vlog that shows Yiseng cleaning and organizing his desk and bookcase. There is no verbal communication involved, and the main purpose of the video is

to present Yiseng as a well-behaved child, which then suggests Ethan's mother's success and dedication in parenting.

Table 26. Information of the six selected videos of Yiseng's mother

	Topic	Language skills	Teaching styles/methods
Video 1 (pinned)	Vocabulary differences between American English and British English	Vocabulary	Reading aloud
Video 2 (pinned)	Ethan and his mom performing on a reality show, singing the English song *Tough Love* from the Disney musical *The Villains Lair*	N/A	N/A
Video 3	English vocabulary words for adult animals and their young	Vocabulary	Reading aloud
Video 4	English vocabulary of numbers from 1-12	Vocabulary	Reading aloud
Video 5	Ethan cleaning and organize his desk and bookcase	N/A	N/A
Video 6	Basic math vocabulary in English	Vocabulary	Reading aloud

Table 27 summarizes information of the 6 selected videos from the account @英国小克里 (British XiaoKeli). The account owner constructs a public image of a bilingual expert in both English and Chinese, a social elite, a professional translator, and someone who loves China.

Video 1 of Xiaokeli is a pinned instructional video that discusses the appropriate ways to check someone's understanding and therefore has a focus on speaking, or more specifically, pragmatic skills. The content creator in the video first explains why the two questions "Do you understand?" and "Do you know my meaning?", which many Chinese learners often use, are

problematic. Then he provides several expressions that can be used to check one's understanding politely.

Video 2 is also a pinned non-instructional video that features Xiaokeli's self-introduction in five languages, Mandarin Chinese, English, Spanish, French, and Cantonese, through which he constructs a multilingual speaker image and a linguistically skilled person. The content of his self-introduction includes where he was born (the U.S.), the place where he grew up (the U.K.), how he acquired Spanish, French, and Cantonese in addition to English and Mandarin Chinese, as well as his transnational marriage. To conclude his self-introduction and encourage *Douyin* viewers to learn English, he states that learning Chinese is life-changing for him, and he hopes audiences will find learning English life-changing for themselves too. This video is primarily designed to promote Xiaokeli as an English language authority, a native speaker, a multilingual, and someone who is closely affiliated with Chinese people and culture in order to attract followers.

Video 3, 4, and 5 are not pedagogically focused, as they are all vlogs featuring Xiaokeli sharing his life in Mandarin Chinese. Video 3 shows Xiaokeli walking on a street in the U.K. and talking about the COVID-19 vaccination progress in the country. Similarly, Video 4 features Xiaokeli walking outside and giving followers an update on his daughter's progress in learning English and Chinese as an early bilingual. Likewise, Video 5 is about the content creator going to a gardening store to purchase flower seeds. He concludes the video by wishing his daughter grow up healthy little by little like plants. Through posting these vlogs, Xiaokeli shares his recent life with audiences, which serves to convey a sense of intimacy. His exclusive use of Mandarin Chinese also helps him to address *Douyin* viewers and enhance his affiliation with them.

The last video, Video 6, is a video of English instruction that emphasizes listening skills. The video features Xiaokeli accompanying his daughter to learn English writing. The conversation between the two people is in English, and both English and Chinese captions are added to the videos close to the center of the screen for an increased information *value* (Kress & van Leeuwen, 2006). Both Xiaokeli and his daughter appear to back face the camera, which creates an objective and detached view for the audience as a listener of their conversation. As those semiotic resources integrated in the video can direct viewers' attention to hearing what they say, I, therefore, consider the content instructional with a focus on listening skills.

Table 27. Information of Xiaokeli's six selected videos

	Topic	Language skills	Teaching styles/methods
Video 1 (pinned)	Expressions used to check if someone understands what you said in English without sounding condescending or rude	Speaking (pragmatics)	Lecture
Video 2 (pinned)	Self-introduction	N/A	N/A
Video 3	The COVID-19 vaccination progress in the U.K.	N/A	N/A
Video 4	Update on Chris's daughter's language acquisition progress as a bilingual	N/A	N/A
Video 5	Going to a gardening store	N/A	N/A
Video 6	Chris talking to his daughter in English	Listening	Real-life conversation

In addition to constructing himself as an English language expert and a friend of Chinese people, another important identity aspect Xiaokeli establishes is his role as a loving father. My

155

long-term observation finds that many of his videos include Xiaokeli mentioning his daughter's growth and development, as well as him accompanying his daughter to learn and play. The image of a caring, dedicated parent Xiaokeli projects also helps him win viewers' approval.

Table 28 presents information of the 6 selected videos from the account @ 🇺🇸Allie 爱丽 💕 (the US flag emoji, Allie Aili, two hearts emoji). The account owner constructs an image of an American teenager who is a native speaker of English and a multilingual.

Most of Allie's videos are teaching-oriented. There are three pinned videos (Video 1, 2, and 3), and all of them are of language instruction. Video 1 shows Allie sharing the names of 18 punctuation marks in English. She begins by asking the question in Mandarin Chinese "Do you know the names of the following 18 punctation marks in English?" to introduce the topic of the video. After that, she lists the names of the 18 punctuation marks. As she reads, the corresponding punctuation marks, along with their English and Chinese names are superimposed on the screen. Finally, she ends the video by checking with the users "Got it?", again, in Mandarin Chinese.

Video 2 features Allie in an American-style home kitchen introducing the English names of different household items in the kitchen. In the video, she walks around the kitchen, pointing at each item she sees and calling the name. She also explains what an item is used for if it is uncommon to be found in a Chinese kitchen, such as French press and food processor. Her instruction is mainly in English, with both English and Chinese captions. However, Allie also speaks Chinese from time to time, especially when mentioning common Chinese kitchen utensils and essential Chinese seasonings.

Video 3 is a recording of a real-life conversation between Allie and a flight attendant, which instructs viewers on how to order drinks on a plane. The English conversation is not only

accompanied by English and Chinese subtitles but also some alternative expressions in written form. For example, when Allie asks for water, the flight attendant responds "Still or sparkling?". Her utterance is captioned in both English and Chinese. Moreover, under the English subtitle, the alternative expression "Would you like still or sparkling water?" is added.

The next three videos (Video 4, 5, and 6) are unpinned. Video 4 is the only non-instructional video among the selected six. This video was originally posted during Chinese New Year 2021, and features Allie sending Chinese New Year wishes to *Douyin* users in both Mandarin Chinese and English in order to establish connections and rapport with the audiences. She also encourages viewers to keep learning English in the new year by telling them that they will be able to "communicate with over 2 billion people from around 70 countries" if they speak both Chinese and English. As the primary focus of this video is to send the content creator's new year wishes and motivate learning, it is classified as non-pedagogical.

Video 5 is an English teaching video that Allie creates to discuss the expressions for common units in the U.S., such as *pound*, *gallon*, *miles per hour*, and *degree Fahrenheit*, as well as how to put them in sentences. Video 6 discusses the names for common units as well, but with a focus on the metric system, which is used in China. In the video, Allies shares the name of units such as *meter*, *gram*, *second*, *liter*, and *degree Celsius*. She also provides sentences to exemplify how these units are used in sentences. Although both Video 5 and 6 have a speaking component, the various words introduced in the videos make them primarily vocabulary focused.

As the analysis shows, many of Allie's videos address English vocabulary. Unlike many other content creators who focus on only a few words in one video when teaching vocabulary, Allie tends to share as many words as possible at a time, and the words she teaches in each video is organized based on a specific topic (e.g., punctuation marks, measurement units, household

157

items in kitchen). By doing so, she manages to present information in a systematic way. In addition, the lessons she creates may impress the viewers with the large amount of information included, through which Allie establishes her English language expertise to attract following. On the other hand, however, it is questionable if those videos of vocabulary are well received by *Douyin* audiences as effective teaching/learning materials because there is a lot of information, and the videos are very fast paced.

Table 28. Information of Allie's six selected videos

	Topic	Language skills	Teaching styles/methods
Video 1 (pinned)	Names of various punctuation marks in English	Vocabulary	Lecture
Video 2 (pinned)	Names of different household items in kitchen in English	Vocabulary	Lecture
Video 3 (pinned)	How to order drinks on an airplane in English	Speaking	Real-life conversation
Video 4	Sending Chinese New Year greetings and encouraging learners to keep learning English	N/A	N/A
Video 5	English expressions for common units in the U.S.	Vocabulary	Lecture
Video 6	How to talk about common measurement units English.	Vocabulary	Lecture

Cross-case Analysis of RQ2

As shown in Table 29, among the 84 selected videos, videos with or without a focus on English language teaching are quite balanced in number. Among the 43 English instruction videos, the number of pinned videos is 13, which is much smaller that of unpinned ones (30). Although the remaining 41 videos do not center on English teaching, many of them are still

relevant to the English language, as shown in the previous tables, in which the topics of those videos are summarized. The English language related topics of these non-instructional videos include singing English songs, voicing characters in English movies, doing accent impressions, and so forth. Although these videos are not designed with a clear teaching purpose, they are used demonstrate the content creators' English language proficiency and thus construct an English expert identity. Other topics of non-instructional videos include self-introduction, demonstrating multilingual competence in languages other than English and Chinese, promoting language related products, and addressing popular social events and issues. These topics are often adopted by content creators to project themselves as someone who is knowledgeable, highly educated, and talented in language learning, as well as help them promote their products and services, and establish affiliations with the viewer.

As previously mentioned, each *Douyin* account is allowed to pin up to three videos. As microcelebrities of English instruction, however, they in general choose to pin more videos that do not have a direct focus on English language teaching. As shown in Table 29, there are 22 pinned videos that do not have an explicit pedagogical orientation, while the number of pinned ones with an instructional purpose is only 13. Since pinned videos are limited in number and are displayed on the top of the video list on the vlogger's profile page, it is reasonable to assume that pinned videos contain the content *Douyin* microcelebrities want the viewer to see the most when they visit their profile page. Most of the pinned videos feature the content creator demonstrating their language competence in various ways (e.g., singing, performing accents, voicing movie characters) without teaching anything, which might suggest the trend of prioritizing language competence over pedagogical knowledge in the construction of an English language expert identity on *Douyin*. In the attention economy, this trend is likely to be motivated by Chinese

159

netizens' perception of a legitimate English teacher. In addition, these pinned videos can be used as an entertainment appeal to catch viewers to attention.

Table 29. English teaching videos vs. other videos among the 84 selected videos

	English teaching videos			**Other videos**		
Count	43			41		
% of total	51%			49%		
		Pinned videos	**Videos not pinned**		**Pinned videos**	**Videos not pinned**
	Count	13	30	Count	22	20
	% of 42 English teaching videos	30%	70%	% of 42 English teaching videos	52%	48%
	% of total	15%	35%	% of total	26%	24%

Table 30. Types of language skills taught in the 43 English teaching videos

Types of language skills taught in the videos	Count
Speaking	25
Vocabulary	13
Grammar	3
Listening	2
Sociolinguistics	1

Table 30 provides an overview of the language skills taught in the 43 instructional videos. As the table shows, speaking is the language skill that is most frequently taught and

discussed in the 43 English teaching videos, which might indicate the popularity of pedagogical videos with a focus on speaking on *Douyin*. Three main reasons may contribute to this phenomenon. First, many Chinese learners of English lack opportunities and resources to improve spoken English in real life. Traditionally, formal English education in China has had its primary focus on tests of reading and writing, while speaking and listening skills having usually not received as much attention. Consequently, many Chinese learners of English need to improve their spoken English for various reasons including job requirements, promotions, travel, studying abroad, and so forth. Thus, creating and sharing videos that focus on speaking skills on *Douyin* meet leaners' needs and are therefore more likely to attract followers.

Second, a focus on speaking skills may create more chances for the content creators to interact with their viewers. For example, vloggers can share practical expressions that viewers can relate to and use immediately in their own life, such as internet slang (e.g., "spill the tea"), travel English, business English, and English for interviews. With this focus, leaners may be motivated to practice, think about the language usage, and ask questions in comments. I have also observed that many English vloggers encourage viewers to do an asynchronous duet with them to practice dialogues (e.g. Vlogger: "practice your English with me. I'll say blue (text), you'll say yellow (text)") and to post their videos. By doing so, content creators promote interactions and increase their publicity. Another reason that these Douyiners produce more spoken English topics might be the affordances and constraints involved in delivering language lesson using videos on *Douyin* platform. Since *Douyin* is a short video sharing platform, it is suitable for content creators to share a tiny piece of knowledge in informal, casual speech that is both easy to process and attention grabbing in each short video. For example, one vlogger teaches the viewer to use "gimme" to replace "give me" to sound more fluent and natural, and

161

uses the expression "gimme five" as a catchy phrase to make the instruction more fun and memorable.

Perhaps it is with a similar reason that vocabulary is the second most popular topic among the videos in my dataset. Vocabulary words in general do not require much time to explain and therefore can fit well in a short video. In addition, vocabulary words, especially speaking vocabulary, goes well with instructional videos that focus on speaking. Among the vocabulary words included in the videos in my dataset, I observed that daily used English words is the most popular category. Many words discussed in the videos might be among those that learners already known for their dictionary meaning or their most common meaning. However, learners might not be aware of the connotations of these words, how their meaning changes depending on the context, or how to use them correctly in a sentence. For example, many Chinese learners use the word "lover" to refer to their spouse as a result of direct translation from the Chinese expression "爱人" (love person). However, "lover" in English is often used to refer to a partner in a sexual relationship outside marriage. Teaching learners to use the words they have some prior knowledge about helps learners comprehend the content with ease. Meanwhile, it is also attention-grabbing to address common mistakes and misperceptions in learners have.

On the other hand, skills such as grammar, reading, and writing are much less frequently addressed in those *Douyin* informal instruction videos in my dataset. For grammar, the reason for its low presence might be attributed to its difficulty. Viewers might find content about grammar requires effort to comprehend. It is also challenging for content creators to explain grammar rules concisely in short videos. However, as my Case Study 3 in Chapter 6 will show, one of the microcelebrities in my dataset was able to successfully present a simple and straightforward grammar lesson that has gained over 1 million likes. Nevertheless, not many content creators of

informal English lessons may be capable of doing so. As for reading and writing skills, modal affordances of *Douyin* might be the main reason for their unpopularity. Unlike Kindle which is designed as a reading device, or Google docs used mainly for writing and editing, *Douyin* as a short video platform is less in its capacity to present written passages as learning materials for reading and writing lessons. In addition, videos focusing on reading and writing might have requirements for the viewers' learning environment, and the videos are more likely longer, and filled with texts which might overwhelm the users who have short attention spans.

Although listening skills are as underrepresented as grammar, reading, and writing, they are in fact addressed to a certain extent in videos that primarily focus on speaking. Table 31 presents the number of videos centering on different aspects of English-speaking skill taught in the 25 videos of spoken English. There are 10 videos that address speaking skills in general, 8 videos that focus on pronunciation, and another 7 videos that highlight pragmatics. Many of these videos can also be used as listening materials as the content often involves conversation dialogues. Similarly, videos with a focus on pronunciation, which is often taught through the method of reading aloud, can also be used to help learners improve their aural word recognition. Nevertheless, listening is not the skill that is emphasized or explicitly addressed the most. This is perhaps also due to the need for a quiet environment to allow the learners to focus on the audio, especially when they are listening to a conversation instead of individual words or individual sentences.

Table 31. The specific aspects of English-speaking skill taught in the 25 videos of spoken English

The specific aspects of speaking skill addressed in the videos	Count
General	10
Pronunciation	8
Pragmatics	7

Table 32. Types of teaching styles and methods in the 43 videos with an English language teaching focus

Types of teaching styles/methods in the videos	Count
Lecture	22
Reading aloud	13
Role play	9
Real-life conversation	2

As for the types of teaching styles and methods used in the 43 videos with an English language teaching purpose, half of them feature a lecture style of teaching (see Table 32), which is characterized by the content producer giving language instruction. The second most used teaching style is reading aloud. Different from the lecture genre, content creators who adopt the method of reading aloud to teach do not provide any metalinguistic explanations in the video. Instead, they simply read words and expressions aloud, usually at a slow pace. This teaching style is often applied to teach vocabulary and pronunciation. In addition, it is shown that role plays are used to teach speaking, and occasionally also vocabulary in order to provide a context for interpretation. Similar to role plays, real-life conversations are integrated into two videos to provide a context for language use. Meanwhile, the real-life conversations are considered more

authentic materials for language learning. However, they are limited in number probably due to the difficulty to record a real-life English conversation in China as predominantly a monolingual context. The two real-life conversations in my dataset are both recorded in an English-speaking context outside of China.

CHAPTER SIX: RQ3 FINDINGS

This chapter is dedicated to answering the third research question:

RQ 3: What multimodal resources are employed by content creators in their videos?

To address this research question, I ask the following sub-questions:

a. How do they use multimodality to construct English expert identities?

b. How do they use multimodal resources to communicate other personae that they present on the platform?

c. How do they exploit multimodal resources to teach English?

d. What multimodal strategies do they use to engage in self-branding and advertising?

This chapter presents a detailed qualitative analysis of four videos selected from the 84 videos sampled for this dissertation project. These four videos are created by four different microcelebrities of informal English instruction among the 14 content creators included in the dataset. As will be explained in my analysis of this chapter, the reasons for the selections of these videos for a scrutiny include the popularity of the video and the content creator, the variety and uniqueness of the multimodal strategies incorporated in the videos, and the power negotiation in knowledge production indexed in the video content. To transcribe these videos, I adopt a multimodal transcription method adapted informed from both Matwick and Matwick (2017), which attends to the interplay to various semiotic resources, as well as Ho (2019), which emphasizes the structure of the pedagogical videos. Table 33 is a transcription excerpt that exemplifies how one of the four selected videos, which is Yang's video of airport English, is transcribed. The full transcriptions of the four selected videos are enclosed as Appendices.

To examine how the content creators incorporate multimodal resources in their videos in order to display a public image appealing to the audience and deliver English lessons, I take a social semiotic approach to multimodal discourse (Kress, 2010; Kress & van Leeuwen, 2001, 2006) to analyze content creators' selection of meaning-making resources that are situated in their socio-cultural context and available for use due to the technological affordances. The diverse identity types microcelebrities display are discussed using a framework adapted from both Zimmerman (2014) and Tracy (2002) and are categorized into transportation identities, discourse identities, situated identities, and personal identities. Furthermore, to study the characteristics of identity construction in *Douyin* short videos, I follow Buchotlz and Hall's (2010) five principles, which are the emergence principle, the positionality principle, the indexicality principle, the relationality principle, and the partialness principle, to examine how identities are evoked, indexed, and shifted as the content creators draw on multimodal resources to interact and communicate with viewers. I present my analysis of the videos as four different case studies as follows.

Table 33. A multimodal view of Yang's video of airport English: 0:07-0:39
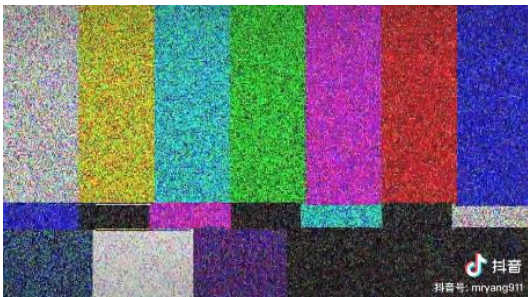
Role play and Explanation (0:07-0:39): How to answer the airport customs question "What is the purpose of your stay?"

Screen capture
0:07

Table 33. (Continued)

On-screen verbal attributes	Off-screen verbal attributes
None	None

Visual attributes	Sound
A TV error effect	A TV static sound effect

Explanatory Notes

A TV error effect and a TV static sound effect are for transition.

Screen capture	

0:08

On-screen verbal attributes	Off-screen verbal attributes
English captions: *Top 1* *What is the purpose of your stay?* **Translated caption in Chinese:** *你逗留的目的是什么*	**Narration by Yang** (in Chinese)**:** *Top 1. What is the purpose of your stay?*

Visual attributes	Sound
Position of captions: Center; English caption above the Chinese caption **Color:** English caption in white; Chinese caption in yellow; black background **Font size:** Both English and Chinese are about the same size	None

Explanatory Notes

The airport customs question is highlighted using **color** (the white and yellow texts contrast sharply with the black background) and the **position of captions** (center).

Table 33. (Continued)

Screen capture	
0:09	0:15
0:20	0:23

On-screen verbal attributes	Off-screen verbal attributes
English captions: (Immigration Officer): Next. What is the purpose of your stay? (Tourist A with a thick Cantonese accent): For school? To see my relative? For work and to find a wif…WiFi!! **Translated caption in Chinese:** (Immigration Officer):*你逗留的目的是什么?* (Tourist A): 去上学? 探亲? 找工作 再找个老… **Chinese captions:** (Tourist A): 怎么回事啊？去哪里啊？ **Translated caption in English:** (Tourist A): What happened? Where am I going?	**Role play conversation**: (Immigration Officer): Next. What is the purpose of your stay? (Tourist A): For school? To see my relative? For work and to find a wif…WiFi!! (Tourist A yells): wei wei wei ('wei' in Chinese is an interjection used to express surprise) (Tourist A in Cantonese): 怎么回事啊？去哪里啊？(Translation: What's going on? Where are you taking me?)

169

Table 33. (Continued)

Visual attributes	Sound
Mid-shot **Background:** Airport customs **Middle ground:** Yang wearing glasses, in a blue plaid shirt, with a backpack (as the role of tourist A); Yang in a blonde wig standing behind a computer monitor at a customs counter; in a professional uniform (playing the role of an immigration officer) **Position of captions:** Close to the center of the screen. **Color:** The captions are highlighted in yellow. **Font size:** The English subtitles are larger in font size than the Chinese subtitles below	Light clarinet music with a moderately quick tempo, playing in the background.

Explanatory notes

In role play when both parties communicate in English, the English and Chinese subtitles are placed right next to the person who is speaking (close to the center of the screen). The English captions are placed above the Chinese captions. The order of the two subtitles is reversed when the speech is in Chinese/Cantonese.

Screen capture
0:25

On-screen verbal attributes	Off-screen verbal attributes
None	None
Visual attributes	**Sound**
TV error effect	A censor beep sound

Explanatory notes

A TV error visual effect and a censor beep sound effect are for transition.

Table 33. (Continued)

Screen capture

0:25

On-screen verbal attributes	Off-screen verbal attributes
Chinese captions: *这些情况都需要特定的签证, 所以如果你只是去玩的话就不要乱说啦!*	**Narration by Yang** (in Chinese)**:** *这些情况都需要特定的签证, 所以如果你只是去玩的话就不要乱说啦*
Translated captions in English: *These answers above need a special visa. So if you're just there for fun? Then play it safe!*	(Translation: You need to present specific types of visas to make these answers sound legit. If you are just visiting a foreign country for leisure, don't use those answers.)

Visual attributes	Sound
Mid-shot **Background:** Room decor, white board **Middle ground**: Yang talking, direct eye contact with viewers. **Position of captions:** Close to the bottom of the screen; English translation placed under the Chinese caption. **Font size:** Chinese caption in larger font, English translation in smaller font.	Light, clarinet music with a moderately quick tempo, playing in the background.

Explanatory Notes

Explanation is done after a brief role play section using spoken language.

Table 33. (Continued)

Screen capture

0:30

On-screen verbal attributes	**Off-screen verbal attributes**
Chinese captions: 正确回应方式	**Narration by Yang:** 正确回应方式
Translated captions in English: *The correct way*	(Translation: The correct response)

Visual attributes	**Sound**
Position of captions: Center; Chinese caption above English caption	Light, clarinet music with a moderately quick tempo, playing in the background.
Color: Both Chinese and English captions in yellow with a red shadow; black background	
Font size: Chinese caption in larger font, English caption in smaller font	

Explanatory Notes

The scene informs the viewers that the next part is about the correct ways to answer the discussed airport customs question, making a transition from the explanation to the next role play episode. The **colors** (the yellow texts with red shadow contrast sharply with the black background) and the **position of captions** (center) are used to emphasize the textual information.

Table 33. (Continued)

Screen capture

0:32	0:33
0:34	0:38

On-screen verbal attributes	**Off-screen verbal attributes**

English captions:
 (Immigration officer): *What is the purpose of your stay?*
 (Tourist B): *I'm here for fun*
        I'm here to shop
 (Two immigration officers): *Welcome*
 Translated caption in Chinese:
 (Immigration officer):*你逗留的目的是什么?*
 (Tourist B): *我是来玩的*
       我是来购物的
 (Two immigration officers): *欢迎欢迎*
	Role play conversation:
	 (Immigration officer): *What is the purpose of your stay?*
	 (Tourist B): *I'm here for fun*
	        I'm here to shop
	 (Yang as the immigration officer): *Ah!*
	 (Two immigration officers together):
	 Welcome! Welcome! Welcome! Welcome!

173

Table 33. (Continued)

Visual attributes	Sound
Mid-shot **Background:** Airport customs **Middle ground:** Yang in red plaid shirt, with a backpack (as the role of tourist B); Yang and another person in a blonde wig and professional uniform (playing the role of immigration officers) **Gesture:** Yang as an immigration officer rubbing two fingers and his thumb together while welcoming tourist B. **Position of captions:** Close to the speaker **Color:** The captions are highlighted in yellow. **Font size:** The English subtitles are larger in font size than the Chinese subtitles below.	Light, clarinet music with a moderately quick tempo, playing in the background until Yang as tourist B finishes saying "I'm here to shop." Then, the Spring Festival Overture, a tune widely used on various festive occasions in China, is playing when the two immigration officers welcome tourist B.

Explanatory Notes
Through the reenactment of the previous role play scene in which the immigration officer asks the question as well as through Yang's performance as Tourist B, this role play episode serves to contrast the language performance of Tourist A and Tourist B. Rubbing two fingers and the thumb together signifies money.

Case Study 1: Yang's Video of Airport English

The first video selected for a multimodal analysis is from one of six videos collected

from the account @MrYang 杨家成 (Mr. Yang Jiacheng Yang), including three pinned videos

and another three videos displayed right after the pinned ones. The selected video is one of

Yang's three pinned videos and has a clear focus on English language teaching. Pinned videos on

*Douyi*n are placed at the top of the video list on one's profile page. Each account is allowed to

pin up to three videos. The rest of the videos are displayed in reserve chronological order. In

order to seize as much attention within as little time as possible, many content creators carefully

choose which videos to pin and show to the viewers first. It is also one of Yang's most popular

videos with over two million likes accumulated by the time the data were collected (February 9th,

2021). The video is about three minutes in length, during which Yang discusses travel English,

specifically, the English that language learners need to know when going through customs at an

airport. Based on the multimodal transcriptions of the video as exemplified in Table 33, the

structure of the lesson is then summarized as shown in Table 34 to provide an overview of the

key "moves" or stages of the video.

Table 34. The structure of Yang's video of airport English

A. **Introduction** (0:00-0:07): frames the video
B. **Role play and Explanation** (0:07-0:39): how to answer the airport customs question "What is the purpose of your stay?"
C. **Role play and Explanation** (0:39-1:01): how to answer the airport customs question "How long do you intend to stay?"
D. **Role play and Explanation** (1:02-1:24): how to answer the airport customs question "Where will you be staying?"
E. **Role play and Explanation** (1:24-1:52): how to answer the airport customs question "What is your occupation?"
F. **Role play and Explanation** (1:53-2:42): how to answer the airport customs question "Do you have anything to declare?"
G. **Concluding remarks** (2:43-2:47)
H. **Appeal to following** (2:47-2:51)
I. **Post-credit scene** (2:52-3:00): what happens to the tourist in the role play after he was detained at an airport

Compared to most of the *Douyin* English instructional videos that are kept within one

minute to cater to the audience's short attention span, a three-minute video is clearly not typical

in terms of length to increase views. However, by strategically integrating a myriad of

175

multimodal resources, Yang manages to keep the content informative, engaging, and entertaining. Below, I provide a detailed analysis of how Yang exploits assorted semiotic resources to display a public self and engage in English teaching. I also address how the combination of semiotic modes generates meaning.

Speech

In Yang's video, speech (as an essential semiotic mode of meaning-making) is realized by several linguistic codes, including Mandarin Chinese, English, and Cantonese. Specifically, Yang uses Mandarin Chinese in the video to introduce the topic of discussion during the opening scenes, to provide explanations on the role play dialogues, to make concluding remarks and to appeal to his following, as well as to perform a parody as a post-credit scene at the end of the video (see Table 34). However, when it comes to role play conversations in which Yang performs a chain of dialogues between an immigration officer and a Chinese tourist, English is used as the main language of communication, with some translanguaging practices between English, Mandarin Chinese, and Cantonese. As the role play episodes are inserted at different points throughout the video, the difference in the adoption of linguistic codes for the role play episodes and the rest of the video serves to indicate the transition between the role play and other content. This alternation in the linguistic resources used for speech also helps direct viewers' attention to things the content creator intends to highlight. For the role play, apart from some translanguaging practices, the speech is primarily constructed in English, which helps frame the video as instructional in the Chinese online context where Mandarin Chinese is the dominant language. The English conversations also make the learning objective salient by drawing the audience's attention to the English language expressions used in the role play. When at other points of the video, Yang mainly uses Mandarin Chinese for his speech. As the official language

in China, Mandarin is spoken by the majority of the Chinese population. Using Mandarin allows Yang to make his content accessible to as many viewers as possible on *Douyin* and communicate ideas effectively.

In addition to using the choice of codes to structure the video and direct the viewers' attention, the way Yang constructs his speech with his multilingual repertoire is also shaped by his purpose of self-branding as I will illustrate below.

The identity of Yang as an English language teacher is constructed linguistically at the very beginning of the video. During the opening scenes, Yang introduces the topic of the video by asking the following question in Chinese:

"大家出国一落地一定会面临一个问题，那就是过关。那警官问你的都是英文，那怎么回答才好呢?"

Translation: "*Right after your flight arrives in a foreign country, there is one thing you must deal with, which is go through immigration. However, all questions the immigration officer will ask are in English. How would you properly answer those questions*?"

In this question, Yang enumerates a circumstance of "出国" (in a foreign country; go abroad), "过关" (go through immigration), and "都是英文" (all questions…are in English). Through this concrete description of an occasion where English is necessary for communication, Yang conveys a personalized message to the viewer by imaging a situation they are likely to be involved in and identifying their language learning objectives. In this sense, Yang's utterance constructs him as both an advertiser and an English language teacher who strives to address the consumer/learner's needs. Together with Yang's direct address of the viewer, including "大家" (the second person plural *you*), and "你" (the second person singular *you*), the contextualized

177

message Yang conveys enhances audiences' self-referencing to the described scenario, which can lead to increased motivation for consumption-related activities (Huang, et al, 2020). Moreover, to establish English as a necessary skill to learn, Yang constructs English as a global language. Specifically, in his description of the scenario during the opening scenes, Yang claims that "那警官问你的都是英文" (all the questions the immigration officer will ask are in English) and associates it with the circumstance of "出国" (being in a foreign country). As such, Yang communicates the assumption that English is used in all foreign countries and is an essential skill one must learn to travel abroad. To further relate the viewer to the scenario of going through immigration, Yang raises an immediate question after his description of the imagined context "How would you properly answer those questions?" Opening with a question is a common strategy to bring audiences' attention to a topic. To further engage the viewer and make them think, the direct address "you" is used in the question. Since the question is based on the airport scenario Yang describes, when it is combined with the classroom setting in which Yang is projected as an English language teacher, the question Yang raises further situates himself as a teacher who facilitates learning by inviting students' interactions and thinking on a hands-on, meaningful language task.

After the introduction, Yang discusses the top five questions that an immigration officer might ask at an airport and how to answer them. The five questions are presented one by one, and each is accompanied by role play episodes and tips for responding appropriately. Drawing on the discussion of the first question "What is the purpose of your stay?" in the video as an example (see Table 33), the discussion of this question is structured as 1) introducing the question, 2) a role play episode featuring Yang playing the role as a nervous Chinese tourist (referred to as Tourist A), who responds to the discussed immigration question inappropriately,

3) an explanation of why Tourist A's response is problematic, 4) a role play episode in which Yang plays the role of a confident Chinese tourist (referred to as Tourist B), who provides correct answers to the same question. The discussions of the remaining four questions all follow a similar structure. As mentioned earlier, during the role play, conversations are primarily in English with a focus on language teaching. However, translanguaging practices between English, Mandarin, and Cantonese in the role play also serve as a key feature that differentiate it from other content in the video. It is revealed in my analysis that Yang strategically uses translanguaging in the role play to perform multiple identities for a self-branding purpose.

   To understand how Yang draws on his multilingual repertoire to construct various identities in the role play, I first introduce the three roles Yang performs in the role play episodes and the linguistic codes Yang assigns to each role. These three roles are an immigration officer who speaks fluent American English, Tourist A who speaks English with a thick Cantonese accent and translanguages between English, Mandarin, and Cantonese, and Tourist B who speaks fluent American English (see Figure 17).

Figure 17. The three roles in the role play episodes
Note. The immigration officer (17/1), Tourist A (17/2), Tourist B (17/3)

   When Yang performs the immigration officer and Tourist B, his speech is characterized by common phonological and phonetic features of Mainstream American English, including rhoticity /ɹ/, using the low back vowel /ɑ/ instead of the mid back vowel /ɔ/ in words like *shop* (/ʃɑp/) and *brought* (/bɹɑt/), as well as replacing the aspirated /tʰ/ with the alveolar flap

/ɾ/ (a quick /d/ sound) in words like *bottle* (/bɑɾl̩/). Different from speaking fluent American

English as the immigration officer or Tourist B, Yang performs pidginized English with a heavy

Cantonese accent when playing the role of Tourist A, and also frequently switches between

English, Mandarin, and Cantonese. Drawing on the linguistic characteristics of 'Asian accents'

Chun (2004) identifies and my insider knowledge of Chinese speakers of English, a close

analysis of Tourist A's speech reveals that Yang adopts various features of Cantonese-accented

English to revoice a character of a Cantonese speaker of English. Phonologically, Tourist A's

speech is marked by its 1) prosodic aspect including final syllable lengthening, syllable-timed

rhythm, sudden rising and falling intonation, and emphatic lengthening (e.g., line 6 of Extract 1,

lines 2-4 of Extract 2, lines 3-4 of Extract 3), 2) vowel quality including monophthongization

(line 2 of Extract 2), 3) consonant substitutions of [v] with [w], [ɹ] with [l], [n] with [l], and [z]

with [s] (e.g., line 6, 8, 12 of Extract 5), 4) syllable structure including coda /ɹ/ deletion and

schwa epenthesis to closed syllable (e.g., line 4 of Extract 1, lines 8 and 12 of Extract 5).

Syntactically, features such as subject omission (line 5 of Extract 5) and sentence-final particle

(line 4 of Extract 2) are also integrated into Tourist A's speech.

Extract 1

1 Immigration Officer:	What is the purpose of your stay?
2 Tourist A:	for school
3	to see my relative
4	for work *(coda /ɹ/* deletion and *schwa epenthesis to closed syllable[43]: work[wɜɹk]→[wɜkə])*

---

[43] Chun (2004) identifies schwa epenthesis to closed syllable as a feature that Margaret Cho, a Korean American comedian, employed to perform a Japanese or Korean accent. However, this phonological feature can be found in Chinese (including Mandarin and Cantonese) accented English as well.

5	and to find a wif
6	wifi::::: *(final syllable lengthening: [waɪfaɪ] → [waɪfaɪ:::::])*
7	(immigration officer frowns with a down-turned smile; a popular *Douyin* sound effect commonly used to express surprise or shock cooccurs with the immigration officer's facial expressions)
8	(in Cantonese)喂喂喂! 怎么回事啊? 去哪里啊? *(hey hey hey! What is going on? Where are you taking me?)*

Extract 2

1 Immigration Officer:	How long do you intend to stay?
2 Tourist A:	I don't know:::: *(monophthongization and final syllable lengthening: I don't know [aɪ doʊnʔ noʊ] → [aɪ donʔ no::::])*
3	Why do you ask me that *(syllable-timed rhythm: **Why do you ask me that**)*
4	Maybe I don't want to go back laak[3] *('laak[3]' ((in Jyntping, or Cantonese Pinyin)) is a sentence-final particle in Cantonese connecting a sense of certainty or finality)* *H L L H H L H H* *(sudden rising and falling intonation: **May**be I **don't want** to **go back*** *H* ***laak[3]*** *(syllable-timed rhythm: **May be I don't want to go back laak[3]**)*
5	(immigration officer frowns with a down-turned smile; a popular *Douyin* sound effect commonly used to express surprise or shock co-occurs with the immigration officer's facial expressions)
6	(in Cantonese)喂喂喂喂! 啊! 啊! *(hey hey hey hey! ah! ah!)*

Extract 3

1 Immigration Officer:	Where will you be staying?
2 Tourist A:	at a hotel
3	with a friend *(emphatic lengthening: friend [fɹɛnd] → [fɹɛːːːːnd])*
4	at an Airbnb *H H L H* *(sudden rising and falling intonation: **Air b n b**)*

Extract 4

1 Immigration Officer:	What is your occupation?
2 Tourist A:	(in Mandarin Chinese) 怎么说来着 我百度一下啊 *(How do I say it in English? Let me search it online)*
3	No work *(coda /ɹ/ deletion and schwa epenthesis to closed syllable:* *work[wɜɹk]→[wɜkə])*
4	I am owned by someone *H L H L H H* *(sudden rising and falling intonation: **I** am **owned** by **some one**.)* *(syllable-timed rhythm: **I am owned by some one**)*
5	(in Cantonese)*咦? 咦? 喂, 喂, 又是你呀? 陈生? 啊?* (What? What? Hey, hey, is that you again? Mr. Chen? Huh?)

Extract 5

1 Immigration Officer:	Do you have anything to declare, Sir?

2 Tourist A:	No, nothing
3	(Tourist A drops a plier on the floor)
4	I'm a car mechanic *coda /ɹ/ deletion: car[kaɹ] ➔ [ka]*
5	so have a si ba la (subject omission: so I have a …➔so have a…) (si ba la: The Cantonese expression "si⁶ baa¹ naa⁴*²", which the transliteration and the loanword of "spanner", with the original tones removed)
6	very normal *([v]➔[w], [ɹ]➔[l],[n]➔[l],and coda /ɹ/ deletion: very normal [vɛɹi nɔɹməl] ➔ [vɛli ləməl])*
7	(Tourist A drops a pack of raw pork meat on the ground)
8	In the nighttime, I am a pig killer *(coda /ɹ/ deletion: killer [kɪləɹ] ➔ [kɪlə])*
9	(Tourist A imitates pig squealing)
10	very normal *([v]➔[w], [ɹ]➔[l],[n]➔[l],and coda /ɹ/ deletion: very normal [vɛɹi nɔɹməl] ➔ [vɛli ləməl]*
11	(Tourist A drops a bottle of alcohol on the ground)
12	Because…I drink…I kill…I *(schwa epenthesis to closed syllable and [z] ➔[s]: because [bikaz] ➔ [bikasə])* *(syllable-timed rhythm: **Be cau se**)*
13	(in Cantonese) 哎呀呀我的拖鞋啊喂! *(Ahh I dropped my slides!)*

Previous raciolinguistic research has used "Mock Asian" (Chun, 2004) or "yellowvoicing" (Chung, 2021) to mainly describe non-Asians' parodic revoicing of an immigrant from Asia to the U.S. through playful renderings of "fresh-off-the-boat" (FOB)

foreign accents and language style. Through imitating and ridiculing the stereotypical speech of an imagined Asian person, the discourse of "Mock Asian" or "yellowvoicing" associates linguistic otherness with racial and national otherness (Chun, 2004). In addition to the use of Mock Asian by non-Asians, Chun's (2004) analysis also reveals that this stereotypical discourse can also be employed by an Asian American to elevate their persona as a legitimate speaker of Mainstream American English (MAE) who is able to perform Mock Asian in public spaces. Furthermore, Chung (2021) uses "in-group yellowvoicing" and "implicit yellowvoice performances" to refer to the imitation of Asian speech by an Asian American who speaks American English as their native.

When Yang impersonates Tourist A in the video, his linguistic practices are similar to in-group "Mock Asian" or "in-group yellowvoicing" in the sense that he caricatures Tourist A as a Cantonese speaker who speaks English with various features of "Mock Asian", or "Mock Chinese" in particular. However, different from being performed by a native-speaker Asian American, Yang's "Mock Chinese" is further "in-group" in the sense that he is not only racially Asian but also linguistically and culturally Chinese. Yang's performance as a thickly accented Chinese tourist might arguably be a realistic impersonation of a Cantonese speaker because Yang speaks Cantonese, and he is familiar with how Cantonese speakers incorporate Cantonese and Mandarin language features into English. A mocking overtone is, nevertheless, implied through the excessive use of "Mock Chinese" features, the answers Yang provides as Tourist A to the immigration officer's questions, as well as other semiotic resources in the video.

As the focus of this section is *speech* as a multimodal resource in Yang's video, I will first move forward to explain how Yang's answers to the immigration officer's questions as

184

Tourist A contributes to the mocking of Chinese speech. I will elaborate on other semiotic resources that coordinate with *speech* to create mockery in later sections.

Voicing as Tourist A to answer the immigration questions, Yang portrays Tourist A as a Chinese person who intends to immigrate to the U.S. illegally. He states that he is currently unemployed (see line 3, Extract 4) and indicates his intention of finding a job and getting married in the States (see lines 4-5, Extract 1) without an immigrant visa in hand. He also fails to specify the length of his visit as he is considering not returning to China (see lines 2-4, Extract 2). The fact that he provides these obviously problematic answers with little awareness of the consequences constructs him as someone without much common sense or intelligence. Yang's linguistic performance as Tourist A, together with his dramatic play and other aspects of the video, communicates a comical image of a Chinese traveler and an illegal immigrant who speaks non-fluent English with a thick Cantonese accent.

While mockery can be used to generate humor that helps increase views and attract followers, Yang also needs to ensure the mocking is acceptable to *Douyin* audiences. Content that is interpreted by viewers as inappropriate or offensive can tarnish the public image of a microcelebrity and lead to a decrease in their popularity. The fact that Yang's airport video has been one of his most popular *Douyin* videos and one of the videos he wishes to promote the most (as it is a pinned video) suggests that the mocking component in the video is well received by *Douyin* viewers. The perceived legitimacy of Mock Chinese in Yang's performance is established through his construction of at least three aspects of his public persona on *Douyin*.

First, Yang as ethnically and culturally Chinese allows him to align with audiences and escape accusations of racism. Yang's Chinese identity is not only suggested by his look, which is associated with his transportable identity, but also indicated through his use of Mandarin and

Cantonese as his native languages. With Yang being an in-group member in mind, *Douyin*

viewers, who are mostly Chinese, would not assume that Yang would place the Chinese

language or Chinese people in a position of inferiority.

Second, Yang promotes himself as an actor and a comedian by paying homage to and

showing affinities with Stephen Chow, a Hong Kong actor who is immensely popular in Hong

Kong and Mainland China. Chow is well known for his performance of a type of slapstick

Cantonese humor known as "mo lei tau" in Cantonese or "wu li tou" in Mandarin, meaning

nonsensical, in Hong Kong comedy movies in the 1990s. The typical characteristics of Chow's

nonsensical humor include "non sequitur logic, slapstick expressions, exaggerated bodily

movements, and comical but mean retorts" (Chew, 2020, p. 127). Adopting Chow's style of

humor, Yang's speech when playing the role of Tourist A is constructed as absurd and lowbrow.

For example, Tourist A unashamedly admits to the immigration officer that the reason why he

does not need to work to make a living is that he is "owned by someone" (see lines 3-4, Extract

4), meaning he, as a sugar baby, receives financial support from a wealthy person in exchange

for companionship. Another example is when he is caught carrying a plier and a pack of raw

pork, he explains that he is a car mechanic and a butcher (see lines1-10, Extract 5), which does

not really justify his decision of not declaring those items. In addition to the language, Yang's

exaggerated facial expression and bodily movements in the video correspond to Chow's

flamboyant nonsensical manners. By aligning his comic performance with Chow, Yang

legitimizes his practice of Mock Chinese as part of the discourse of nonsensical humor.

Moreover, the poster appearing behind Yang on the top left side of the screen during the open

scenes (see Figure 18) features Stephen Chow in the movie *A Chinese Odyssey*, which has

received phenomenal popularity in China and has become a significant part of Chinese popular

culture. This visual element, together with the label "drama king" Yang uses in his bio to refer to himself, helps frame Yang's video as non-serious and playful. This interpretive frame also serves to legitimize Yang's Mockery of Cantonese speakers of English. Drawing on various multimodal resources to construct meaning relevant to the *Douyin* context, Yang projects an image of a comedian inspired by Stephen Chow. In so doing, Yang not only manages to convey his mocking humor in a non-offensive manner, but also establishes commonality and affiliation with viewers who are fans of Stephen Chow and Chinese popular culture.

Third, Yang's established identity as an English language instructor on *Douyin* and the pedagogical frame (as will be discussed in the following sections) of his video serves to justify his caricatures of Chinese/Cantonese speech as well. The excessive use of Chinese/Cantonese speech features in Tourist A's English and the problematic answers he provides stand in stark contrast to the language practice and responses of Tourist B. It can be argued that Tourist A's speech is deliberately constructed as an extreme case in order to draw learners' attention to the distinctions and facilitate learning. The purpose of teaching entailed by Yang's teacher identity and the pedagogical frame of the video can also help excuse Yang's Mock Chinese from ideological critique by mainstream Chinese viewers.

On the other hand, the viewers' appreciation of Yang's playful revoicing of a Cantonese speaker hinges upon the perception that Yang, as an English teacher, can speak a dominant variety of English. Using only Cantonese-accented English in the video may risk disqualification of the English teacher and expert status Yang puts great efforts in establishing. Through his linguistic performance as the immigration officer and Tourist B, Yang authenticates himself as a fluent speaker of MAE. He simultaneously uses mockery to deauthenticate his use of Cantonese accented English. In so doing, Yang elevates his persona and further commodifies himself as his

187

use of an "inner-circle" variety of English adds additional value to him as a *Douyin* influencer

who specializes in English language teaching. Although Yang's employment of Mock Chinese in

the video might be mainly motivated by his self-branding and teaching purposes, his adoption of

this discourse nevertheless problematizes accented English and the speakers of it. As such, a

hierarchical relationship is created within Chinese speakers between those who speak a socially

powerful variety of English and those who speak heavily Cantonese-accented English. This type

of linguistic in-group othering reinforces the standard language ideology, which, based on my

long-time observation of various kinds of *Douyin* informal English instruction videos, has been

widely circulating on this platform.

      Along with Mock Chinese features, translanguaging is used to characterize a "deficient"

Chinese learner of English. Tourist A's speech in the video is marked by his translanguaging

practice between English, Mandarin, and Cantonese. When Tourist A struggles to express

himself in English to answer the immigration officer's questions, he switches to Mandarin or

integrates features of different linguistic codes in one expression. For instance, when asked about

his occupation, Tourist A first speaks in Mandarin that he does not know how to put it in English

and asks to use his cellphone to search online (see line 4, Extract 4). What the role play suggests

here is that tourist A lacks English proficiency to allow him to carry a conversation in English

and he has to resort to the internet to translate Chinese into English. Moreover, when the

immigration officer asks the tourist if he has anything to declare, the tourist does not disclose any

items. A second later, however, a plier suddenly drops on the tourist's foot. The tourist picks up

the plier and explains, "I'm a car mechanic, so have a si ba la, very normal," (see lines 4-6,

Extract 5) meaning that as a car mechanic, the tourist normally carries a plier with him. The

Cantonese expression "si^6 baa^1 naa^{4*2}" (written in Mandarin as "士巴拿") is the transliteration

and the loanword of "spanner", which in the video is used to refer to the plier. However, by removing the original tones and applying flat tones on the three syllables, Tourist A pronounces this Cantonese expression nearly as "sī bā lā" in order to make it sound more like English with the hope that the immigration officer will be able to understand. Furthermore, at the end of the role play dialogues between Tourist A and the immigration officer, the tourist cries and complaints in Cantonese when he is being taken away by two other officers for possible detention. As such, Yang's playful translanguaging performance invokes an image of a Cantonese speaker who is unable to communicate in English fluently, through which translanguaging practice is stigmatized.

The juxtaposition of Mock Chinese features and translanguaging in Tourist A's linguistic performance is indexically linked to Chinese speakers whose spoken English "sounds Chinese" and who have limited English proficiency. Together with Yang's direct address of the viewer throughout the video using "大家" (*you* plural) and "你" (*you* singular), the role play episodes featuring Tourist A encourage viewers who wish to improve their English to relate themselves to the described scenario. With audiences' self-reference, showing them the scenes of Tourist A being detained and denied entry to the U.S as a result of not being able to communicate in English appropriately can serve to convince the viewer of the importance of learning English, especially a dominant variety of English. As a result, only certain kinds of English (i.e., prestigious varieties of English) are valorized and commodified as the "sociolinguistic scale" moves from local to translocal (e.g., an international airport) (Blommaert, 2010).

Following the discussion of the five commonly asked airport immigration questions with role play episodes and Yang's explanations, Yang concludes the informal lesson by reminding viewers of speaking to an immigration officer at an airport as a high-stakes situation and the

importance to prepare well before they travel abroad. Here again, Yang encourages following and promotes his language lessons by helping the viewer identify their potential learning needs. After the concluding remarks, Yang appeals to his following through a witty commentary on himself in Mandarin Chinese:

> 我是明明可以靠脸的 但偏偏不要脸的 *Mr. Yang*。飞。
>
> Translation: *I'm Mr. Yang. My face could be the moneymaker, but I insist on not saving it. See you.*

The Chinese phrase "不要脸", which literally means "not need face", is a euphemistic expression for "shameless". However, in Yang's message, "不要脸" is used as a pun to construct multiple public identities. When "不要脸" is directly interpreted as "shameless", Yang invokes his comedian identity and his nonsensical, exaggerated style of humor. On the other hand, by first stating that he could have earned money by simply attracting followers through his appearance ("My face could be the money the moneymaker…"), Yang indicates that he is good-looking, which is also suggested through his cover photo on his profile page in which Yang is surrounded by many enthusiastic young Chinese female fans. Within this context, what Yang actually means by "不要脸" is that he chooses not to rely on his physical appearance to make money effortlessly even if this option is available. As such, what Yang suggests about himself is that he is a hardworking and competent individual whose success is well-deserved. Although Yang does not explicitly encourage viewers to follow his account, his humorous evaluative message about himself is utilized to construct a positive image that can gain the favor of the audience.

Text, typography, and layout interplay

In Yang's video, both English and Chinese subtitles are provided for most parts of the speech. When Yang speaks in Mandarin or Cantonese, the Chinese subtitles and the translated captions in English are juxtaposed. Similarly, when Yang speaks in English, the English subtitles and the translated captions in Chinese are both featured in the video as well. However, the typography (including font size, color, and background) and layout of the subtitles vary to convey different degrees of *information value*, *salience,* and *framing* (Kress & van Leeuwen, 2006).

When Yang self-broadcasts in Chinese, Chinese and English subtitles are placed at the bottom of the screen. Captions in both languages are in white, while the font size of the Chinese subtitles is larger than that of the English subtitles (see Figure 18). In terms of salience, the texts in both languages are not arranged to draw the viewer's major attention because of their peripheral position and relatively small font sizes. Similarly, when it comes to framing, the white texts, the whiteboard next to Yang, and the grey background present a visual 'rhyme' in which the linguistic and other semiotic elements are brought together cohesively (Kress & van Leeuwen, 2006). The non-contrastive color of the subtitles to the rest of the visual elements in the scene also downgrade the salience of the texts. At a more detailed level, the white color applied to both subtitles creates a sense of connectedness between these two linguistic codes as a unified piece of information, which conveys a similar level of salience. Nevertheless, the Chinese subtitles are still evidently larger than the English captions in size. In fact, when viewing the video on a cellphone, the translated captions in English placed at the bottom of the screen are a bit too small to read. However, the corresponding Chinese subtitles are in a reader-friendly size. Moreover, the Chinese subtitles are placed above the English translations, through which

the Chinese text appears slightly closer to the center of the screen. This arrangement and design of subtitles increase the salience and information value of Chinese captions over the English substiles, as the Chinese text is immediately relevant to Yang's speech in terms of linguistic code. The Chinese subtitles near the bottom of the screen are used to make sure the content Yang broadcasts is accessible to most *Douyin* viewers, even if they are watching the video without the audio. Compared to the Chinese captions, the translated subtitles in English are much smaller. The main purpose of the English text is apparently not communication, but rather a means to demonstrate Yang's English language expertise and attract more followers. Yet still, the English captions can help Yang reach out to as many viewers as possible by including those who speak English but not Chinese. Moreover, the English subtitles contribute to the pedagogical frame of the video and provide resources for learning. As learners on *Douyin* have the agency to decide what to learn and how to learn, it is likely that some of them might use the English subtitles as learning materials. It is also worth noting that a black outline and shadow are applied to both texts. Therefore, although these white subtitles at the bottom of the screen are not designed to receive main attention, they are nonetheless far from being trivial.

However, when Yang presents the five commonly asked airport customs questions in English, each English question and its Chinese translation are positioned in the center of the screen to enhance information value. While both language subtitles are large in similar font sizes to be salient to the viewers, the translated captions in Chinese are placed under the English texts and different colors are applied to the two texts (see Figure 19). Compared to the layout of subtitles when Yang is self-broadcasting as shown in Figure 18, the position of the two linguistic codes in Figure 19 is reserved. This adjustment is consistent with the linguistic code of Yang's speech—the subtitles of the language Yang is speaking are always placed above the translated

192

caption in the alternative language (either Chinese or English). Additionally, although the

English text remains in white just like the subtitles appearing at the bottom of the screen, the

Chinese translation in Figure 19 is highlighted in yellow. Therefore, despite being placed under

the English text and slightly off-center, the Chinese translation in yellow is no less prominent or

important than the English question itself. By doing so, Yang emphasizes both the language form

in English and the meaning in Chinese to demand the viewer's attention with a pedagogical

orientation. Furthermore, to maximize the salience of the texts, the background of the scene is

colored in white to contrast sharply with the white and yellow subtitles.

Figure 18. The opening scene Yang's video of airport English

Figure 19. Introducing the first immigration question of discussion

When it comes to role plays in which two parties communicate in English, the English and Chinese subtitles are placed right next to the person who is speaking (see Figure 17). Specifically, the immigration officer Yang impersonates (see Figure 17/1) always appears on the left side of the screen and faces toward the right side, which indicates the position of the interlocutor. The English and Chinese subtitles for the immigration officer's speech are positioned on the right side near him, and at a similar level to the character's face. Similarly, Tourist A and Tourist B always show on the right side of the screen and face toward the left side, where presumably stands the immigration officer. Aligning with the direction that the tourists are facing, the English and Chinese subtitles for the tourists' speech are arranged on the left near them, and at a similar level to the characters' faces too. For each role, the subtitle positions make it seem that the texts originated in the character's mouth. The positions of the two characters in

194

each episode, together with the layout of the subtitles of their speech, help connect individual scenes that involve only one person at a time altogether and unify them as a conversation of two different people from a temporal dimension. As the role play dialogues are in English mainly, the English subtitles are placed above the translated captions in Chinese to match the linguistic code in speech. This arrangement of relational positions between subtitles is consistent throughout the video most of the time—the text that matches the speech in the code choice is placed above the translated captions. However, one exception is that when Tourist A translanguages to Mandarin Chinese or Cantonese in role play, the English text, or in other words, the translated captions in English, remains above the Chinese text of his utterances (e.g., see Figure 20). By prioritizing the relevant English expressions that Tourist A fails to express, Yang intends to keep the viewers' focus on the English language forms in order to facilitate learning. In addition, Yang communicates the expectation that English is the only language that should be used in a real-life scenario of going through airport customs that the role play simulates. Apart from the layout, the typography of the texts in role play snippets serves to enhance the prominence of the language as well. Both English and Chinese subtitles are highlighted in yellow with a black outline and shadow to increase the visual appeal.

By providing both English and Chinese subtitles, and through combining texts with various semiotic resources to direct the viewer's attention to the information Yang considers most important and relevant to learning, Yang strategically adopts translanguaging as a pedagogy into his online language teaching videos. Translanguaging also helps Yang create an inclusive environment for viewers with different English proficiency levels. Using translanguaging as a teaching method, Yang also demonstrates his multilingual competence,

195

pedagogical knowledge, as well as his multimodal design knowledge. All of these contribute to a language expert identity in digital contexts.

The only section in which translated captions are absent is the post-credit scene (see Figure 22). As will be illustrated in a later section, what Yang says and presents in the post-credit scene mainly serves to entertain the viewer and address of the importance of learning English. Its lack of relevance to teaching is suggested but the subtitles as well—there are only Chinese subtitles for the Chinese speech while English translations are not given. Therefore, the code choice of subtitles can also play a role in framing the video content in relation to its pedagogical orientation.

Apart from the texts superimposed on the video, there is also a written description as part of the add-on information appearing at the bottom of the cellphone screen (see Figure 21). The written description of a video is presented throughout the playing time of the video on *Douyin*, A downloaded video does not include the written description. However, it is important to address this text as content creators rely on video descriptions to communicate ideas to the viewer. For Yang's airport English video, the video description Yang adds is:

> *"出国过关怎么办？快看机场过关必备英语#机场#英文#杨老师的英语@DOU+小助手"*

> Translation: *"what do you need to know when going through the customs? Check this video for English you must know to go through the airport customs. #airport #English #Teacher Yang's English @Dou+ marketing tool".*

In this written description, Yang first asks a question in order to raise audiences' interest and make them think. Then Yang summarizes the focus of the video ("English you must know to go through the airport customs") and encourages the viewer to watch the video to look for ways to

address the question he asks. Following the description are three hashtags. While the first

hashtag *#airport* addresses the language context and the second hashtag *#English* emphasizes

which language Yang teaches, the third hashtag *# Teacher Yang's English* focuses on Yang's

English teacher identity. Moreover, it is used to mark Yang's English by creating to label for it.

As content searchable by others, the hashtag *# Teacher Yang's English* Yang adds serves to

promote him as a language teacher and his English.
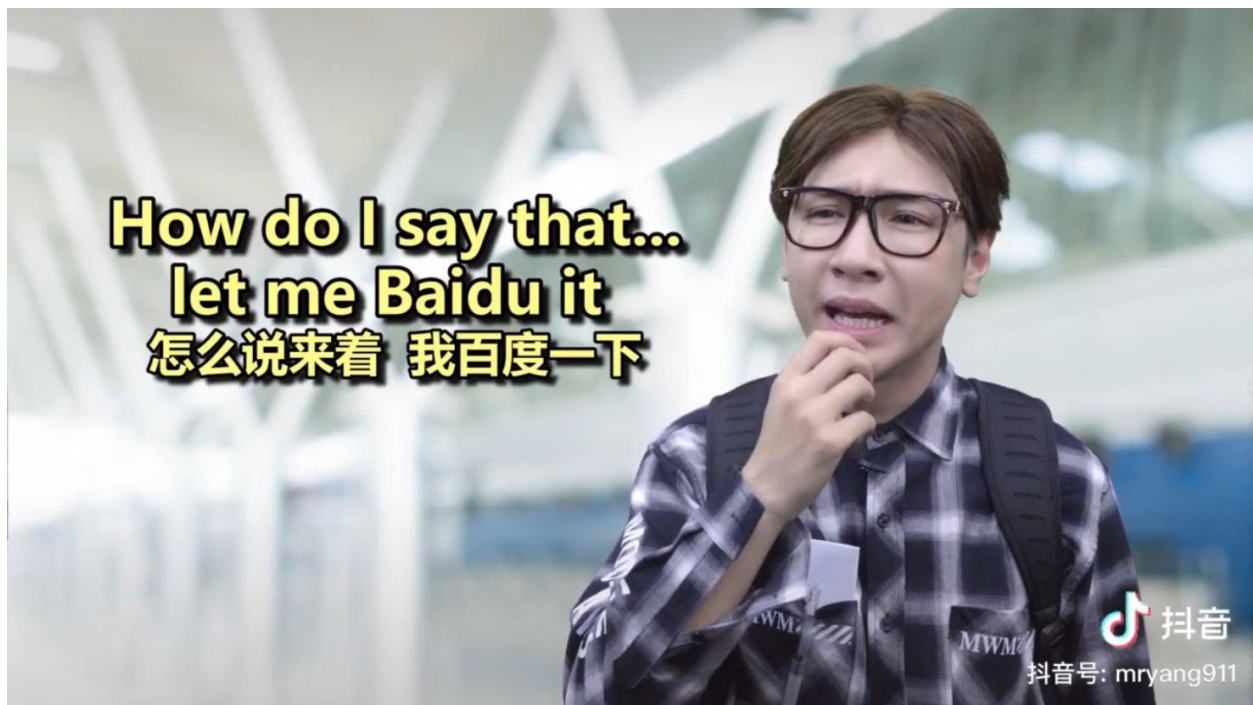
Figure 20. Tourist A speaks in Mandarin Chinese

Note. Tourist A speaks in Mandarin Chinese "怎么说来着，我百度一下" while the Chinese
subtitles are placed under the translated captions of English "How do I say that…let me Baidu
it". Baidu is a dominant internet search engine in China. Similar to the word "google" in English,
"baidu" can be used as a verb to refer to the action of using the Baidu search engine to search for
information online.

Username: @MrYang 杨家成(Mr. Yang Jiacheng Yang)
video description: what do you need to know when going through the customs? Check this video for English you must know to go through the airport customs. #airport #English #Teacher Yang's English @Dou+ marketing tool

Time when the video was posted (only visible when the video opened from the video list, not available when the video appears on the "For You" feed): 07/17/2019

Audio information: Mr. Yang Jiacheng Yang's original sound

Figure 21. Add-on information of Yang's video lesson of airport English

Image

      To visually frame the video as instructional during the opening scenes (see Figure 18),

Yang is standing in the center of the frame while some Chinese and Western cultural items are

arranged in the background surrounding Yang. The whiteboard on the left side, behind Yang,

creates a classroom-like setting. Exactly what kind of class is then signified by a series of

semiotic resources, including the Stop sign and the phrase "I hate flying" on the whiteboard, the

Western cultural items (e.g., a faux dear head, a framed poster with the U.S. map, a mini

chalkboard with the sign "Mr. Yang" and "welcome") as wall decorations. In terms of *salience*

(Kress & van Leeuwen, 2006), although these semiotic elements are smaller in their visual

'weight' due to their sizes and placement, they nevertheless are combined to suggest the video

lesson is about travel English. The central position of Yang is likely an indication of greater

information value, especially in the Confucian cultural context such as China where centering is

commonly associated with a higher social status (Kress & van Leeuwen, 2006). In Chinese

traditional values, teachers are highly respected as the authority of knowledge in the classroom.

As such, Yang's position to the camera, together with the instructional and cultural items

surrounding him, creates a classroom environment specific to the Chinese educational context

where teachers are held in high esteem. These semiotic arrangements also help Yang

communicate his identity as a respectable teacher at the beginning of the video. Meanwhile,

Yang points at the viewer with one hand, which creates a visual form of direct address. It

acknowledges the viewers explicitly, addressing them with a visual "you". In addition, Yang's

gaze invites the viewer into an imaginary relation with him. The specific type of this relation is

then understood through other modes, including what he says, the size of the frame, and his

facial expression. As he is explicitly addressing the viewer both linguistically and visually, is

smiling at the viewer, and the size of the frame is a medium close shot, his direct eye contact subsequently establishes social affinity with the audience. In addition to the beginning scenes of the video lesson, the same classroom setting is also repeatedly used almost every time Yang self-broadcasts in the video. In fact, when the video features Yang facing the camera to give explicit instruction, he is always situated with the classroom setting. With this meaningful integration of visual elements, the viewers are constantly reminded of the instructional purpose of this video, and Yang's identity as a language teacher is further reinforced.

In the role play episodes, visual cues are essential to mark and portray the different roles played by Yang. When Yang is dressed in his professional uniform, wears a blond wig, stands, behind a computer monitor, appears on the left side of the frame facing the right side, and speaks fluent American English, he performs as a U.S. immigration officer. When the video features him off-center to the right side, he is playing a role of a Chinese tourist. Specifically, when he wears a pair of thick-rimmed glasses and a blue plaid shirt buttoned all the way up, stands there with a forward head and hunched shoulders, and speaks English with a thick Cantonese accent and imperfect grammar, he constructs an image of a nervous Chinese tourist (Tourist A) with low English proficiency. Thick-rimmed glasses in Chinese culture are often used to index a nerdy, socially awkward, and unattractive person. This indexical meaning conveyed through visual elements matches with the unnatural postures (e.g., hunched shoulder, exaggerated facial expressions) of Tourist A and his pragmatic failure. When Yang wears no glasses and a red plaid shirt open, carries his backpack with one shoulder, and speaks fluent American English, he performs another Chinese tourist (Tourist B) who is composed and with advanced English language proficiency.

Through role play, Yang constructs multilayered, multifaceted discourse identities. By engaging in the production of *Douyin* videos, Yang's discourse identity in general is a content producer. It is through focusing on English language teaching in his video that Yang constructs a more specific layer of identity as an English expert/teacher. Role play is a method Yang, as a language instructor, adopts to create meaningful, learning context. As Yang plays three different roles in the role play, his discourse identity shifts between an immigration officer and two Chinese tourists with different English proficiencies.

Moreover, the immigration officer in the video wears a blonde wig, the color of which is a visual index of the prevailing stereotype of an authentic, native English teacher—a White native speaker of English. However, Yang, who is ethnically Chinese, is not automatically granted authenticity by his non-Western appearance, or in other words, his transportable identity. By playing the role of a US immigration officer, who is designed by Yang to be blonde and speak in MAE, what Yang implies is that his English proficiency is commensurate with that of a native speaker despite his non-Western look. By demonstrating that he can pass as a native speaker of English, he claims his authenticity in English language use.

It is also worth noting that the T-shirt Yang wears in the classroom setting may also serve to index his authenticity as English language expert. The English expression "TrueFake" printed on his T-shirt (see Figure 18) may humorously allude to the different discourse identities he constructs in the role play episodes. Apparently, Yang's performance in the role of the immigration officer and Tourist B is perhaps a true reflection who he is—a competent English language user who can pass as a native speaker, and his portrayal of role of Tourist A is clearly fake.

Sound and image interplay

Transitions between lecture and role play episodes in Yang's video are not made through explicit reminders in speech or text. Instead of saying something like "I will show you a role play" or present text such as "discussion time", the alternation of these different parts of video is often realized through auditory and visual modes. For instance, after Yang raises the question during the opening scene, the video presents an image of a TV error effect (see Figure 22). A TV static sound effect is playing in the background simultaneously. This combination of visual and audio resources is used to imitate what a vintage television would be like when changing channels, through which viewer are reminded of the transition in video content. The signposting realized through audio and visual cues not only helps avoid unnecessary interruption of the flow of the lesson, but also allows the content creator to present information in a more coherent, structured way to facilitate viewers' understanding.
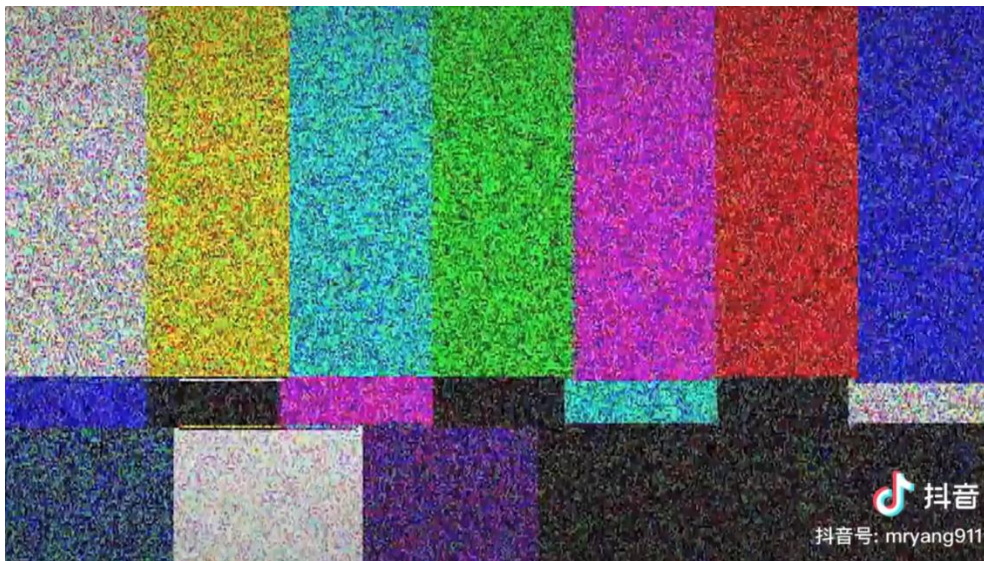
Figure 22. Transition scene

Language and image interplay

The final part of the video is constructed as a post-credit scene telling viewers what happened to Tourist A after he was detained (see Figure 23). In this scene, Yang adopts a type of Guangxi (a southwest province adjacent to Canton) accent when performing Tourist A, who admits in Mandarin that he will never get through immigration in his entire life because he does not speak English. Yang's performance of various accents in this video is a clear indication that he is very linguistically adept. This scenario constructed by Yang originates from and intertextually references a viral video interview (see Figure 24) on Chinese social media, in which a thief named Liqi Zhou, was seen detained at a police station, with one hand chained to a window. During the interview, Zhou proudly proclaims in Mandarin with a strong Guangxi accent that the reason why he repeatedly committed larceny is that he would never work for others in his entire life and that he does not know how to run a business either. The intertextual reference of the post-credit scene to Liqi Zhou's interview is invoked not only through Yang's revoicing of Zhou's Guangxi Mandarin and his format of response but also through visual elements as Yang's right hand chained to a window just like Liqi Zhou's. The combination of the language use and visual resources in the post-credit scene together constructs this part as a parody based on Chinese online culture, which serves to entertain the viewer. Meanwhile, through Tourist A's voice, by associating the ability to speak English with the possibility to connect to the world ("pass through customs"), Yang highlights the significance of learning English. What happens next in the role play—the handcuffs open suddenly (possibly due to an issue with the locking mechanism) and Tourist A mistakenly takes it as an official notice that he is allowed entry to the foreign country. This adds an additional layer of humor.

Figure 23. Post-credit scene

Note. Yang as Tourist A says: "It's impossible to pass through customs in my entire life and I don't speak English either."

Figure 24. Interview with Liqi Zhou that went viral on Chinese social media

Note. Liqi Zhou claims in the interview: "It's impossible to work for others in my entire life and I don't know how to run a business either."

Linguistic, visual, and auditory interplay

In the role play section, when Tourist B tells the immigration officer that the purpose of his visit is to shop in the States, the officer immediately becomes thrilled and welcoming. Another immigration officer with a blond wig comes out of nowhere into the screen joining the

previous officer to welcome Tourist B (see Figure 25). Both of them are cheering and shouting

out loud "Welcome" multiple times. Yang as one of the immigration officers is also rubbing two

fingers and his thumb together, which is a Chinese hand gesture signifying money. While they

are cheering, colorful confetti are falling at the same time. Meanwhile, the Spring Festival

Overture, a tune widely used on various festive occasions in China, is playing in the background.

With the knowledge that Spring Festival Overture used for festive occasions cannot be played at

a U.S. airport customs, and immigration officers would not behaved as what is presented in the

video because airport immigration is a formal, political context, the viewer is able to appreciate

the humor generated by multimodal stylistic incongruity (Jiang & Vásquez, 2020; Piata, 2019;)

realized by the interplay between language, visual, and auditory resources. In terms of identity

construction, this playful scene is also a reflection of Yang's non-sensical humor and helps

project him as a comedian.

   In addition, the viewer's appreciation of this humor may also hinge on the stereotype

of Americans being economic materialism and capitalism and of Chinese tourists always

spending extravagantly on luxury goods when travelling overseas. As suggested in the role play,

the immigration officer seems not care about whether Tourist B is eligible for the entry into the

U.S. upon hearing that his purpose of visit is to shop. Instead, the officer simply judges the

tourist by his economic value. This negative stereotype of Americans and the stereotype of

Chinese tourists being rich and spendthrift are both, to certain extent, acknowledged and

accepted by Chinese online users. With the rise of Chinese nationalism (Weiss, 2019), criticism

on the U.S. and other countries which are considered political alliances of the U.S. are

propagated on Chinese mass media and social media. Meanwhile, under the influence of national

discourse, many Chinese people take pride in the purchasing power the Chinese consumers have

exhibited to the world and take it as a sign of China's modernization and economic power. In China, the negative media coverage about China from the U.S. and other Western countries is widely propagated. With the awareness of being viewed as a trade competitor and an emerging military rival, as well as the negative accounts of the U.S. from China's state media, many Chinese people in turn perceives the U.S. unfavorably and as their major competitor as well. By playing the role of Tourist B who is welcomed by the U.S. immigration officers for being financially capable, Yang aligns himself with viewers who celebrate the negative stereotype of Americans and the perceived positive image of Chinese that Yang invokes in the video. As such, Yang's stance is established through the identities he performs in the role play, which reflects the positionality principle in Buchotlz and Hall's (2010) framework for identities in interactions.

Figure 25. Immigration officers welcoming Tourist B

Summary of the analysis

The multimodal analysis of Yang's video of airport English has shown that Yang harnesses various linguistic (e.g., different languages and accents) and other semiotic resources (e.g., background, clothing, accessories) to construct different discourse and situational

identities, including a narrator, an English teacher, a comedian, an U.S. immigration officer, and two different Chinese tourists. As for personal identity, he also draws on multimodal elements to perform an image of an English language expert and an attractive male. For instance, Yang's passing as a native speaker of English are realized through the interaction of speech and visual elements (i.e., a blond wig), through which he establishes authority and expertise in English. It is also found that the multimodal construction of the hybrid genre of the video, which consists of lecture, role play, and movie (post-credit scene) serves to present different identities, deliver language instruction, and entertain viewers. In addition, multimodal resources are utilized to make references to online and popular culture in China, which creates a sense of playfulness to keep viewers engaged with the content. Furthermore, the adoption of both English and Chinese subtitles as well as the variation in their layout, font size, and color play an important role in framing the video as instructional.

Case Study 2: Dabai's Video of Vocabulary

The second video I selected for an analysis of its modal resources used for identity construction and language teaching is from @大白外教英语 (Foreign Teacher Dabai's English), the second most followed account among the 300 collected accounts. The video is one of Daibai's non-pinned videos and has received the most likes among the three non-pinned videos collected from his account. There are another three pinned videos collected. However, because none of them focuses on English language teaching, they cannot be used to investigate the content creator's employment of semiotic resources for delivering language instruction. The selected video from Dabai focuses on vocabulary. In the video, Dabai presents several pairs of English words that are similar in spelling but differ in meaning. The video is about 20 seconds in

length. Below, I analyze Dabai's adoption of multimodal resources to construct a public self and teach English in order to attract followers.

Speech

In the video, Dabai introduces vocabulary words by reading those words and their Chinese meaning out loud. He also concludes the video and appeals to likes in Mandarin exclusively at the end of the video. As a non-Chinese person, Dabai's ability to speak the Chinese language helps him establish a close relationship with the Chinese audience by relating to them linguistically and culturally. Moreover, speaking Chinese allows him to communicate and deliver language instruction to Chinese earners more effectively.

Text, typography, and layout interplay

At the beginning of the video, Dabai introduces the topic of the video through text. As shown in Figure 26, the Chinese text "加不加 s 单词意思大不同" (meaning: adding an -s ending may completely change the meaning of a word) superimposed on the opening scene summarizes the teaching focus and indicates the language skill addressed in the video. The text is positioned under the images of Dabai, but still remains in the center of the video frame to indicate its information value. The font size is large and the yellow font color contrasts with the white text background, enhancing the salience of the text to direct the viewer's attention to it.

Figure 26. The opening scene of Dabai's video of vocabulary

Note. Translation of the Chinese text superimposed on the opening scene: adding an *-s* ending may completely change the meaning of a word

The four pairs of English words Dabai compares for their different meanings are: *good* vs. *goods*, *green* vs. *greens*, *new* vs. *news*, and *glass* vs. *glasses*. The comparison is visualized relying on layout as well as typography (see Figure 27 for an example). In terms of the layout of the text, for each pair of words, Dabai first presents one that does not end with an *-s* or *-es* on the left of the screen by reading the word and its Chinese translation out loud. As he reads, the English word and the Chinese meaning as subtitles appear around the bottom of Dabai's left

image. During this time, the right side of the screen is frozen. The left side of the screen freezes immediately when Dabai finishes uttering the word and its meaning in Mandarin, and the captions remain. Then the right side of the screen shows Dabai presenting the second word that ends with an -s or -es and its Chinese translation. The English and Chinese subtitles follow along with Dabai's speech. As soon as Dabai reads the word and its Chinese meaning, the right side of the screen freezes again. Once both of the words in a pair are presented, the captions are removed, which suggests the end of the comparison and that a new pair of words will be introduced. Then the video moves on to show the next pair of words following the same structure: the word without an -s or -es suffix is introduced first on the left side of the screen, while the right side of the screen is frozen. Then the left side of the screen freezes, directing viewers' attention to the right side where the second word with an -s or -es ending appears, and so on.

When it comes to typography, the spelling difference between each pair of each is marked by yellow. For instance, when the text "new" and "news" are presented, the letter "s" in the second word "news" is highlighted in yellow while the word "new" and the first three letters in "news" are in white (see Figure 27). The spelling differences in the rest of the three pairs of words are marked following the same pattern: the letter "s" in "food" is highlighted in yellow when the words "food" and "foods" are juxtaposed; the letter "s" in "greens" is in yellow when "green" and "greens" are compared; the letters "es" in "glasses" are made salient in yellow when the words "glass" and "glasses" are shown in the text.

While comparing the spellings and meanings of those words, the Chinese explanation "加不加 s 单词意思大不同 (meaning: adding an -s ending may completely change the meaning of a word) that is positioned under the images of Dabai at the opening scene is moved to the very top

of the screen, which appears to be marginalized. The text stays in yellow but slightly shrinks in the font size. These adjustments in typography and layout, probably motivated by fact that the information is no longer new, result in reduced value and salience.

Figure 27. Dabai compares the spelling and meaning of "new" and "news"

Text and image interplay

Back to the opening scene as shown in Figure 26, the white background of the text and the still images of Dabai create a sharp line as a boundary between the photos and the writing, treating them as two separate sections. A vertically structured composition with an upper and a lower section signals "polarization", which "conveys a sense of contrast or opposition" between

the two distinct spaces (Kress & van Leeuwen, 2021, p. 190). When there is a contrast between the top and bottom, the information appearing in the upper section is presented as the ideal, which refers to "the idealized or generalized essence of the information" and thus is more prominent and noticeable (Kress & van Leeuwen, 2021, p. 191). On the other hand, the information placed at the bottom coincides with the real, which is more detailed, informative, and practical. The images in the video screenshot presented in Figure 26 are positioned in the upper section, playing the "ideal role" to visually foreground the layout of the images showing up in the rest of the video. Specifically, the image section in the opening scene is designed as a split-screen structured along the horizontal axis. Both the left side and right side of the screen feature Dabai facing directly toward the camera, remaining in the same posture, and wearing the same white baseball cap. The main difference between Dabai's two images lies in the color of his shirt. It is also through the two different colors that the boundary between the two sides of the split-screen is visually marked. The visual composition of the image section during the main body of the video, in which Dabai introduces and compares four pairs of English words with similar spellings but completely different meanings, is consistent with the opening scene as a horizontally oriented split-screen. In contrast, the text under the images in the introduction scene plays a real role, which serves as an explanation of the purpose of adopting a split-screen.

When Dabai reads the English words and then their Chinese meaning, the subtitles are separated by the two linguistic codes into two rows, aligning with the vertical layout of the video.The English words are placed above their Chinese translation following the general reading convention from top to down, through which the spatial arrangement matches the temporal order (i.e., each English word is introduced first and followed by its Chinese meaning). The English caption is superimposed on the bottom of either the left or right side of the screen

depending on which side of the screen is playing, while the Chinese translation is positioned underneath, situated in the black margin at the bottom part of the video frame. Both the English and Chinese texts are in white. The coherence in color visually constructs these two linguistic codes as one unified piece. The English text as part of the subtitle section is superimposed on the image, yet the Chinese part of the subtitle is positioned outside of the image. In doing so, both the image and the textual frames become porous and blurry, creating a spatial interaction between the images and the texts. In the meantime, however, the demarcation generated between the images and the black margin at the bottom of the screen separate the English and the Chinese text. Therefore, the meaning conveyed from the interaction between the subtitles and the images is two-fold. First, both the English words and their Chinese meanings are the focus of the learning content. Second, the English text is more immediately relevant to what Dabai is teaching—English vocabulary words. Nevertheless, the Chinese text is important in its own way. Although both language texts are in white, the Chinese caption forms a greater contrast with its black background. As such, Dabai directs viewers' attention to the meaning differences between each pair of words with similar spellings.

It is also worth noting that the image frame during the discussion on vocabulary (see Figure 27) does not take the top and bottom of the screen. Rather, it is arranged in the center, leaving the upper and lower part of the whole video frame as black margins. The arrangement can be interpreted from top to bottom as a margin-center-margin structure. Needless to say, the multimodal text positioned in the center is most important and salient. Similarly, during the opening scene (see Figure 26), although there is a clear dividing line to separate the text and images into two sections, they both appear in a relatively central position. Together with the black margins on the top and bottom of the screen, the compositional structure of the whole

scene can be divided into margin-center-center-margin. In addition to focusing viewers' attention on the most important and prominent information, this type of structure also serves to optimize audiences' viewing experience and avoid unnecessary overlap between the add-on information (e.g., account name, video description, see Figure 17) superimposed on the bottom and the video content.

Since the layout of Dabai's video is vertical, splitting the image frame into left and right sides makes the two parts of the frame vertical in layout as well. This structural alignment or layout coherence is partially contributed by the platform affordance. As a video-sharing mobile app, most videos on *Douyin* videos are shot and viewed in a vertical view. Therefore, a split-screen view with left and right sides caters more to the viewing convention compared to a top and bottom split-screen. More importantly, the horizontal juxtaposition of the two parts of the screen as opposed elements, which are realized both through image and text as discussed above, also denotes a polarization between the *Given* and the *New* (Kress & van Leeuwen, 2021). The left side is presented as information already known. In contrast, the right side is where the new information is placed and demands more attention from the viewer. Comparing the four words presented on the left side of the screen ("good", "green", "new" and "glass") to the other four introduced on the right side of the screen ("goods", "greens", "news", and "glasses"), it is not hard to tell that the four words shown on the right side are used relatively less frequently and thus are more likely to be new information to the viewer. Furthermore, by following the sequence of given to new, easy to difficult to teaching vocabulary, Dabai demonstrates his awareness of content scaffolding and pedagogical knowledge, which contributes to his public image as an English language teacher.

Username: @大白外教英语 (Foreign Teacher Dabai's English)

Video description: adding an -s ending may completely change the meaning of a word #English #English vocabulary #Beginner English @ Douyin Asistant

Time when the video was posted (only visible when the video opened from the video list, not available when the video appears on the "For You" feed): 3 days ago

Audio information: Foreign Teacher Dabai's English's original sound

Figure 28. Add-on information of Dabai's video of vocabulary

As far as the linguistic codes for the text, Dabai uses both English and Chinese in the

video. Compared to the English text that is only used to present English vocabulary words,

Chinese text is exclusively used by Dabai for other purposes including introducing the topic,

explaining the meaning of the English words, concluding the lesson, appealing to likes, and

providing a video description which appears at the bottom of the screen on *Douyin* platform. In contrast to his image as a non-Chinese person, the fact that Dabai is competent in using Chinese text to communicate ideas effectively, together with his fluent speech in Chinese, likely results in appreciation from and establish an affiliation with the Chinese viewers on *Douyin*.

Linguistic, visual, and auditory interplay

Right after the comparison of the four pairs of words, the split-screen feature is removed and so are the black margins, indicating the end of the instructional content. As shown in Figure 29, the screenshot features a medium shot of Dabai from a frontal and an eye-level angle, through which Dabai invites the audience to get involved in a non-hierarchical relationship with him. The transition from a centered, split-screen featuring two images of Dabai to a single image of Dabai that takes up the entire screen creates a zoom-in effect that reduces the distance between the content creator and the viewer. In terms of language, Dabai first asks the question "got it?" in Mandarin Chinese to the viewer while looking at the camera directly with a smile, through which he presents a caring, approachable teacher identity. Then, he as a *Douyin* content producer ends the video by appealing to likes ("Don't forget to leave a like"), again, in Chinese. While he keeps smiling and having direct eye contact with the viewer, he raises one hand close to the camera, making a "finger heart" gesture by crossing his thumb and index finger. Originated from South Korean popular culture, the "finger heart" gesture in recent years has become popular among the young generations of Chinese and is widely used as an adorable way to express affection, love, and friendship. As the *Douyin* like button is a heart-shaped symbol, the "finger heart" gesture Dabai makes serves to remind the viewers to click the like button. Moreover, by conveying affection to the audience through the use of a culturally specific gesture, Dabai establishes affinity and common ground with the followers through emotional appeal and

216

cultural competence. Additionally, when Dabai encourages the audience to like his video and makes the "finger heart" gesture, a popular *Douyin* sound effect is played in the background. The name of this sound effect is "打卡成功" (meaning: mission complete), which can be found in *Douyin*'s official editing app *Jiangying*. As the name of the sound effect suggests, it is often used to suggest a task is completed, or a goal is accomplished. Combining with Dabai's hand gesture and his verbal appeal to likes, the sound effect further motivates the viewers to leave their likes to the video. Moreover, the adoption of a popular *Douyin* sound effect reflects Dabai's familiarity with Chinese digital contexts, which affiliates him with the *Douyin* audience.

Figure 29. Dabai appealing to likes

Summary of the analysis

As a Caucasian person, Dabai's English language expert identity is indexed in his race, which is associated with a native speaker of English image. Therefore, instead of focusing on demonstrating English language proficiency, the content creator prioritizes creating simple and accessible instructional content in order to demonstrate his competence in teaching Chinese learners English on *Douyin*. To ensure the accessibility of the lesson to most Chinese viewers, Dabai narrates in Mandarin Chinese except for when reading the English words he is teaching. Both English and Chinese subtitles are provided the vocabulary words of discussion. Furthermore, the comparison of the four pairs of English words with similar spellings but different meanings is visually enhanced through a combination of different modal resources including a split-screen of left and right configuration, motion of the pictures, and color. The multimodal design of the instructional content serves to present information in a systematic and coherent manner, through which Dabai demonstrates his expertise in online English teaching.

Using Chinese is an important way through which Dabai as a non-Chinese person build connections with Chinese audiences. In addition, he also adopts visual and auditory elements from Chinese online culture to establish affiliations with *Douyin* users.

Case Study 3: Scott's Video of English Prepositions

The next video I discuss for its linguistic and other semiotic resources used to serve a self-branding and language teaching purpose is chosen from @ Scott 学口语's (Learning Spoken English with Scott). In this video, the content producer, Scott, provides a short lesson of two prepositions "in" and "on", and when to use which to talk about transport.

Speech

In his video about using the two English prepositions when discussing means of transport, Scott narrates in Mandarin Chinese as the instructional language. The English language is only used to provide example collocations with the two prepositions discussed in the video. Adopting Mandarin Chinese as the instructional language ensures the accessibility of the content to Chinese online users as well as Chinese learners of English of different levels. Meanwhile, it is important to point out that although Scott speaks in Mandarin in the video, his Chinese speech shifts among three styles to construct different identities.

At the beginning of the video, Scotts introduces the topic of the informal lesson by asking the viewer to select the correct preposition between "in" and "on" to collocate with "the bus" (see Extract 6, line 1). The question is asked in Mandarin. However, Scott stylizes his speech by pushing his tongue forward between the front teeth, changing an unaspirated alveolar dental affricate sound /t͡s/, an aspirated retroflex affricate sounds /t͡ʂʰ/, a voiceless retroflex fricative sound /ʂ/, and an unaspirated alveolar-palatal affricate sound /t͡ɕ/ into interdental sounds, which do not exist in standard Mandarin. Speaking Mandarin with interdental sounds leaves the impression that the speaker talks with a lisp. Considering the pedagogical frame of this video, Scott's use of interdental sounds in Chinese might serve to draw learners' attention to one of the phonological differences between English and Mandarin, which is the existence of interdental sounds in English. Many Chinese learners of English find interdental sounds particularly challenging because these sounds do not exist in their native phonetic inventory. As a result of intensive practice, Chinese learners may find themselves overapplying interdental sounds in their spoken English or even Chinese at a certain stage. Through the playful adoption of the

interdental articulatory feature from English to Chinese, Scott relates to learners' struggles and generates a sense of humor.

Preposition usage in English is another challenging aspect for Chinese learners. Thus, immediately after Scott raises the question of choosing between "in" and "on" at the opening of the video, he reassures the viewer that the problem is easy to solve and they just need to remember one rule (see Extract 6, line 2). Here, instead of applying interdental sounds to his speech, he switches to Sichuan-accented Mandarin Chinese. With the influence of the frequent adoption of Sichuan dialect in Chinese movies to generate a comic effect (Yao, 2013), Sichuan dialect has become a popular and recognizable Chinese regional dialect in Chinese popular culture and online spaces. In addition to conveying a sense of playfulness, Scott's use of Sichuan accented Mandarin also indexes his regional identity as a Sichuanese. This identity is also indicated in his bio on the profile page, in which he marks his location as Chengdu, the capital of Sichuan. Sichuan as a southeast province in China is famous for a variety of things (e.g., Sichuan cuisine, national parks, the habitat of giant pandas, a laid-back lifestyle) and thus a popular tourist destination. Therefore, by bringing viewers' attention to his regional identity through various semiotic resources, Scott also styles his persona in addition to creating humor.

Following his Sichuan accented Mandarin speech, Scotts switches to standard Mandarin to explain the grammar rule relevant to the preposition "in" (see Extract 6, line 3). This linguistic code choice helps Scott deliver language instruction that is most readily understood by the most Chinese viewers. Using standard Mandarin also indicates a higher degree of formality of the speech and allows Scott to adopt a "teacher's talk", as teachers in China are expected to speak Mandarin in class. In doing so Scott conveys a sense of seriousness about the grammar rule he is explaining and invokes his language teacher/expert identity. Furthermore, Scott manages to

explain the rules for collocation of "in" with the types of transportation vehicles in one sentence. Different from traditional grammar books and classroom instruction that involve complicated rules and tedious elaborations, Scott's explanation is concise, straightforward, and easy for learners to follow. His ability to turn complicated grammar rules into highly accessible instructional content demonstrates his expertise in both the English language and teaching, through which he proves himself as a qualified English language teacher.

Extract 6

1	(in Mandarin Chinese with an interdental lisp): "在公交车上"是 in the bus 还是 on the bus? *Translation: Which is correct? "in the bus" or "on the bus"?*
2	(in Sichuan accented Mandarin): 简单得很嘛 一条规则 *Translation: It's very simple. You just need to remember one rule.*
3	(in Standard Mandarin): 你必须坐在里面的就是 in *Translation: You use 'in' before the object that requires you to sit inside.*
4	in a car
5	in a taxi
6	in a helicopter
7	(in Sichuan accented Mandarin):*这些交通工具你几乎没有办法站立起来* *Translation: You can barely stand in those transportation vehicles.*
8	*身体可以立起来的就是 on* *Translation: You use 'on' before the object that you can stand inside.*
9	on a ship
10	on a bus
11	on a bike
12	on a skateboard

13   (in Sichuan accented Mandarin)点赞!
            Translation: Give me a like!

To help learners comprehend the rule he just summarized, he speaks in SAE and provides three examples of collocation of the preposition "in" with a specific type of transportation vehicle, which are "in a car", "in a taxi", and "in a helicopter" (see Extract 6, lines 4-6). After giving these examples, he further explained in Sichuan accented Mandarin that the transportation means included in the three example phrases do not have enough space to allow a passenger to stand inside, and therefore the preposition "in" is used. The reason for him to switch back to Sichuan accented Mandarin to provide a further explanation is probably because this part of the information is only supplementary and therefore less important, and because Scott intends to entertain the viewer and keep them engaged.

Moving on to present the rule for the collocations of "on" with the types of transportation vehicles, Scott again uses standard Mandarin to ensure the clarity of the message. One more time, the grammar rule is summarized in one sentence ("You use 'on' before the object that you can stand inside.") (see Extract 5, line 8) to make it easy to understand. Following the grammar rule, Scott lists fours phrases as examples of collocation of "on" with a transportation means, which are "on a ship", "on a bus", "on a bike", and "on a skateboard".

At the end of his video, Scott switches again to Sichuan-accented Chinese and encourages audiences to like his video by saying "give me a like!". The imperative sentence structure, together with his forceful tone, frames the message as a command and communicates an authoritarian stance through which Scott positions himself as a respectful English expert who deserves good evaluations. In the meantime, the Sichuan accent he adopts serves to soften his tone and make his request less demanding, as speaking a Chinese dialect other than Mandarin suggests a certain degree of informality.

It is also worth noting that despite the "non-standard" ways he adopts for part of his Chinese speech, whenever Scotts speaks in English, even if it is an utterance of one single English word, he speaks in MAE. The playful part of the language is reserved for Chinese. This choice is motivated by the expectations for a qualified, competent Chinese teacher of English in the Chinese ELT industry, who needs to speak a dominant, prestigious variety of English in order to be considered a legitimate language teacher.

Text, typography, and layout interplay

When Scott self-broadcasts, subtitles in the language Scott is using at that moment are provided. Specifically, when Scott narrates in Chinese or narrates in Chinese with a few English words mentioned, the captions reflect what exactly Scott says. For instance, at the opening of the video when Scott asks a question about choosing the correct preposition for a phrase, the text superimposed on the screen is composed of two languages to reflect the code choices in Scott's speech (see Figure 30). However, when Scotts provide collocation examples in English, both English caption and the translated captions in Chinese are provided (see Figure 32) to help learners understand the meaning of those expressions.
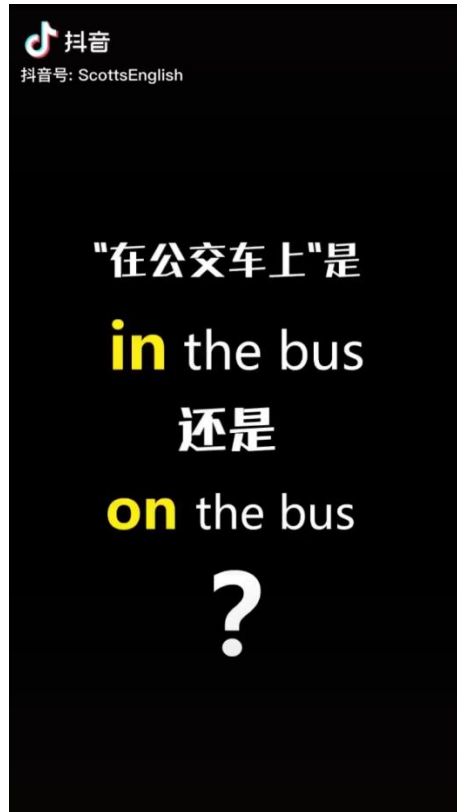
Figure 30. The opening scene of Scott's video of English prepositions

Note. Translation of the caption: Which one is correct? "in the bus" or "on the bus"?

In addition to code choices, the variation in typography and layout of the text also contribute to different degrees of information value and salience. In the opening scene (see Figure 30), the question Scott asks about choosing a correction preposition between "in" and "on" to collocate with "the bus" is presented at the center of the screen to highlight its information value. Meanwhile, the white and yellow colors of the text contrast sharply with the background of the screen, enhancing the salience of the text. Additionally, the audiences' attention is further drawn to the two prepositions of "in" and "on" bolded and highlighted in yellow within the white text.

Figure 31. A cat and a patio in Scott's background

Note. Translation of the caption: "You use 'in' before the object that requires you to sit inside."

In the rest of the video, subtitles are always placed under or near the bottom of the image of Scott (see Figures 31 & 34). When Scott presents and explains grammar rules (e.g., see Figure 31), the corresponding Chinese captions are highlighted, but the English words mentioned stay in white, through which the code choice is emphasized visually. When Scotts give examples in English, both the English subtitles and the translated captions are provided. However, both captions are in white to create a visual rhyme and convey an equivalent level of salience (e.g.,

see Figure 32). The English captions are placed above the Chinese translations to receive a slightly higher prominence.

Figure 32. Three collocation examples with the preposition "in"

While viewers are watching this video on *Douyin*, they will also be able to see the add-on information (see Figure 33). In the video description, Scott marks the video as episode 6, which is one of the video lessons of the collection named "Nail English Grammar in One Go" appearing at the very bottom of the screen. Users will be able to see other lessons in this collection by tapping on the collection link. In addition to labeling the video as episode 6, Scott writes in the video description that he likes using Sichuan-accented Mandarin, which references his Sichuan-accented Mandarin speech in the video lesson. He also includes the hashtag "#英语黑科技" (literally meaning: English black technology) to promote his lesson as useful and effective. "Black technology" (黑科技) is a Chinese online slang term to describe something that is unexpectedly enlightening, futuristic, and of advance science. Furthermore, there is also a

product link: "product:
Getting addicted to
learning English Grammar
like watching American
TV dramas"

Username: @ Scott 学口
语's (Learning Spoken
English with Scott)

Video description
translation: Episode 6: I
just like speaking
Sichuan-accented
Mandarin. What can you
do to me? #cool tech of
English

Time when the video was
posted (only visible when
the video opened from the
video list, not available
when the video appears on
the "For You" feed):
August 5th, 2019

Audio information:
Learning Spoken English
with Scott's original sound

Collection link: "Collection: Nail English Grammar in
One Go"

Figure 33. Add-on information of Scott's video of English prepositions

product link appearing above Scott's account name, which features a short advertising message

about the commodity. In the written message, Scott claims that his lessons will make the learner

become addicted to learning English grammar like watching American TV dramas, appealing to

227

consumers by suggesting his language learning materials are fun. If users tap on the link, they will be directed to the product page where they can purchase Scott's video courses on English grammar. Through the product link superimposed on the video, Scott constructs his identity as an advertiser/entrepreneur is emerged. As such, his video functions like a demo lesson through which Scott demonstrates his language and teaching expertise in order to promote his language teaching products.

Image

When Scott delivers the lesson on English prepositions, the video features him holding the camera and gazing at the lens in a way similar to taking selfies. The close shot shows the head and shoulders of Scott, demanding viewers to engage in an intimate, personal relation with him (Kress & van Leeuwen, 2006). His gaze at the camera lens serves as a direct address to the viewer. Combining with the close shot, the visual address initiated by the gaze invites the audiences into an imagined close relation with the content creator. Meanwhile, the camera is placed at a low angle that makes viewers "look up to" him, denoting a hierarchical relation in which Scott has power over the viewer. The hierarchical relation is legitimatized by the Chinese traditional values that teachers are the authority in the classroom and should be highly revered. Through the low camera angle, Scott indicates his identity as a teacher and an authority on English.

While giving instruction, the background of the video suggests Scott is constantly changing. Instead of remaining in one place, the video shows that Scott is moving around in his home while talking to the camera. As he moves, different features in his home are included and presented to the viewer, which include a cat, a patio (see Figure 32), a pile of storage boxes (see Figure 34), and so forth to reveal a piece of his life, which constructs a sense of authenticity. By

doing so, Scott mixes the genre of a video lesson with a vlog that features content creators and

their life as the main subject. This hybrid genre serves to entertain the viewer and makes the

language lesson more interesting. In the meantime, the audience might have an impression that

Scott is capable of doing improvised teaching when he has other things in his life that he needs to

attend to. In addition, the vlog features embedded in his instructional video also construct the

lesson as informal and help Scott establish intimacy with the viewer.

Figure 34. A pile of storage boxes on the right side of the background in Scott's video

Figure 35. Scott demands viewers to leave likes

Note. Translation of the caption: Give me a like

At the end of the video when Scott appeals to likes by formulating his request like a command. When he utters "Give me a like!" in Sichuan-accented Mandarin, he raises his hand with his palm facing the camera and quickly waves at the camera vertically, which is a Chinese hand gesture meaning "come here" (see Figure 35). This gesture is typically used when summoning someone, especially those who hold an equal or lower position in the relation to the summoner. Together with the hierarchical relation denoted by the low camera angle, Scott constructs himself as an authority, who can legitimately tell the viewers to like his video.

Language and image interplay

Except for the opening scene of the video, an image frame that features Scott is provided in the upper part of the screen, while the rest of the screen features a grey background. However, when Scott utters collocation examples with the preposition "in" and "on", an image is added to the lower section of the screen to illustrate the meaning of the example. For instance, as shown in Figure 32, when Scott provides the example "in a car", the bottom image features a car with four people setting in it. Similarly, when he utters "in a taxi", an image of a woman sitting in a taxi is provided; when he says "in a helicopter", the image in the lower part of the screen shows a helicopter with people sitting inside of it. The upper and lower image on one screen establish a relation of the *idea* and the *real* (Kress & van Leeuwen, 2006). Specifically, the image position on the top that features Scott coincides with the *ideal*, which denotes a higher important value and serves to provide a generalized sense of the information. On the other hand, the image at the bottom provides *real* information that serves to elaborate on the example Scott utters.

Meanwhile, English subtitles and translated captions are juxtaposed with the top and bottom images. The captions are positioned between the two images and overlap with both of them, blurring the boundaries of the image frames. This visual composition combines the text with both images together as a unified piece of information. The reason for this particular visual arrangement is discussed below.

When Scott provides the three collocation examples with the preposition "in", The translated caption in Chinese for each English expression uses "在…里" (to be … inside) instead of "在…上"(to be…on), although both are acceptable in Chinese. The images showing people sitting in the transportation vehicles and the word choice in the translated caption in Chinese are combined to associate "里" (meaning: inside) with "in" and objects that require people to sit

231

inside. In other words, the relation between "in" and transportation means that do not have enough space to allow people to stand inside is not explicitly stated in these scenes (although Scott explicated stated the rule right before giving examples). Rather, this connection is implied through the interaction between the image and the texts. Thus, presenting the text and the information as a coherent unit serves to better help learners perceive the connections and facilitate understanding.

In a similar vein, when Scott provides four different phrases in which the preposition "on" is used to attach to a mode of transportation. The translated caption in Chinese for each English expression uses "在…上" (to be … on). The images at the bottom showing people standing on a mode of transportation and the word choice in the translated caption in Chinese are combined to associate "上" (meaning: on) with "on" and objects that people can stand on.

Sound

Though a bit subtle, there is a short clip of bluegrass music played with a quick tempo in the background throughout the video. Developed in the United States, bluegrass music on *Douyin* is often used for videos with a light-hearted theme. The adoption of bluegrass music for Scott's video spices up his video and adds interest to the content.

Summary of the analysis

In Scott's video of English prepositions, his expert identity of online English language teaching is highlighted through the design and delivery of a highly accessible grammar lesson that addresses a common confusion many Chinese learners have. Specifically, he explains grammar rules in Mandarin Chinese in a concise manner and provides examples in English for illustration. The adoption of a translanguaging teaching method is realized through the use of both Chinese and English in speech and written text. Typography (i.e., font size, color,

background) and layout of the text are utilized to visually convey different degrees of information value, salience, and framing (Kress & van Leeuwen, 2006) in order to guide learners to digest the content in a particular way. In addition, the cohesion of the information is facilitated by the visual composition of texts and images as an integrated unit. Furthermore, visual resources are deployed to create a hybrid genre of a video lesson with a vlog that features the content creator and his life, signaling improvised teaching. By doing so, Scott constructs an expert identity of English language teaching by demonstrating his knowledge of multimodal design, pedagogical knowledge, and linguistic knowledge.

In addition to an English language expert/teacher identity, Scott also performs an authority image by creating a hierarchical relation between him and his viewers. This relation is denoted by a number of multimodal resources including camera angle, gesture, and speech style.

Case Study 4: Jiadecang's English Learning Vlog

In this case study, I focus on one of the videos sampled from the account @嘉德仓自学 English 早 7:50 播 (Jiadecang Self-learning English, Stream Starts at 7:50 a.m.). Although Jiadecang is not among the most popular, or the most followed *Douyin* microcelebrities who offer informal English instruction, he and his videos are unique in three ways. First, Jiadecang is the only Tibetan content producer in my data. In fact, he is the only Chinese ethnic minority I have seen from years of observation who posts videos of himself practicing English on *Douyin*. Second, different from most other microcelebrities of English language teaching content who portray themselves as language expert and teacher, Jiadecang claims a learner identity. Nevertheless, his videos still have an instructional component. Third, I consider his language learning videos as a type of vlog that features him as the vlogger and his life. I label this type of vlog as English learning vlog. In the analysis, I also examine how this genre is constructed

233

through both linguistic and other semiotic resources and how this genre is used for a self-branding purpose. I present my analysis of Jiadecang's use of multimodal resources in one of his English learning vlogs below. The vlog features him cooking outdoors, reading aloud a motivational message and explaining in English what type of food he is cooking.

Image

As a Tibetan person, Jiadecang's ethnicity is not only revealed through his appearance, or in other words, his transportable identity, but is also constructed through a range of visual elements that he integrates into the video. For instance, the video features Tibetan-style residential architecture, which is characterized by flat red roofs, square-shaped overhanging decorations, multiple windows, and an elevated construction site of (see Figure 36). In addition, the video also presents scenes of Jiadecang stewing yak meat in a large pot on a campfire. Yak meat is a staple of Tibetan cuisine and is included in a variety of Tibetan dishes. Chinese people of other ethnicities, however, rarely eat yak meat. Therefore, the scene of Jiadecang stewing yak meat also indexes his Tibetan identity.

Another identity aspect he constructs through images is relevant to rurality. As shown in the opening scene featured in Figure 36, Jiadecang is sitting on the ground in a yard full of dirt and small rocks. Next to him is a large pot placed on a campfire made of wood and stone bricks. Inside the pot are many big chunks of bone-in yak meat in plain water (as shown in Figure 36). The lack of landscaping for the yard and the rustic style of cooking without the use of kitchen

234

Figure 36. The opening scene of Jiadecang's English learning vlog

implements or urban cooking techniques depict the life that Jiadecang lives as rural and wild. In

addition, the tractor in the background also is associated with agriculture and a rural area. Harsh

and raw as the rural conditions portrayed in the video might be, the rural life represented in the

video is nevertheless likely to be attractive to many Chinese online users. As Li (2020) pointed

out, the rapid urbanization and industrialization in the recent decades in China have led to

increased social stratification as well as development-related anxiety. Consequently, city

dwellers have grown a nostalgic longing for a slow-paced, simplistic, and preindustrial

landscapes. Facilitated by mobile devices and social media platforms, rural-content videos

featuring an idyllic countryside lifestyle and an uncontaminated natural setting have gained

growing popularity (Li, 2020). To many netizens, the rural life depicted in the videos offers a

therapeutic effect and a temporary escape from "cramped lifestyles, poor air quality and an

exhausting and mechanistic life replete with growing social disparities" (Li, 2020, p. 3782). As

such, Jiadecang's documentation of his simple, rural life may actually be captivating to many

urban-based online users, especially considering the blue sky featured in the background scene

and his humble mode of living. Compared to the first three videos discussed earlier in this

chapter, Jiadecang's video is relatively simple without much postproduction editing. However, as

Figure 37. Jiadecang stewing yak meat

Li (2020) argued, it is precisely the crude video quality that substantiates one's rural identity and therefore contributes to a sense of genuineness and authenticity. Furthermore, Jiadecang's construction of his rural identity and demonstration of his rural lifestyle can also authenticate his Tibetan identity by relying on the widely perceived discrepancy between Tibet as a remote, underdeveloped region and the east part of China as being populous and modernized.

While Jiadecang's construction of an authentic Tibetan identity embraces and invokes a stereotype about Tibet as a pre-modern, uncivilized region, he simultaneously contests this widely circulating ideology in the Chinese context by portraying a modern and international public image. In terms of visual resources, the video features Jiadecang wearing modern and athletic style clothing (e.g., a winter jacket, a shirt with English letters on it, and sneakers) (see Figure 38). In terms of language, as will be further discussed under the modes of speech and text, Jiadecang's demonstrates his linguistic competence in both Mandarin Chinese and English as a native speaker of the Tibetan language, through which he projects an image of an educated, internationalized Tibetan man.

The public self Jiadecang constructs in the video stands in stark contrast to the negative perception of Tibetans being poverty-stricken and uncivilized. However, the audience's appreciation of this contrast hinges on the assumption that the majority of people in Tibet live a pre-modern life, which is also validated by the rurality demonstrated in the video. The image of an educated, modernized Tibetan person help Jiadecang establish a unique, unconventional identity on *Douyin* to attract followers. At the same time, however, his self-branding practice that appropriates China-Tibet differences can reproduce and reinforce stereotypes about Tibet and Tibetan people.

In addition, his ethnic identity, when combined with his multilingual competence, invokes a hard-working persona. As will be discussed in this and later sections, Jiadecang draws on various semiotic resources to frame his video as a learning vlog featuring him practicing English by himself. Having the impression that Jiadecang has been teaching himself English in a rural area that lacks educational resources and that he manages to speak fairly proficient English makes the viewer appreciate him as a hard-working and resourceful learner.

To frame his video as a vlog using visual resources, Jiadecang presents the video content through a dynamicized distance (Kress & van Leeuwen, 2006). In Jiadecang's video, long shots are used at the beginning. Then the camera gradually zooms into a close-up shot. This camera-initiated change in distance and angle communicates a neutral stance of recording what is taking place like a documentary and thus frames the video as an objective, authentic representation of Jiadecang's mundane life.

Near the end of the video, Jiadecang is shown in medium shots, which cut off the subject approximately at the knees. Together with the high angle of the camera, it conveys a sense of closeness and respect for the audience. The direct eye contact Jiadecang is making with the viewer invites the audience into an imaginary relation with him. Exactly what kind of relation is then signified by the other semiotic resources, including his smile, winking, and tongue clicking, through which Jiadecang further establishes intimacy with the viewer in order to appeal to his following.

Figure 38. Jiadecang smiling and winking

Sound

     In addition to visual elements, Jiadecang also harnesses auditory resources to index his ethnic identity. Specifically, Jiadecang chooses a Tibetan folk song as the background music playing throughout the video. Although the viewer might not understand in what language the lyric is, the sound is produced by a Dramyin, a traditional Tibetan plucked string instrument that is relatively easier for many Chinese viewers to identify.

Speech

In terms of linguistic code, Jiadecang's speech in the video is in English the entire time. He first narrates in English to share his thoughts about life "Life is a kind of enjoyment, so don't complain about anything." By sending this motivational message to the viewer, Jiadecang position himself as a mentor and constructs a positive, resilient personality. When the video shows close-up shots of big chunks of bone-in yak meat covered by water in a large pot, Jiadecang again uses English to explain that he is making a yew stew by saying "cook the yak meat." This expression might sound slightly unnatural to proficient users of English, and they might interpret it as an imperative. The definite article "the" inserted before the expression "yak meat" is redundant as well. In fact, what Jiadecang intends to express is something like "I am cooking yak meat" or "I am making yak stew."

Although Jiadecang's English is non-standard, he is protected from potential critiques about his language production. Since Jiadecang draws on a variety of modal resources to identify himself as a language learner who uses short videos to document his journey of learning the English language, this relational identity (Buchotlz & Hall, 2010) exonerates him from the potential criticism of providing inaccurate information which a language teacher/expert would receive. On the other hand, for non-proficient users of English who wish to improve, many may still find Jiadecang's learning vlog useful in helping them learn new vocabulary words, practice speaking, or increase their language learning motivation. From this perspective, Jiadecang also serves the role of a knowledge disseminator, or an expert in relation to these learners.

Text, typography, and layout interplay

In addition to visual resources, Jiadecang employs textual information to construct his video as an English learning vlog as well. As shown in screenshots included in Figures 36-38,

there is a line of Chinese text that stays on the top of the screen throughout the video: "自学英语 第 959 天", meaning "the 959th day of learning English by myself". What Jiadecang indicates is that this video is the 959th video he has made featuring him practicing English and he has been learning English for at least 959 days. Thus, the discourse identity Jiadecang establishes through the text is that of an English language learner. In addition, Yang's persistence in learning English contributes to his personal identity as a motivated and determined person. Furthermore, Jiadecang demonstrates his language proficiency in Chinese by adding Chinese text to the video. Most Tibetan Chinese people grow up speaking Tibetan as their first language at home. Those who can afford education will learn Mandarin Chinese as their second language at school. By adopting Chinese text, Jiadecang constructs himself as an educated Tibetan. Moreover, this code choice also reflects the main group of the audience he is targeting, which are those users who speak and write in Chinese. From a cultural perspective, although Jiadecang is ethnically and culturally Tibetan, he establishes group affiliation with other Chinese online users through his code choice.

When it comes to subtitles near the bottom of the screen, although Jiadecang narrates only in English, both English captions and translated captions in Chinese are provided (see Figure 37). The English captions are placed above the translated captions in Chinese to indicate a higher level of relevance to Jiadecang's speech. Same as the Chinese text superimposed on the top of the screen, both English and Chinese subtitles are in yellow to create a visual rhyme in all textual information. Providing subtitles in both languages also allows Jiadecang to frame his video as instructional. By doing so, he is no longer just a learner, but also an expert to a certain extent who has the authority with respect to the language content.

In addition to the textual information embedded in the video, Jiadecang composes a description for his video as part of the add-on information that viewers can read when they watch the video on Douyin. In the video description section at the bottom of the screen (see Figure 39), Jiadecang first puts the Chinese translation of the motivational English message ("life is a kind of enjoyment, so don't complain about anything) he shares with the viewers in the video. Following the quote he adds two hashtags. The first one is #记录美好生活 (Recording wonderful moments), which is a trendy hashtag on *Douyin* that can help promote the video for more views. The second hashtag is #嘉德仓吃货 (Jiadecang the foodie), which has an intertextual reference to the video content featuring Jiadecang cooking yak meat. The label "foodie" Jiadecang assigns to himself is also what many people relate themselves to, through which Jiadecang constructs an ordinary persona and conveys a sense of intimacy. There are also two *Douyin* accounts mentioned after the hashtags. The first account is @嘉爷 (Jia Ye, meaning "master Jia"), which is Jiadecang's other *Douyin* account[44]. The Chinese word "爷" (Ye) in the account name is an honorific to address a man who holds a higher social status for his age, wealth, or power. In the Chinese online context, this honorific expression is adopted by some users to either address themselves to convey masculinity or to address others in a respectful, manly way. Compared to the account included in my data, his second *Douyin* account does not have a focus on English language teaching/learning. Rather, it functions more like a "supplementary" account for Jiadecang to provide his followers with additional information about him such as daily activities, through which Jiadecang becomes more accessible to the

---

[44] The fact that Jiadecang has a different *Douyin* account is contradictory to what he claimed in his bio on the profile page that he does not another *Douyin* account. This might be due to his delay in updating his bio information. Another possible reason might be that he does not want viewers to misrecognize other *Douyin* accounts with similar account names as his.

Username: @嘉德仓自学
English 早 7:50 播
(Jiadecang Self-learning
English, Stream Starts at
7:50 a.m.)

Video description
translation: Enjoy your life
and stop complaining.
#Recording wonderful
moments #Jiadecang the
foodie @Jia Ye @Dou+
marketing tool

Time when the video was
posted (only visible when
the video opened from the
video list, not available
when the video appears on
the "For You" feed): 18
hours ago

Audio information:
Jiadecang Self-learning
English, Stream Starts at
7:50 a.m.'s original sound

Figure 39. Add-on information of Jiadecang's English learning vlog

viewer. The next *Douyin* account included in the description is @Dou+ marketing tool, which is

automatically generated and added to the description of a video that has used the paid service on

Douyin to increase views and gain followers. Although it was not Jiadecang's choice to include

this information, it nevertheless shows that as a content creator on *Douyin*, Jiadecang is very conscious about attracting followers and maintaining his microcelebrity status.

Text and image interplay

Considering the layout of the Chinese text "自学英语第 959 天" (the 959th day of learning English by myself), its position on the top of the screen makes it the *ideal* information, appealing to viewers' emotions and providing a generalized essence of the information. The perception of Jiadecang as a Tibetan person living in an underdeveloped region that lacks learning resources serves to increase the emotional appeal of the text. As such, audiences are more likely to appreciate Jiadecang as a hardworking individual. On the other hand, the lower section of the screen, as opposed to the Chinese text superimposed on the top, visualizes the *real* information that specifies what Jiadecang learning English by himself looks like. In a top-bottom spatial organization that is arranged along what is being perceived as *ideal* vs. *real*, the *ideal* information carries more salience while the *real* information plays a subservient role (Kress & van Leeuwen, 2006). The typography in Jiadecang's video is designed to align with the degree of salience of the text in relation to the *ideal* and the *real* as well. The Chinese text remaining on the top of the screen is made considerably larger than the English and Chinese captions on the bottom part of the screen, through which viewers are consistently reminded of Jiadecang's learner identity and his persistence constructed through the top text.

Summary of the analysis

My analysis has revealed that Jiadecang's multimodal construction of English learning vlog as a hybrid genre enables him to engage in a hybrid form of talk as both an English expert and a learner. The instructional aspect of the video is realized through the juxtaposition of both English and Chinee subtitles on the video. Performing an amateur-expert identity helps him

promote his English language expertise, and at the same time, protects him from potential critiques about his language production. In addition, a fellow language learner identity allows the content creator to establish membership and connections with his followers.

Other main identity roles Jiadeceang projects include an educated and international Tibetan Chinese, a rural person, and a friend. Being a Tibetan person is Jiadecang's transportable identity. However, this identity is also addressed through the use of various visual and auditory resources (e.g., Tibetan-style architecture, Tibetan food, and Tibetan folklore) in order to create a commodified public image, which is exploited to promote the content creator's products. Meanwhile, Jiadecang's Chinese identity is demonstrated through his adoption of Mandarin Chinese in the written text to establish affiliations with *Douyin* viewers. Moreover, the content producer's rural identity indexed in visual elements serves to authenticate his Tibetan identity by invoking the stereotype of Tibet as an underdeveloped area. Similarly, it is mainly through visual elements, such as gesture and camera angle that Jiadecang positions himself as a friend to convey a sense of closeness.

CHAPTER SEVEN: DISCUSSION AND CONCLUSION

The aim of the dissertation was to understand the self-branding and language teaching practices of *Douyin* microcelebrities of informal English language instruction, which was motivated by three major factors. First, the study wanted to improve our knowledge of multimodal digital discourse practices, in general, and multimodal construction of short videos on *Douyin*, specifically. Second, the project sought to understand microcelebrities' self-branding practices and the related notion of authenticity in Chinese digital and socio-cultural contexts. Third, the research aimed to inform the current and future online teaching and advertising practices. In this chapter, I first summarize and discuss the findings of the study. I will then present the research implications for and contributions to digital communication and ELT field, followed by an overview of the study limitations. I will end this chapter with a discussion of future research directions.

Summary and Discussion

In Chapter 4, I illustrated various multimodal resources the 14 selected *Douyin* microcelebrities of informal English instructional videos used for self-presentation and product promotion on their profile pages. Similar to what has been discussed in some of the prior studies on microcelebrity strategies (e.g., Jerslve, 2016; Raun, 2018), most of these 14 content producers communicate varying degrees of connectedness, accessibility, and availability by sharing their other social media accounts, listing their live stream schedules, and encouraging followers to join fan groups. The cultivation of microcelebrities is found to emphasize the interaction and connections between vloggers and their audiences. Consistent with Bhatia's (2022) analysis of

246

the identity construction of YouTube make-up artists, the *Douyin* content producers of informal English instruction videos in my sample construct multiple identity roles, each of which serves a specific communicative goal, such as establishing credibility, communicating authenticity, and socializing.

The diversity of the content creators in race and culture leads to different labor in identity construction. While Chinese content creators prioritize demonstrating a professional aspect of public identity in order to be considered as a legitimate teacher of English, vloggers who are Caucasian emphasizes more on constructing a personal aspect of their image as well as building cultural affiliations with *Douyin* viewers.

It is also found that various kinds of non-standard written practices such as using emojis, homophones, and combination of symbol and English letters, are adopted to evade *Douyin*'s commercial censoring of its competitors. Previous studies on Chinese online users' practices of bypassing censorship mostly focus on politics (e.g., Jiang & Vásquez, 2020; Rea, 2013; Wang, 2012). The present study, however, discovered a case of evading commercial censorship in the Chinese digital context.

In Chapter 5, I presented the content topics discussed in the 84 videos sampled from the 14 selected accounts, as well as the language skills addressed, and the teaching methods used in those videos. Then I performed a cross-case analysis to look for patterns. It was found that the language components and skills taught in the videos were closely associated with the platform affordances, language learners' needs, and the content producer's motivation to generate interactions and discussions with the viewers. Specifically, speaking is the language skill that is most frequently taught and discussed, and vocabulary the second. There are only a few videos address grammar, reading, and writing skills due to the constrains of *Douyin* platform

affordances and viewers' short attention span. Although videos that center primarily on listening skills are limited in number, videos of speaking skills often implicate listening.

In terms of teaching styles and methods, although lecture appeared to be the most frequently used style, reading aloud, role-play, and real-life conversation altogether are implemented in more than half of the pedagogical videos. This might be an indication of learners' preference for teaching materials they can follow closely and for learning English in a context.

The content analysis of this chapter also revealed that videos that do not directly focus on English language teaching are more frequently pinned than videos that do so. Most of the pinned videos feature the content creators showing off their linguistic competence but without a pedagogical orientation, which might indicate the perception that how well one can speak English outweighs the importance of their pedagogical knowledge and professional credentials.

In Chapter 6, I elaborated on the linguistic and other semiotic resources that four different microcelebrities marshal to construct public identities through delivering English language instruction in the four selected videos. My detailed multimodal analyses of the videos revealed that identity constructions come about as a result of complex interactions between different modal resources that generate meaning together as multimodal ensembles (Jewitt, 2009, 2013; Kress, 2010). As the "designers" of the multimodal text, *Douyin* microcelebrities of informal English instruction videos selected resources with an awareness of the specific context of meaning making and an understanding of the platform affordance for their videos in order to meet their *interest* (Kress, 2010), or their self-branding purpose. Meanings associated with the signs that appear in the videos are not pre-assigned. Rather, they are generated and negotiated in and through interactions. As such, the microcelebrities' identities emerge from the multimodal

meaning-making practices rather than being preestablished. For instance, in Yang's video of airport English, when he spoke MAE as a U.S. immigration officer with a blond wig, he invoked a native speaker image and indicated his ability to pass as a native speaker.

As for authenticity as a microcelebrity, it was found that not all the four microcelebrities of the four selected videos for a case study strived to be perceived as "real" in their videos. If the Tibetan vlogger, Jiadecang, exhibited most authenticity among the four Douyiners, then Yang was perhaps the least. Different from Jiadecang who featured his mundane life in his video, Yang performed staged, comical, exaggerated identities to generate humor in a highly entertaining and highly edited video in order to attract followers. However, when Jiadecang incorporated all kinds of linguistic (i.e., mentioning of yak meat), auditory (i.e., using a Tibetan folklore as the background music), and visual resources (e.g., Tibetan style residential architecture, a large pot of yak meat) to index his Tibetan ethnicity within the frame of business transaction (e.g., including his live stream schedule in his account name and bio, selling Tibetan specialties on his *Douyin* personal store), his authenticity became commodified.

Similar to Jiadecang, Scott also performed a commodified authenticity in his video. Specifically, when he was giving English instruction, the video featured him moving around in his home, during which he disclosed different things and features of his personal life through the background. Although this hybrid genre of lecture and vlog added to his authenticity, his authenticity is commodified because of the product link superimposed on the video.

In terms of authenticity in ELT, the ideologies of native speakerism as well as standard English were evident on *Douyin*. Although the standard language ideology was embraced by many of the content creators of English language instruction, their stance toward native speakerism is perhaps more conflicted and ambivalent. Take Yang's video of airport English as

an example. Yang's Mock Chinese practice was used to playfully portray an image of a learner who spoke Chinese-accented English. The mocking frame allowed Yang to disauthenticate his Mock Chinese speech and therefore distanced himself from the performed persona. Meanwhile, he authenticated his own MAE by playing the role of an U.S. immigration officer who had blond hair, which was a visual index of a stereotypical white native speaker of English. What Yang indicated is that his English proficiency was native-like, and he could pass as a native speaker if he were Caucasian. In so doing he contested the ideology of native speakerism. However, at the same time he is reproducing the ideology when he speaks MAE and wears a blond wig.

When it comes to the design of the language teaching content, the analyses demonstrated that microcelebrities deployed a wide range of modes to frame the video as pedagogical, to structure the lessons into discrete sections, to signal the different degrees of importance of information, to direct leaners' attention, as well as to guide viewers to read and digest the content in particular ways. One feature that has been widely adopted in *Douyin* informal English instruction videos is subtitles. When content creators speak in English, there is almost always both English captions and translated captions in Chinese. The layout and typography of the subtitles vary depending on the information value, prominence, and framing (Kress & van Leeuwen, 2006) the content creator intended to convey for a pedagogical purpose.

<div align="center">

Implications and Contributions

</div>

By exploring the various resources used by *Douyin* content creators to craft instructional English videos for a self-branding purpose, this study generates several implications.

First, the findings of this research suggest that a consistent identity representation plays an important role in attracting followers on *Douyin*. Content creators are expected to perform in their videos what they claim to be on their profile pages and vice versa. My analysis has shown

that the 14 microcelebrities in my sample all kept a consistency between the identities they constructed in their videos and the public image they established through profile page. Due to the different affordances of *Douyin* profile pages and videos, identity construction on the profile pages was generally more concise and straightforward through written text, while videos served as the elaboration of the same identity roles. For instance, Yang on his profile page used "English", "music", and "drama king" to label himself in a catchy and appealing way. These words may give viewers a general idea of who Yang is. However, it is only until viewers watched his videos (featuring him teaching English, singing Chinese and English songs, and performing exaggerated persona) that they develop a better understanding of the three labels and the image Yang projected. Likewise, when viewers finished watching a video of Yang on their *Douyin* feed and wish to know more about the content creator, the information on the profile page give the viewers an overall understanding of who the vlogger is and what types of content are featured in his videos. An inconsistency in identity construction in between profile pages and videos can cause confusion to the audience. When viewers are unable to tell whether a content creator is consistently projecting a public image and producing content they like, it is unlikely that they will follow the account.

Second, a unique hybrid identity can serve to enhance *Douyin* content creators' competitiveness in the attention economy. The construction of a unique public image is often associated with prevailing racial, linguistic, and ethnic stereotypes. Furthermore, the racial, linguistic, or ethnic ideology microcelebrities invoke may also contribute to language commodification. For instance, in Jiadecang's English learning vlog, he constructed an authentic Tibetan identity through reference to the stereotype about Tibet as a pre-industrialized region. However, by portraying himself as a modernized Tibetan through the use of the English

language, he projected a unique hybrid identity to attract attention. As a result, his English language added extra value to his public persona and thus played a role in his self-commodification. Moreover, when viewers purchase local specialties listed in his *Douyin* store, what is actually helping to market these products is Jiadecang's public character. In this sense, Jiadecang's English language also add value to the products he sold, albeit in an indirect way.

Third, an effective self-promotion requires socialization and engagement with both audiences and other content creators. The collaboration between content creators and the co-production of videos helps them grow their fanbase through receiving and offering the same audience to fellow influencers. Meanwhile, the interactions between content creators allow them to establish social groups that impose standards on membership requirement, through which they establish authority and credibility.

Next, the findings also suggest that language pedagogies need to be developed with a serious consideration of power redistribution and learner's agency. As Kress (2010) pointed out, learning has transcended the boundaries of institutions and take place regardless of time, location, and format. As such, pedagogies that are exclusively dependent on traditional authority held by experts and institutions might not be able to meet learners' social, personal, and affective needs. Knowledge is acquired through receiving information from authority. However, in the digital era, knowledge can also be disseminated and acquired through individual's content creations and reproductions. For instance, despite the imperfect English expressions Jiadecang provides in his English learning vlogs, many viewers may still find his videos useful in helping them learn new vocabulary words, practice speaking, or increase their language learning motivation. On the other hand, as the content creator, his agency in relation to the production of

knowledge allows him to construct identities for a self-branding purpose and sustain his motivation in language learning.

Digital technologies facilitate new emerging forms of professional communication. There is a blurring of boundaries between traditional genres in digital spaces (Bhatia, 2022; Kress, 2010). The on-going negotiation between professionalism and amateurship in self-representation can be realized through genre, which is an expression of the social relation (Kress, 2010). The construction of hybrid genres on *Douyin* reflects a narrowing divide between experts and novices (Selwyn 2015), which leads to pedagogic democratization (Kellner & Kim 2010), and the remaking of power relations in knowledge production and information transmission.

In terms of designing effective informal English instruction videos, translanguaging as a pedagogy is potentially powerful in creating an inclusive learning environment for language learners with varying levels of proficiency. For example, by providing subtitles in both leaners' first language and the target language, leaners may find it easier to keep up with the instructional content. The interactions between written text, typography, and layout were shown to play an essential role in framing a video as instructional and directing learners' attention. Furthermore, the integration of different genres into a video can not only structure the lesson, but also entertain the learners to keep then engaged.

In addition, the findings suggest that many microcelebrities tended to prioritize their language competence over pedagogical knowledge and professional credentials when constructing a language expert/teacher identity on *Douyin*, it might be a sign of deprofessionalization of language teachers motivated by social media affordances and Chinese netizens' perception of a legitimate English teacher.

Furthermore, the study uncovers the uneven distribution of language resources in *Douyin* informal English instruction videos to different audiences. For instance, Yang in his video of airport English did not just teach English as language in general. What he was actually teaching is a certain kind of English that is bureaucratically useful and is a prestigious variety. As such, Yang was targeting at specific group of learners—learners who already know some English, but need some specific set of "linguistic, pragmatic and metapragmatic skills" to upscale in a globalized age (Blommaert, 2010, p. 36).

Finally, it is worth noting that the successful application of Kress and van Leeuwen's (2006, 2021) grammar of visual design in this study shows that despite traditionally being used to analyze Western visual composition, this framework is, in general, applicable to *Douyin* content.

The findings of the study uncover the widespread nature of self-branding strategies and online identity construction practices used by online vloggers that appeal especially to Chinese consumers. The detailed multimodal analysis of the videos contributes to our understanding of multimodal digital discourse practices and multimodal construction of short videos.

In addition, this research can help language educators identify strategies for creating engaging online learning /advertising content. Some of these strategies include providing subtitles in both learners' L1 and the target language, generating multimodal humor, mixing different genres, creating a personalized experience, and establishing affinity with the consideration of the viewer's short attention span. Furthermore, this research adds to the growing body of research on computer-assisted language learning and also, to our understanding of authenticity in digital self-branding, a topic of interest in marketing and media studies as well as in digital text analysis. Lastly, this study sheds light on the broader media landscape and user-

generated participatory culture in China, which then increases our understanding of contemporary life.

Limitations

This research, however, is subject to several limitations. The section will address potential concerns about data transcription, the interpretation process, and the presentation of multimodal features in writing.

First, although I adopted a multimodal transcription method to capture how various modal resources generate meaning jointly in the selected videos as multimodal ensembles, the transcription does not detail content creators' voicing and prosody relevant to identity construction. For example, in the multimodal transcription of Yang's video of airport English, I did not include phonetic and phonological features that Yang deploys in his Mock Asian speech. The main reason for this decision is that including a more in-depth transcription of speech would take too much space and leads to the spatial disconnection between the presentation of different modal resources used in one frame. This disconnection is undesirable for a multimodal analysis as it makes the interplay between modes harder to see. In the meantime, however, a transcription of the linguistic features in Yang's Mock Asian performance is essential in investigating his identity construction practices and self-branding strategies. Therefore, I supplemented the multimodal transcription with additional transcripts that attends to detailed prosodic and voicing features, and I presented those supplementary transcripts in the body text when discussing Yang's speech. In doing so I provided a more in-depth analysis of language features in the video, which is what Bhatia (2022) calls for when analyzing online videos. On the other hand, my supplementary transcripts are less in their capacity to demonstrate the interactions between modal resources. For instance, Yang's comic performance of a stereotypical image of an illegal

Chinese immigration to the States is characterized not only by the linguistic features in his speech, but also through visual elements such as clothing, accessories, and body language. Therefore, I needed to constantly go back and forth between the multimodal and the additional transcripts in order to make sure that I had dual focuses on in-depth analysis of the speech as well as the interplay between different modal resources. It is likely that during the process there were other important features I should have attended to but failed to capture.

Another point of consideration is the interpretation of data. There are numerous semiotic resources interacting in all kinds of ways to jointly represent and communicate meaning. It is impossible to give even attention to all modal resources and to identify all possible ways modes interact. My selections of modes and the semiotic resources realizing those modes, as well as the interplay between modes for a close investigation is primarily based on my interpretation of their relevance to online identity construction and language teaching practices. My analysis was guided by a social semiotic approach to multimodality and identity frameworks informed by discourse analysis and sociolinguistics. However, my interpretation of data lacks an emic perspective from the content creators of informal English instruction videos on *Douyin* as well as their fans. Microcelebrities' interactions with followers go beyond the profile page and videos, such as comments and replies under videos, interactions during live streaming, in *Douyin* fan groups, and on other platforms such as *WeChat*. However, as a detached, non-obtrusive observer I did not engage in any forms of interpersonal communications with content producers or other audiences. Thus, my interpretation of the data may be less social than it is to the core participants, including the designer of the multimodal text and the viewers who regularly interact with the vlogger (Recktenwald, 2018). For example, I may not be able to tell whether a video is posted to address questions learners raised during live stream or in the comment section of a

previous video, or how might a content creator's selection of semiotic resources be influenced by their interactions with core followers.

Finally, the last limitation of the study is related to the presentation of multimodal features in writing. Simultaneously afforded and constrained by the genre of academic writing, the presentation of my written analysis is essentially linear, one word, one sentence following another, realized as a sequence in time. On the contrary, videos are based on both the logic of time and space. Therefore, the connections of modal resources in videos are both temporal and spatial. Because writing is linear, the conundrum I faced was how I should structure my analysis. I originally started with the logic of time and tried to present a detailed multimodal analysis based on video frames, following how I transcribed the videos for their multimodal features. However, soon I found writing based on fames in a time sequence is chaotic and repetitive. Most importantly, it did not provide a clear, systematic account of the role each mode plays to contribute to the overall meaning. Aware of this limitation, I then decided to use modal resources and types of interplay between different modes as units to present analysis. Although doing so yielded a much clearer prose, I should point out that I have no intention to suggest the modal resources I examined as separate sections (e.g., speech, image) in the four case studies generate meaning independently (even if they might appear to be) without the need to consider their interactions with other modes. Rather, it is a compromise between the linearity of academic writing and the combined logics of time and space in videos.

Directions for Future Research

In the last section of this dissertation, I close the discussion of the current study by proposing three directions for future research.

The first direction concerns learners' agency in using short video applications. The fast-evolving digital technologies have provided language learners with tremendous opportunities to access language learning resources anywhere and anytime they like. The blurring distinction between everyday life and language learning in the modern language learning landscape has created favorable conditions for learners to engage in self-directed learning. As such, the agency of learners demands more serious attention from both a social and pedagogical perspective (Kress, 2010). Since my dissertation explores the content creation side, in my future work, I envision investigating the "consumer" side—in other words, how do users interact with this English language learning content and other users. Specifically, I will examine learners' selection and evaluation of learning resources, identity construction, and power negotiation in relation to knowledge production. This type of research can contribute to a better understanding of online users' self-directed learning practices and how learners' agency is mediated by technological affordances, as well as by personal and social-cultural factors.

The second research direction seeks to explore the discourse of online live streaming of English classes. The rapid expansion of China's live streaming market and the tremendous growth in its revenue have led to a growing number of brands and individuals using this new media genre to attract potential customers. Following this trend, many content creators of informal English language teaching videos on *Douyin* actively utilize the live streaming feature embedded in the applications to interact with viewers, during with they deliver demo language lessons, promote language learning products (e.g., books, video lessons, live lectures), and receive fan donations. It would be interesting and necessary to explore streamer-viewer interactions during the live streams of English classes. For instance, how do streamers use language and other semiotic resources, as well as technology to keep language learners engaged

and attached to their channels, as well as to deliver instruction, promote sales, and encourage donations? How do viewers interact with streamers and with other viewers for a variety of purposes, such as for seeking clarifications and establishing personal bonds? Studies addressing these questions will have potential to add to a growing body of research addressing digital advertising and self-promotional discourse, internet pragmatics, and technology-assisted language teaching/learning.

Finally, future projects could also explore the perceptions of effective informal instruction content from the perspective of language professionals as well as content creators/streamer. Together with the future findings of the two research directions described above, it might be possible to propose some recommendations or guidelines about how online informal language instruction content can be evaluated.

REFERENCES

A new target market for Chinese short-video platforms: the "Silver generation". (2020, June).

    Marketing to China. Retrieved from https://marketingtochina.com/a-new-target-market-
for-chinese-short-video-platforms-the-silver-generation/

Abidin, C. (2016), 'Aren't these just young, rich women doing vain things online?': Influencer

    selfies as subversive frivolity, *Social Media + Society*, *2*(2), 1–17.

Adami, E. (2018). Styling the self online: Semiotic technologization in weblog publishing. *Social*

    Semiotics, *28*(5), 601–622. https://doi.org/10.1080/10350330.2018.1504713

American Educational Research Association (2011. February). *Code of Ethics*. Retrieved

    November 30, 2020 from

    http://www.aera.net/Portals/38/docs/About_AERA/CodeOfEthics(1).pdf

Arndt, H. L., & Woore, R. (2018). Vocabulary learning from watching YouTube videos and

    reading blog posts. *Language Learning*, 22(1). 124–142. https://doi.org/10125/44660

Association of Internet Researchers. (2019, October 6). Internet Research: Ethical Guidelines

    3.0. *Ethics*. https://aoir.org/ethics/

Benson, P. (2015a). Commenting to learn: Evidence of language and intercultural learning in

    comments on YouTube videos. *Language Learning & Technology*, *19*(3), 88–105.

Benson, P. (2015b). YouTube as text: Spoken interaction analysis and digital discourse. In R. H.

    Jones, A. Chik, & C. A. Hafner (Eds.), *Discourse and digital practices: Doing discourse*

    *analysis in the digital ag*e (pp. 81-96). Taylor and Francis Group.

Benson, P. (2006). Autonomy in language teaching and learning. *Language Teaching*, *40*(1), 21–40. https://doi:10.1017/S0261444806003958.

Benson, P., & Chan, N. (2010). *TESOL* after YouTube: Fansubbing and informal language learning. *Taiwan Journal of TESOL*, *7*(2), 1–23.

Berryman, R., & Kavka, M. (2017). 'I guess a lot of people see me as a big sister or a friend': The role of intimacy in the celebrification of beauty vloggers. *Journal of Gender Studies*, *26*(3), 307–320. https://doi.org/10.1080/09589236.2017.1288611

Bezemer, J., & Jewitt, C. (2010). Multimodal analysis: Key issues. In L. Litosseliti. (Ed.), *Research methods in linguistics* (pp. 180–197). Continuum.

Bhatia, A. (2022). Analyzing online videos. In C. Vásquez (Ed.), *Research Methods for Digital Discourse Analysis* (pp. 177-196). Bloomsbury.

Block, D., & Corona, V. (2016). Intersectionality in language and identity research. In S. Preece (Ed.), *The Routledge handbook of language and identity* (pp. 507–522). Routledge.

Blommaert, J. (2010). *The sociolinguistics of globalization*. Cambridge.

Blommaert, J., & Rampton, B. (2011). Language and superdiversity. *Diversities*, *13*(2), 1–21.

Blommaert, J, Ying Lu, Y., & Li, K. (2019). "From the self to the selfie." *Tilburg Papers in Culture Studies*, 222. https://www.tilburguniversity.edu/sites/tiu/files/download /TPCS_222_Blommaert-Lu-Li.pdf (accessed 5 May 2019).

Bolander, B., & Locher, M. A. (2014). Doing sociolinguistic research on computer-mediated data: A review of four methodological issues. *Discourse Context and Media*, *3*, 14–26.

Bourdieu, P. (1991). *Language and symbolic power*. Polity Press & Basil Blackwell.

Braine, G. (1999). *Non-Native Educators in English Language Teaching*. Mahwah, NJ: Erlbaum.

Braine, G. (2010). *Nonnative speaker English teachers: Research, pedagogy, and professional growth*. Routledge.

Bucholtz, M. (2003). Sociolinguistic nostalgia and the authentication of identity', *Journal of Sociolinguistics*, *7*(3), 398–416.

Bucholtz, M., & Hall, K. (2004). Language and identity. In A. Duranti (Ed.), *A Companion to Linguistic Anthropology* (pp. 369–394). Blackwell Pub.

Bucholtz, M., & Hall, K. (2010). Locating identity in language. In C. Llamas & D. Watt (Eds.), *Language and identities* (pp. 18–28). Edinburgh University Press.

Buendgens-Kosten, J. (2014). Authenticity. *ELT Journal*, 68(4), 457–549.

Cai, Y. (2017). *Universities offer lifelong learning to China's elderly. Sixth Tone: Fresh voices from today's China*. Retrieved from http://www.sixthtone.com/news/1840/universities-offer-lifelong-learning-to-chinas-elderly

Cheng, L. (2008). The key to success: English language testing in China. *Language Testing*, 25(1), 15–37.

Chew, M. M. (2020). Discovering the digital Stephen Chow: The transborder influence of Chow's films on the Chinese Internet in the 2010s. *Global Media and China*, *5*(2), 124–137. https://doi.org/10.1177/2059436420928058

Choi, L. J. (2016). Revisiting the issue of native speakerism: 'I don't want to speak like a native speaker of English.' *Language and Education*, *30*(1), 72–85. https://doi.org/10.1080/09500782.2015.1089887

Chik, A., & Ho, J. (2017). Learn a language for free: Recreational learning among adults. *System*, *69*, 162–171. https://doi.org/10.1016/j.system.2017.07.017

Chinese Internet Network Information Center. (2020, April). *Statistical report on internet development in China*. Retrieved November 30, 2020 from https://cnnic.com.cn/IDR/ReportDownloads/202008/P020200827549953874912.pdf

Cook, V. (2013). The language of the street. *Applied Linguistics Review*, *4*, 43–81.

Creese, A., & Blackledge, A. (2015). Translanguaging and identity in educational settings. *Annual Review of Applied Linguistics*, *35*, 20–35. https://doi.org/10.1017/S0267190514000233

Chun, E. W. (2004). Ideologies of legitimate mockery: Margaret Cho's revoicings of mock Asian. *Pragmatics. Quarterly Publication of the International Pragmatics Association (IPrA)*, *14*(2–3), 263–289. https://doi.org/10.1075/prag.14.2-3.10chu

Chung, H. S. (2021). From I'm the One That I Want to Kim's Convenience: The paradoxes and perils of implicit in-group "yellowvoicing." In G. Planchenault & L. Poljak (Eds.), *Pragmatics & Beyond New Series* (Vol. 327, pp. 189–204). John Benjamins Publishing Company. https://doi.org/10.1075/pbns.327.08chu

Darvin, R. (2016). Language and identity in the digital age. In S. Preece (Ed.), *The Routledge handbook of language and identity* (pp. 507–522). Routledge.

De Veirman, M., Cauberghe, V., & Hudders, L. (2016), Marketing through Instagram influencers: impact of number of followers and product divergence on brand attitude, *International Journal of Advertising*, *36*(5), 798–828.

Dervić, M., & Bećirović, S. (2019). Native and non-native EFL teachers dichotomy: Terminological, competitiveness and employment discrimination. *Journal of Language and Education*, *5*(3), 114–127. https://doi.org/10.17323/jle.2019.9746

Duffy, B. E., & Hund, E. (2019). *Gendered visibility on social media: Navigating instagram's authenticity bind*. 13. 4983–5002.

263

Duguay, S. (2019). "Running the numbers": Modes of microcelebrity labor in queer women's self-representation on Instagram and Vine. *Social Media + Society*, *5*(4), 1–11. https://doi.org/10.1177/2056305119894002

Fairchild, C. (2007). *Building the authentic celebrity: The 'idol' phenomenon in the attention economy*. *Popular Music and Society*, *30*(3), 355–375.

Fairclough, N. (2001). *Language and Power*. Routledge.

Gao, X. (2012). The study of English in China as a patriotic enterprise. *World Englishes*, *31*(3), 351–365.

García, O., & A. M. Y. Lin. (2017). Translanguaging in bilingual education. In A. L. García & S. May (Eds.), *Bilingual and Multilingual Education* (pp. 117–130). Springer.

Garcia, O., & Li, W. (2014). *Translanguaging: Language, bilingualism and education*. Palgrave Macmillan.

Godwin-Jones, R. (2018). Chasing the butterfly effect: Informal language learning online as a complex system. *Language Learning & Technology*, *22*(2), 8–27. https://doi.org/10125/44643

Hall, S. (1996). Introduction: who needs 'identity'?, In S. Hall and P. du Gay (Eds.), *Questions of cultural identity* (pp. 1–17). SAGE.

Heller, M. (2010). The Commodification of Language. *Annual Review of Anthropology*, *39*(1), 101–114. https://doi.org/10.1146/annurev.anthro.012809.104951

Hill, J. (1998). Language, race, and white public space. *American Anthropologist*. *100*(3), 680–689. https://doi.org/10.1525/aa.1998.100.3.680

Ho, W. (2019). Developing professional communication: The construction of a multimodal

    understanding of job interviews. *Languages*, *4*(5), 1–12.

    https://doi.org/10.3390/languages4010005

Ho, W. Y. J., & Tai, K. W. H. (2020). Doing expertise multilingually and multimodally in online

    English teaching videos. *System*, *94*, 1–12. https://doi.org/10.1016/j.system.2020.102340

Holborow, M. (2015). *Language and neoliberalism*. Routledge.

Holliday, A. (2005). *The struggle of teach English as an international language*. Oxford

    University Press.

Holliday, A. (2006). "Native-Speakerism." *ELT Journal*, *60*, 385–387.

Houghton, S. A., & Rivers, D. J. (2013). *Native-Speakerism in Japan: Intergroup Dynamics in*

    Foreign Language Education. Multilingual Matters.

Hu, G. (2005). English language education in China: Polices, progress, and problems. *Language*

    Policy, *4*, 5–24.

Huang, H., Blommaert, J., & Praet, E. V. (2020). "Oh my god! Buy it!": A multimodal discourse

    analysis of the discursive strategies used by Chinese ecommerce live-streamer Austin Li.

    HCI International 2020 - Late Breaking Papers: Interaction, Knowledge, and Social

    Media, *12427*, 305–327. https://doi.org/10.1007/978-3-030-60152-2_24

Isbell, D. R. (2018). Online informal language learning: Insights from a Korean learning

community. *Language Learning & Technology*, *22*(3), 82–102. https://doi.org/10125/44658

Jenkins, H. (2009). *Confronting the challenges of participatory culture: Media education for the*

    21st century. The MIT Press. https://doi.org/10.7551/mitpress/8435.001.0001

Jerslev, A. (2016). *In the Time of the Microcelebrity: Celebrification and the YouTuber Zoella*.

    5233–5251.

Jewitt, C. (2009). *The Routledge handbook of multimodal analysis*. London: Routledge.

Jewitt, C. (2016). Multimodal analysis. In A. Georgakopoulou, & T. Spiloti (Eds.), *The Routledge handbook of language and digital communication* (pp. 69–84). Routledge.

Jiang, Y. & Vásquez, C. (2020). Exploring local meaning-making resources: A case study of a popular Chinese Internet meme (*biaoqingbao*). *Internet Pragmatics*, *3*(2), 260-282. https://doi.org/10.1075/ip.00042.jia

Jin, S. V., & Muqaddam, A. (2019). Product placement 2.0: "Do Brands Need Influencers, or Do Influencers Need Brands?" *Journal of Brand Management*, *26*(5), 522–537.

Jin, S. V., Muqaddam, A., & Ryu, E. (2019). Instafamous and social media influencer marketing. *Marketing Intelligence & Planning*, *37*(5), 567–579. https://doi.org/10.1108/MIP-09-2018-0375

Johnstone, B. (2009). Pittsburghese shirts: Commodification and the enregisterment of an urban dialect. *American Speech, 84*, 157–175

Jones, R. H. (2013). Multimodal discourse analysis. In C. A. Chapelle (Ed.) *The Encyclopedia of Applied Linguistics* (pp. 3992–3996). Wiley-Blackwell.

Jufferman, K. (2013). Multimodality and audiences: Local languaging in the Gambian linguistic landscape. *Sociolinguistic* Studies, *6*, 259–285.

Kang, H., & Lin, X. (2019). Lifelong learning on the go: English language mobile learning in China. *New Directions for Adult and Continuing Education*, 2019(162), 49–60. https://doi.org/10.1002/ace.20325

Kellner, D., & Kim, G. (2010). YouTube, critical pedagogy, and media activism. *Review of Education, Pedagogy, and Cultural Studies*, *32*(1), 49–60.

Khamis, S., Ang, L., & Welling, R. (2017). Self-branding, 'micro-celebrity' and the rise of

    Social Media Influencers. *Celebrity Studies*, *8*(2), 191–208.

    https://doi.org/10.1080/19392397.2016.1218292

King, R., & Wicks, J. (2009). "Aren't we proud of our language?": Authenticity,

    commodification, and the Nissan Bonavista television commercial. *Journal of English*

    Linguistics, *37*(3), 262–283. https://doi.org/10.1177/0075424209339924

Kress, G. (1993). Against arbitrariness: The social production of the sign as a foundational issue

    in critical discourse analysis. *Discourse and Society*, *4*(2), 169–191.

Kress, G. (2009). What is mode? In C. Jewitt (Ed.), *The Routledge handbook of multimodal*

    analysis (pp. 54-67). Routledge.

Kress, G. (2010). *Multimodality: A social semiotic approach to contemporary communication*.

    London: Routledge.

Kress, G., & van Leeuwen, T. (2001). *Multimodal discourse: The modes and media of*

    contemporary communication. Oxford University Press.

Kress, G., & van Leeuwen, T. (2006). *Reading images: The grammar of visual design* (2nd ed.).

    Routledge.

Kress, G., & van Leeuwen, T. (2021). *Reading images: The grammar of visual design* (3rd ed.).

    Routledge.

Kubota, R. (2011). Learning a foreign language as leisure and consumption: Enjoyment, desire,

    and the business of eikaiwa. *International Journal of Bilingual Education and*

    Bilingualism, *14*(4), 473–488.

Kubota, R. (2013). 'Language is only a tool': Japanese expatriates working in China and

    implications for language teaching. *Multilingual Education*, *3*(1), 1–20.

Kuure, L. (2011). Places for learning: Technology mediated language learning practices beyond the classroom. In P. Benson & H. Reinders (Eds.), *Beyond the language classroom* (pp. 35-46). Palgrave Macmillan.

Lam, A. S. L. (2005). *Language education in China: Policy and experience from 1949*. Hong Kong University Press.

Lamarre T (2017). Platformativity: Media studies, area studies. *Asiascape: Digital Asia*, *4*(3), 1–31.

Lange, P. G. (2007). The vulnerable video blogger: Promoting social change through intimacy. *The Scholar and Feminist Online*, *5*, 1–5.

Lee, T.-K. (2015). Translanguaging and visuality: Translingual practices in literary art. *Applied Linguistics Review*, *6*(4), 441–465. https://doi.org/10.1515/applirev-2015-0022

Leeman, J., & Modan, G. (2009). Commodified language in Chinatown: A contextualized approach to linguistic landscape. *Journal of Sociolinguistics*, *13*(3), 332–362. https://doi.org/10.1111/j.1467-9841.2009.00409.x

Li, A. K. (2019). Papi Jiang and microcelebrity in China: A multilevel analysis. *International Journal of Communication*, *13*. 3016–3034.

Li, G., Jee, Y., & Sun, Z. (2018). Technology as an educational equalizer for EFL learning in rural China? Evidence from the impact of technology-assisted practices on teacher-student interaction in primary classrooms. *Language & Literacy: A Canadian Educational E-Journal*, *20*(3), 159–184. https://doi.org/10.20360/langandlit29415

Li, H. (2020). Form disenchantment to reenchantment: Rural microcelebrities, short video, and the spectacle-ization of the rural lifescape on Chinese social media. *International Journal of Communication*, 14. 3769–3787.

Li, W. (2016). New Chinglish and the post-multilingualism challenge: Translanguaging ELF in China. *Journal of English as a Lingua Franca*, *5*(1), 1–25. https://doi.org/10.1515/jelf-2016-0001

Li, W. (2018). Translanguaging as a practical theory of language. *Applied Linguistics*, *39*(1), 9-30. https://doi.org/10.1093/applin/amx039.

Lirola, M. M., & Chovanec, J. (2012). The dream of a perfect body come true: Multimodality in cosmetic surgery advertising. *Discourse & Society*, *23*(5), 487–507. https://doi.org/10.1177/0957926512452970

Lowe, R. J., & Pinner, R. (2016). Finding the connections between native-speakerism and authenticity. *Applied Linguistics Review*, *7*(1), 27–52. https://doi.org/10.1515/applirev-2016-0002

Lu, X., Lu, Z. (2019, June). Fifteen seconds of fame: A qualitative study of Douyin, a short video sharing mobile application in China. In: Meiselwitz, G. (Ed.), *HCII 2019: Social computing and social media: Design, human behavior and analytics* (pp. 233–244). Springer. https://doi.org/10.1007/978-3-030-21902-4_17

Macaro, E., Handley, Z., & Walter, C. (2012). A systematic review of CALL in English as a second language: Focus on primary and secondary education. *Language Teaching*, *45*(1), 1-45. https://doi.org/doi:10.1017/S0261444811000395

Matwick, K., & Matwick, K. (2017). Cooking at home: A multimodal narrative analysis of the Food Network. *Discourse, Context & Media*, *17*, 20–29. https://doi.org/10.1016/j.dcm.2017.03.003

Markham, A., & Buchanan, E. (2012). *Recommendations from the AoIR Ethics Working Committee (Version 2.0)*. 1–19.

Marszałek, W. (2020, March, 29). A thorough guide to influencing on Douyin –For individuals

    and Businesses [Blog]. Retrieved from

    https://www.nanjingmarketinggroup.com/blog/influencers-brands-douyin

Marwick, A. E. (2013). *Status update: Celebrity, publicity, and branding in the social media*

    age. New Haven, CT: Yale University Press.

Marwick, A. E. (2015). Instafame: Luxury selfies in the attention economy. *Public Culture*,

    27(1): 137–160. https://doi.org/10.1215/08992363-2798379

Medgyes, P. (1994). *The Non-Native Teacher*. London: Macmillan.

Merriam, S. B., & Bierema, L. L. (2014). *Adult learning: Linking theory and practice*. Jossey-

    Bass.

Moor, P. J., Heuvelman, A., & Verleur, R. (2010). Flaming on YouTube. *Computers in Human*

    Behavior, *26*, 1536–1546.

Mortensen, K. K., & Maegaard, M. (2019). Meeting the Greenlandic people – Mediated

    intersections of colonial power, race and sexuality. *Discourse, Context & Media*, *32*, 1–9.

    https://doi.org/10.1016/j.dcm.2019.100348

Nissenbaum, H. (2010). *Privacy in context: Technology, policy, and the integrity of social life.*

    Stanford University Press.

Norris, S. (2004). *Analyzing multimodal interaction: A methodological framework*. Routledge.

Otheguy, R., García, O., & Reid, W. (2015). Clarifying translanguaging and deconstructing

    named languages: A perspective from linguistics. *Applied Linguistics Review*, *6*, 281–

    307. https://doi.org/10.1515/applirev-2018-0020.

Over 76% young Chinese live sped-up lives: survey. (2019, October). *Chinadaily.com.cn*.

> Retrieved from

> https://www.chinadaily.com.cn/a/201910/21/WS5dacf6d5a310cf3e355718c7.html#:~:text=BEIJING%20%2D%20More%20than%2076%20percent,percent%20and%2077.6%20percent%2C%20respectively

Park, J. S. Y., & Lo, A. (2012). Transnational South Korea as a site for a sociolinguistics of

> *g*lobalization: Markets, timescales, neoliberalism. *Journal of Sociolinguistics*, *16*(2), 147-

> 164.

Patterson, A. S. (2020). Freelance language teachers: Negotiating authenticity and legitimacy at

> the periphery. *TESOL Journal*, *11*(2). https://doi.org/10.1002/tesj.490

Pauwels, L. (2012) A multimodal framework for analyzing websites as cultural expressions.

> *Journal of Computer-Mediated Communication*, *17*, 247-265.

> https://doi.org/10.1111/j.1083-6101.2012.01572.x

Rahman, T. (2009). Language ideology, identity and the commodification of language in the call

> centers of Pakistan. *Language in Society*, *38*(2), 233–258.

> https://doi.org/10.1017/S0047404509090344

Relaño Pastor, A. M. (2015). The commodification of English in 'Madrid, comunidad bilingüe':

> Insights from the CLIL classroom. *Language Policy*, *14*(2), 131–152.

> https://doi.org/10.1007/s10993-014-9338-7

Piata, A. (2020). Stylistic humor across modalities: The case of Classical Art Memes. *Internet*

> *Pragmatics*, *3*(2), 174–201. https://doi.org/10.1075/ip.00031.pia

Qian, K., Owen, N., & Bax, S. (2018). Researching mobile-assisted Chinese-character learning

    strategies among adult distance learners. *Innovation in Language Learning and*

    Teaching, *12*(1), 56–71. https://doi.org/10.1080/17501229.2018.1418633

Raun, T. (2018). Capitalizing intimacy: New subcultural forms of micro-celebrity strategies and

    affective labour on YouTube. *Convergence: The International Journal of Research into*

    New Media Technologies, *24*(1), 99–113. https://doi.org/10.1177/1354856517736983

Recktenwald, D. (2018) *This discourse of online live streaming on Twitch: Communication*

    between conversation and commentary [Unpublished doctoral dissertation]. The Hong

    Kong Polytechnic University.

Roux, S., & Peck, A. (2019). The commodification of women's empowerment: The case of

    Vagina Varsity. *Discourse, Context & Media*, *32*, 1–9.

    https://doi.org/10.1016/j.dcm.2019.100343

Rubdy, R., & Tan, P. K. W. (Eds.). (2008). *Language as commodity: Global structures, local*

    marketplaces. Continuum.

Rymes, B. R. (2010). Classroom discourse analysis: a focus on communicative repertoires, In N.

    Hornberger and S. McKay (Eds.), *Sociolinguistics and language education* (pp. 528–

    546). Multilingual Matters.

Senft, T. M. (2008). *Camgirls: Celebrity and community in the age of social networks*. Peter

    Lang.

Selwyn, N. (2015). The discursive construction of education in the digital age. In R. H. Jones, A.

    Chik, & C. A. Hafner (Eds.), *Discourse and digital practices: Doing discourse analysis*

    *in the digital ag*e (pp. 81-96). Routledge.

Shutsko, A. (2020, July). User-generated short video content in social media: A case study of

    TikTok. In: Meiselwitz, G. (Ed.), *HCII 2020: Social computing and social media:*

    Participation, user experience, consumer experience, and applications of social

    computing (pp. 233–244). Cham, Switzerland: Springer. https://doi.org/10.1007/978-3-

    030-49576-3_8

Stevens, A. (2010). *Study on the Impact of Information and Communications Technology (ICT)*

    and New Media on Language Learning. European Commission. Retrieved from

    http://eacea.ec.europa.eu/llp/studies/documents/study_impact_ict_new_media_language_

    learning/final_report_en.pdf

Stevenson, M. P., & Liu, M. (2010). Learning a language with Web 2.0: Exploring the use of

    social networking features of foreign language learning websites. *CALICO Journal*,

    27(2), 233–259.

Sauro, S., & Zourou, K. (2017). Call for papers. *Language Learning & Technology*, *21*(1), 186.

    https://doi.org/10125/44603

Sauro, S., & Zourou, K. (2019). What are the digital wilds? *Language Learning & Technology*,

    23(1), 1–7. https://doi.org/10125/44666

Selvi, A. F. (2010). All teachers are equal, but some teachers are more equal than others: Trend

    analysis of job advertisements in English language teaching. *WATESOL NNEST Caucus*

    Annual Review, 1, 156–181.

Sockett, G. (2014). *The online informal learning of English*. New York: Palgrave Macmillan.

Song, Y. (2018). Translingual strategies as consumer design: A case study of multilingual

    linguistic landscapes of urban China. *Multilingua*, *37*(5), 455–482.

    https://doi.org/10.1515/multi-2017-0057

Sun, Y., Franklin, T., & Gao, F. (2017). Learning outside of classroom: Exploring the active part

of an informal online English learning community in China. *British Journal of

Educational Technology*, *48*(1), 57–70. https://doi.org/10.1111/bjet.12340

Tagg, C. (2015). *Exploring digital communication: Language in action*. Routledge.

Tan, T. R. F. (2008). The English language as a commodity in Malaysia: The view through the

medium-of-instruction debate. In P. K. W. Tan & R. Rubdy (Eds.), *Language as

commodity: Global structures, local marketplaces* (pp. 106–121). Continuum.

Tan, C. K., Wang, J., Wangzhu, S., Xu, J., & Zhu, C. (2020). The real digital housewives of

China's *Kuaishou* video-sharing and live-streaming app. *Media, Culture & Society*, *42*(7–

8), 1243–1259. https://doi.org/10.1177/0163443719899802

Tracy, K. (2002). *Everyday talk: Building and reflecting identities*. The Guilford Press.

Tupas, T. R. F. (2008). Anatomies of linguistic commodification: The case of English in the

Philippines vis-à-via other languages in the multilingual marketplace. In P. K. W. Tan &

R. Rubdy (Eds.), *Language as commodity: Global structures, local marketplaces* (pp.

89–105). Continuum.

Wang, W. (2015). Teaching English as an international language in China: Investigating

university teachers' and students' attitudes towards China English. *System*, *53*, 60–72.

Wang, H., & Chen, C. W. (2020). Learning English from YouTubers: English L2 learners' self-

regulated language learning on YouTube. *Innovation in Language Learning and

Teaching*, *14*(4), 333–346. https://doi.org/10.1080/17501229.2019.1607356

Weiss, J. C. (2019). How Hawkish Is the Chinese Public? Another Look at "Rising Nationalism"

and Chinese Foreign Policy. *Journal of Contemporary China*, *28*(119), 679–695.

https://doi.org/10.1080/10670564.2019.1580427

Williams, C. (1994). *Arfarniad o Ddulliau Dysgu ac Addysgu yng Nghyd-destun Addysg Uwchradd Ddwyieithog* [An evaluation of teaching and learning methods in the context of bilingual secondary education]. Unpublished Doctoral Thesis (University of Wales, Bangor).

Xiao, Y., Wang, L., & Wang, P. (2019). Research on the influence of content features of short videos marketing on consumer purchase intensions. *NMETSS 2019 – Advances in Social Science, Education and Humanities Research*, *351*, 415–422.

Xiong, T., & Yuan, Z. (2018). "It Was Because I Could Speak English That I Got the Job": Neoliberal Discourse in a Chinese English Textbook Series. *Journal of Language, Identity & Education*, *17*(2), 103–117. https://doi.org/10.1080/15348458.2017.1407655

Xu, Z. (2010). *Chinese English: Features and implications*. Hong Kong: Open University of Hong Kong Press.

Yao, S. (2014.). *Dialects into films—The element of "dialect" in Chinese films*. [Unpublished master's thesis]. University of Massachusetts Amherst

Yi, Y., & Jang, J. (2020). Envisioning possibilities amid the COVID-19 pandemic: Implications from English language teaching in South Korea. *TESOL Journal*, *11*(3). https://doi.org/10.1002/tesj.543

Yoo, I. W. (2014). Nonnative teachers in the expanding circle and the ownership of English. *Applied Linguistics*, *35*(1), 82–86

Zhang, L. T., & Cassany, D. (2021). "The murderer is him ✓": Multimodal humor in *danmu* video comments. *Internet Pragmatics*. *4*(2). 272–294. https://doi.org/10.1075/ip.00038.zha

Zhang, Y., & Ren, W. (2020). 'This is so skrrrrr' – creative translanguaging by Chinese micro-

    blogging users. *International Journal of Multilingualism*, 1–16.

    https://doi.org/10.1080/14790718.2020.1753746

Zimmerman, D. H. (2014). Identity, context and interaction. In A. Jaworski & N. Coupland

    (Eds.), *The discourse reader* (pp. 351-371). Routledge.

APPENDICES

Appendix A: A Multimodal Transcription of Yang's Video of Airport English

---

Introduction (0:00-0:07): frames the video.

Screen capture

0:00

On-screen verbal attributes	Off-screen verbal attributes
Chinese captions: *大家出国一落地一定会面临一个问题*	**Narration by Yang** (in Chinese)**:** *大家出国一落地一定会面临一个问题*
Translated captions in English: *Whenever you land in another country*	(Translation: Right after your flight arrives in a foreign country, there is one thing you must deal with.)

Visual attributes	Sound
Mid-shot	None
Background: Chinese and Western cultural items as room decor, white board	
Middle ground: Yang wearing a black T-shirt with "TureFake" printed on it, talking, having direct eye contact with viewers.	
Position of captions: close to the bottom of the screen; English translation placed under the Chinese caption	
Font size: Chinese caption in larger font, English translation in smaller font.	

Explanatory notes

Yang as the broadcaster starts an introduction to the topic of the video. Yang is speaking without smiling and making direct eye contact with viewers, indicating the seriousness of the topic.

Screen capture

0:02	0:03

On-screen verbal attributes	Off-screen verbal attributes
Chinese captions: *那就是过关. 那警官问你的都是英文,那怎么回答才好呢?* **Translated caption in English:** *You must go through immigration. And it could be scary since most likely they will ask you questions in English?*	**Narration by Yang** (in Chinese)**:** *那就是过关. 那警官问你的都是英文,那怎么回答才好呢?* (Translation: You will need to go through customs and immigration. However, all questions the immigration officer will ask are in English. How would you properly answer those questions?)

Visual attributes	Sound
Mid-shot **Background:** Airport customs, people in lines **Middle ground:** Yang wearing glasses, with a backpack (as the role of a tourist); Yang in a blonde wig, behind a computer monitor at a customs counter; Yang in professional uniform (playing the role of an immigration officer) **Position of captions:** Close to the bottom of the screen; English translation placed under the Chinese caption **Font size:** Chinese caption in larger font, English translation in smaller font.	None

Explanatory Notes
Role play is indicated by a variety of semiotic resources such as an airport background, the customs setting with a computer, people in lines, the customs officer uniform worn by Yang.

Role play and Explanation (0:07-0:39): How to answer the airport customs question "What is the purpose of your stay?"

Screen capture

0:07

On-screen verbal attributes	Off-screen verbal attributes
None	None
Visual attributes	**Sound**
A TV error effect	A TV static sound effect

Explanatory Notes

A TV error effect and a TV static sound effect are for transition.

Screen capture	

0:08

On-screen verbal attributes	Off-screen verbal attributes
English captions: *Top 1* *What is the purpose of your stay?* **Translated caption in Chinese:** *你逗留的目的是什么*	**Narration by Yang** (in Chinese)**:** *Top 1. What is the purpose of your stay?*
Visual attributes	**Sound**
Position of captions: Center; English caption above the Chinese caption **Color:** English caption in white; Chinese caption in yellow; black background **Font size:** Both English and Chinese are about the same size	None

Explanatory Notes

The airport customs question is highlighted using **color** (the white and yellow texts contrast sharply with the black background) and the **position of captions** (center).

Screen capture	
0:09	0:15
0:20	0:23

On-screen verbal attributes	Off-screen verbal attributes
English captions: (Immigration Officer): Next. What is the purpose of your stay? (Tourist A with a thick Cantonese accent): For school? To see my relative? For work and to find a wif…WiFi!! **Translated caption in Chinese:** (Immigration Officer):*你逗留的目的是什么？* (Tourist A): 去上学？探亲？找工作 再找个老… **Chinese captions:** (Tourist A): 怎么回事啊？去哪里啊？ **Translated caption in English:** (Tourist A): What happened? Where am I going?	**Role play conversation**: (Immigration Officer): Next. What is the purpose of your stay? (Tourist A): For school? To see my relative? For work and to find a wif…WiFi!! (Tourist A yells): wei wei wei ('wei' in Chinese is an interjection used to express surprise) (Tourist A in Cantonese): 怎么回事啊？去哪里啊？ (Translation: What's going on? Where are you taking me?)

Visual attributes	Sound
Mid-shot **Background:** Airport customs	Light clarinet music with a moderately quick tempo, playing in the background.

Middle ground: Yang wearing glasses, in a blue plaid shirt, with a backpack (as the role of tourist A); Yang in a blonde wig standing behind a computer monitor at a customs counter; in a professional uniform (playing the role of an immigration officer)
Position of captions: Close to the center of the screen.
Color: The captions are highlighted in yellow.
Font size: The English subtitles are larger in font size than the Chinese subtitles below

Explanatory notes

In role play when both parties communicate in English, the English and Chinese subtitles are placed right next to the person who is speaking (close to the center of the screen). The English captions are placed above the Chinese captions. The order of the two subtitles is reversed when the speech is in Chinese/Cantonese.

Screen capture

0:25

On-screen verbal attributes	Off-screen verbal attributes
None	None
Visual attributes	**Sound**
TV error effect	A censor beep sound

Explanatory notes

A TV error visual effect and a censor beep sound effect are for transition.

Screen capture

0:25

这些情况都要特定的签证
The answers above need a special visa

On-screen verbal attributes	Off-screen verbal attributes
Chinese captions: *这些情况都需要特定的签证, 所以如果你只是去玩的话就不要乱说啦！*	**Narration by Yang** (in Chinese)**:** *这些情况都需要特定的签证, 所以如果你只是去玩的话就不要乱说啦*
Translated captions in English: *These answers above need a special visa. So if you're just there for fun? Then play it safe!*	(Translation: You need to present specific types of visas to make these answers sound legit. If you are just visiting a foreign country for leisure, don't use those answers.)

Visual attributes	Sound
Mid-shot **Background:** Room decor, white board **Middle ground**: Yang talking, direct eye contact with viewers. **Position of captions:** Close to the bottom of the screen; English translation placed under the Chinese caption. **Font size:** Chinese caption in larger font, English translation in smaller font.	Light, clarinet music with a moderately quick tempo, playing in the background.

Explanatory Notes
Explanation is done after a brief role play section using spoken language.

Screen capture	
0:30	

On-screen verbal attributes	Off-screen verbal attributes
Chinese captions: *正确回应方式*	**Narration by Yang:** *正确回应方式*
Translated captions in English: *The correct way*	(Translation: The correct response)

Visual attributes	Sound
Position of captions: Center; Chinese caption above English caption **Color:** Both Chinese and English captions in yellow with a red shadow; black background	Light, clarinet music with a moderately quick tempo, playing in the background.

Font size: Chinese caption in larger font, English caption in smaller font

Explanatory Notes
The scene informs the viewers that the next part is about the correct ways to answer the discussed airport customs question, making a transition from the explanation to the next role play episode. The **colors** (the yellow texts with red shadow contrast sharply with the black background) and the **position of captions** (center) are used to emphasize the textual information.

Screen capture

0:32

0:33

0:34

0:38

On-screen verbal attributes	**Off-screen verbal attributes**
English captions:	**Role play conversation**:
(Immigration officer): *What is the purpose of your stay?*	(Immigration officer): *What is the purpose of your stay?*
(Tourist B): *I'm here for fun*	(Tourist B): *I'm here for fun*
I'm here to shop	*I'm here to shop*
(Two immigration officers): *Welcome*	(Yang as the immigration officer): *Ah!*
Translated caption in Chinese:	(Two immigration officers together):
(Immigration officer):*你逗留的目的是什么？*	*Welcome! Welcome! Welcome! Welcome!*
(Tourist B): *我是来玩的*	
我是来购物的	
(Two immigration officers): *欢迎欢迎*	

Visual attributes	**Sound**

Mid-shot
Background: Airport customs
Middle ground: Yang in red plaid shirt, with a backpack (as the role of tourist B); Yang and another person in a blonde wig and professional uniform (playing the role of immigration officers)
Gesture: Yang as an immigration officer rubbing two fingers and his thumb together while welcoming tourist B.
Position of captions: Close to the speaker
Color: The captions are highlighted in yellow.
Font size: The English subtitles are larger in font size than the Chinese subtitles below.

Light, clarinet music with a moderately quick tempo, playing in the background until Yang as tourist B finishes saying "I'm here to shop." Then, the Spring Festival Overture, a tune widely used on various festive occasions in China, is playing when the two immigration officers welcome tourist B.
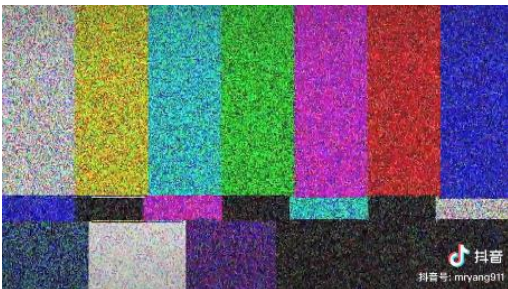
Explanatory Notes

Through the reenactment of the previous role play scene in which the immigration officer asks the question as well as through Yang's performance as Tourist B, this role play episode serves to contrast the language performance of Tourist A and Tourist B. Rubbing two fingers and the thumb together signifies money.

Role play and Explanation (0:39-1:01): How to answer the airport customs question "How long do you intend to stay?"

Screen capture

0:39

On-screen verbal attributes	Off-screen verbal attributes
None	None
Visual attributes	**Sound**
TV error effect	TV static sound
Explanatory notes	

A TV error effect and a TV static sound effect are for transition.

Screen capture

0:08

On-screen verbal attributes	Off-screen verbal attributes
English captions: *Top 2* *How long do you intend to stay?* **Translated caption in Chinese:** *你计划待多久?*	**Narration by Yang** (in Chinese)**:** *Top 2.* *How long do you intend to stay?*

Visual attributes	Sound
Position of captions: Center; English caption above Chinese caption **Color:** English caption in white; Chinese caption in yellow; black background **Font size:** Both English and Chinese are about the same size	Light, clarinet music with a moderately quick tempo, playing in the background.

Explanatory notes

The airport customs question is highlighted using **color (**the white and yellow texts contrasts sharply with the black background) and **position of captions** (center).
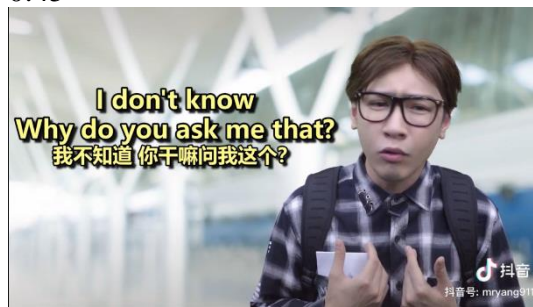
Screen capture

0:43	0:45
0:48	0:50

On-screen verbal attributes	Off-screen verbal attributes
English captions: (Immigration officer): *How long do you intend to stay?* (Tourist A): *I don't know Why do you ask me that? Maybe I don't want to go back* **Translated caption in Chinese:** (Immigration officer):*你计划呆多久?* (tourist A): *我不知道 你干嘛问我这个? 或许我就不想回去了嘞*	**Role play conversation**: (Immigration officer): *How long do you intend to stay?* (Tourist A with a thick Cantonese accent): *I don't know Why do you ask me that? Maybe I don't want to go back lei* ('lei' is a sentence-final modal particle to express acknowledgment) (Tourist A yells): *wei wei wei wei ah ah* ('wei' and 'ah' in Chinese are both interjections used to express surprise)

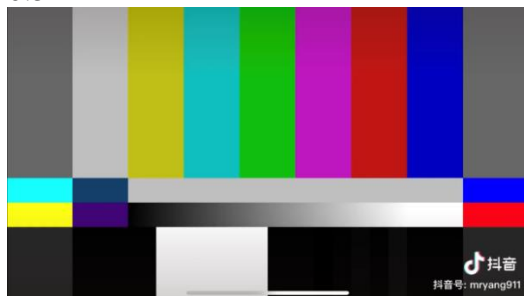| Visual attributes	Sound
Mid-shot **Background:** Airport customs **Middle ground:** Yang with hunched shoulders, wearing glasses, in a blue plaid shirt, with a backpack (as the role of Tourist A); Yang in a blonde wig behind a computer monitor at a customs counter; in professional uniform (playing the role of an immigration officer) **Position of captions:** Close to the center of the screen. **Color:** The captions are highlighted in yellow. **Font size:** The English subtitles are larger in font size than the Chinese subtitles below **Gesture:** Yang as tourist A has his shoe slipper sandals slide off and protests when he gets detained	Light, clarinet music with a moderately quick tempo, playing in the background.

Explanatory notes

In role play Yang performs some problematic ways to answer the airport customs question "how long do you intend to stay?"

Screen capture

0:51

On-screen verbal attributes	Off-screen verbal attributes
None	None
Visual attributes	**Sound**
A TV error effect	A TV censor beep sound
Explanatory notes	
A TV error visual effect and a censor beep sound effect are for transition.	

Screen capture	

0:51

On-screen verbal attributes	Off-screen verbal attributes
Chinese captions: *除非你的配偶真的在美国 或者是你有签证 否则就不要瞎说了* **Translated captions in English:** *Unless you REALLY do have a wife or husband there already or you have a visa Otherwise, stay in your lane*	**Narration by Yang** (in Chinese)**:** *除非你的配偶真的在美国 或者是你有签证 否则就不要瞎说了* (Translation: Unless your spouse is currently in the U.S. or your visa allows you to stay there for a certain amount of time, you should be specific when answering this question.)
Visual attributes	**Sound**
Mid-shot **Background:** Room decor, white board **Middle ground**: Yang talking, direct eye contact with viewers. **Position of captions:** Close to the bottom of the screen; English translation placed under the Chinese caption.	Light, clarinet music with a moderately quick tempo, playing in the background.

Font size: Chinese caption in larger font, English translation in smaller font.

Explanatory notes
Explanation is done after a brief role play section using spoken language in a classroom-like setting.

Screen capture
0:51

On-screen verbal attributes	**Off-screen verbal attributes**
Chinese captions: *正确回应方式* **Translated captions in English:** *The correct way*	**Narration by Yang** (in Chinese)**:** *正确回应方式* (Translation: The correct response)

Visual attributes	**Sound**
Position of captions: Center; Chinese caption above English caption **Color:** Both Chinese and English captions in yellow with red shadow; black background **Font size:** Chinese caption in larger font, English caption in smaller font	Light, clarinet music with a moderately quick tempo, playing in the background.

Explanatory notes
The scene informs that the next part is about the correct ways to answer the second airport customs question discussed in the video, making a transition from the explanation to the next role play episode. The **colors** (the yellow texts with red shadow contrast sharply with the black background) and the **position of captions** (center) are used to emphasize the textual information.

Screen capture	
0:57	0:59

1:00

On-screen verbal attributes	Off-screen verbal attributes
English captions:	**Role play conversation**:
(Immigration officer): *How long do you intend to stay?*	(Immigration officer): *How long do you intend to stay?*
(Tourist B): *I'm here for one week*	(Tourist B): *I'm here for one week*
Translated caption in Chinese:	*I'm here for one week* (The scene repeats, while the speech is slowed as the video is slowed down)
(Immigration officer): *你计划呆多久？*	
(Tourist B): *我要在这边待一个星期*	

Visual attributes	Sound
Mid-shot	Light, clarinet music with a moderately quick tempo, playing in the background. A "ting" sound effect occurs with the repeated frame (1:00).
Background: Airport customs	
Middle ground: Yang in red plaid shirt, with a backpack (as the role of Tourist B); Yang and another person in a blonde wig and professional uniform (playing the role of an immigration officer)	
Position of captions: close to the speaker and the center of the screen	
Color: The captions are highlighted in yellow and there is a green check mark. The repeated video frame in (1:00) is in black and white.	
Font size: The English subtitles are larger in font size than the Chinese subtitles below.	

Explanatory notes
In this role play episode, the example response to the airport customs question is highlighted using the repeated frame as well as colors (the big green check mark under the captions directs

attention to the captions; the black and white scene accentuates the yellow captions) and sound (the "ting" sound effect).

Role play and Explanation (1:02-1:24): How to answer the airport customs question "Where will you be staying?"

Screen capture

1:02

On-screen verbal attributes	**Off-screen verbal attributes**
None	None
Visual attributes	**Sound**
TV error effect	TV static sound
Explanatory notes	

A TV error effect and a TV static sound effect are for transition.

Screen capture

1:03

On-screen verbal attributes	**Off-screen verbal attributes**
English captions:	**Narration by Yang:** *Top 3. Where will you be staying?*
Top 3	
Where will you be staying?	
Translated caption in Chinese: 你将会住在哪里?	
Visual attributes	**Sound**

291

Position of captions: Center; English caption above Chinese caption **Color:** English caption in white; Chinese caption in yellow; black background **Font size:** about the same size of English and Chinese	Light, clarinet music with a moderately quick tempo, playing in the background.

Explanatory notes

The airport customs question is highlighted using **colors** (the white and yellow texts contrasts sharply with the black background) and the **position of captions** (center).

Screen capture

1:06	1:07
1:08	1:09
1:10	

On-screen verbal attributes	**Off-screen verbal attributes**
English captions: (Immigration officer): *Where will you be staying?* (Tourist A): *At a hotel. With a friend. At an Airbnb.*	**Role play conversation**: (Immigration officer): *Where will you be staying?* (Tourist A with a thick Cantonese accent): *At a hotel. With a friend. At an Airbnb.*

Translated caption in Chinese:
(Immigration officer):*你将会住在哪里?*
(Tourist A): *住在酒店里。住在朋友家里。住在民宿。*

Visual attributes	Sound
Mid-shot **Background:** Airport customs **Middle ground:** Yang wearing glasses, in a blue plaid shirt, with a backpack (as the role of tourist A); Yang in a blonde wig behind a computer monitor at a customs counter; in professional uniform (playing the role of an immigration officer) **Position of captions:** Close to the center of the screen; next to the person who is speaking **Color:** The captions are highlighted in yellow. There is a red cross mark in the last frame of the role play in black and white. **Font size:** The English subtitles are larger in font size than the Chinese subtitles below	Light, clarinet music with a moderately quick tempo, playing in the background. At 1:10, the video presents a black and white freeze frame of Yang as tourist A saying "at an Airbnb." A wrong buzzer sound effect occurs with the red cross mark and that still frame.

Explanatory notes

In role play, Yang performs some problematic ways to answer the airport customs question "Where will you be staying?" The last answer "At an Airbnb" is highlighted using colors and sound. For colors, the big red cross mark under the captions directs attention to the captions; the black and white scene accentuates the yellow captions and the cross mark. For sound, there was a "wrong" buzzer sound effect.

Screen capture

1:11

On-screen verbal attributes	Off-screen verbal attributes
Chinese captions: *这些回答可能都会引起警官的注意哦 所以就 uh uh…别乱说哦* **Translated captions in English:** *Those answers will probably raise a red flag for officers So uh uh… Don't talk rubbish*	**Narration by Yang** (in Chinese)**:** *这些回答可能都会引起警官的注意哦 所以就 uh uh…别乱说哦* (Translation: Those responses are likely to raise suspicion, so don't use them.)

293

Visual attributes	Sound
Mid-shot **Background:** Room decor, white board **Middle ground**: Yang talking, direct eye contact with viewers. **Position of captions:** Close to the bottom of the screen; English translation placed under the Chinese caption. **Font size:** Chinese caption in larger font, English translation in smaller font.	Light, clarinet music with a moderately quick tempo playing in the background.

Explanatory notes
Explanation is done after a brief role play section using spoken language in a classroom-like setting.

Screen capture
1:15

On-screen verbal attributes	Off-screen verbal attributes
Chinese captions: *正确回应方式* **Translated captions in English:** *The correct way*	**Narration by Yang** (in Chinese)**:** *正确回应方式* (Translation: The correct response)

Visual attributes	Sound
Position of captions: Center; Chinese caption above English caption **Color:** Both Chinese and English captions are in yellow with a red shadow; black background **Font size:** Chinese caption in larger font, English caption in smaller font	Light, clarinet music with a moderately quick tempo, playing in the background.

| Explanatory notes |
|---|
| The scene informs that the next part is about the correct ways to answer the second airport customs question discussed in the video, making a transition from the explanation to the next role play episode. The **colors** (the yellow texts with red shadow contrast sharply with the black background) and the **position of captions** (center) are used to emphasize the textual information. |

Screen capture

1:17

1:20

1:21
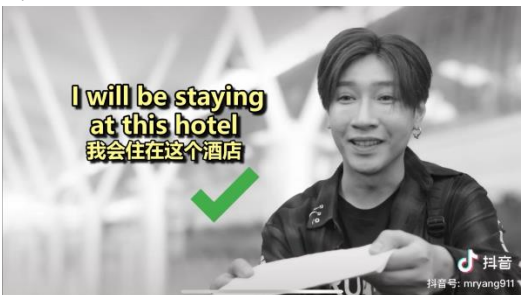
On-screen verbal attributes	Off-screen verbal attributes
English captions: (Immigration officer): *Where will you be staying?* (Tourist B): *I will be staying at this hotel* **Translated captions in Chinese:** (Immigration officer): *你将会住在哪里?* (Tourist B): *我会住在这个酒店*	**Role play conversation**: (Immigration officer): *Where will you be staying?* (Tourist B): *I will be staying at this hotel.*

Visual attributes	Sound
Mid-shot **Background:** Airport customs **Middle ground:** Yang in a red plaid shirt, with a backpack (as the role of tourist B); Yang and another person in blonde wigs and in a professional uniform (playing the role of an immigration officer) **Position of captions:** Close to the speaker and the center of the screen during the role play **Color:** The captions are highlighted in yellow; the check mark is green; the still frame is in black and white. **Font size:** The English subtitles are larger in font size than the Chinese subtitles below.	Light, clarinet music with a moderately quick tempo, playing in the background. At 1:21, the video presents a black and white frozen frame of Yang after the tourist B saying "I will be staying at this hotel". A 'ting' sound effect occurs with the green check mark and that still frame.

295

Gesture: Yang as tourist B presenting a piece
of paper with two hands

Explanatory notes

In this **role play** episode, the example response to the airport customs question is highlighted
using a still frame as well as colors (the big green check mark under the captions directs
attention to the captions; the black and white scene accentuates the yellow captions) and sound
(the "ting" sound effect).

Screen capture

1:22

On-screen verbal attributes	Off-screen verbal attributes
Chinese captions: (Yang as the broadcaster): *不管你住在哪里都要准备好详细的地址和有效讯息* **Translated captions in English:** *No matter where you're gonna stay,* *Be sure to have a detailed address and valid information*	**Narration by Yang as the broadcaster** (in Chinese)**:** *不管你住在哪里都要准备好详细的地址和有效讯息* (Translation: Wherever you will stay, make sure you provide a detailed address and valid contact information.)

Visual attributes	Sound
Position of captions: Captions of explanations are close to the bottom of the screen.	Light, clarinet music with a moderately quick tempo, playing in the background.

Explanatory notes

Explanations are provided to clarify how to respond to the airport customs question "Where
will you be staying?" appropriately and why the example response in the role play is effective
to answer it. Explanations are presented immediately after the role play through the narration
by Yang as the broadcaster while the still frame is showing. It is showed differently from the
previous parts of the video in which Yang's explanations are juxtaposed with the classroom-
like setting where he faces the camera. This time Yang delivers instruction mainly through
linguistic resources such as narration and captions. The transition between the role play and
the explanation after the role play is realized through the switch of linguistic codes (from
English during the role play to Mandarin Chinese for instructional language) as well as visual
resources including the different positions of captions, colors, and font sizes.

Role play and Explanation (1:24-1:52): how to answer the airport customs question "What is your occupation?"

Screen capture	
1:24	

On-screen verbal attributes	Off-screen verbal attributes
None	None
Visual attributes	**Sound**
TV error effect	TV static sound
Explanatory notes	
A TV error effect and a TV static sound effect are for transition.	

Screen capture	
1:25	

On-screen verbal attributes	Off-screen verbal attributes
English captions:	**Narration by Yang:** *Top 4. What is your occupation?*
Top 4	
What is your occupation?	
Translated caption in Chinese: 你的职业是什么？	

Visual attributes	Sound
Position of captions: center; English caption above Chinese caption	Light, clarinet music with a moderately quick tempo, playing in the background.
Color: English caption in white; Chinese caption in yellow; black background	

Font size: about the same size of English and Chinese

Explanatory notes

The airport customs question is highlighted using **color (**the white and yellow texts contrasts sharply with the black background) and **position of captions** (center).
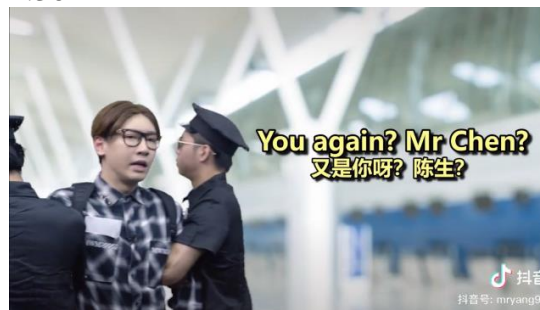

Screen capture

1:28	1:29
1:31	1:32
1:34	1:36

On-screen verbal attributes	Off-screen verbal attributes

English captions:
(Immigration officer): *What is your occupation?*
(Tourist A): *How do I say that… let me Baidu it*
    No work
    I am owned by someone
    You again? Mr Chen?
Translated caption in Chinese:
	Role play conversation：
	(Immigration officer): *What is your occupation?*
	(Tourist A):
	   [in Mandarin] *怎么说来着 我百度一下啊*
	   (Translation: How do I say it in English? Let me search it online.)
	   [in English] *No work*
	      I am owned by someone

298

(immigration officer): *你的职业是？*	[in Cantonese] *咦? 咦? 喂, 喂, 又是你呀? 陈生? 啊?*
(tourist A): *怎么说来着 我百度一下*	(Translation: What? What? Hey, hey, is that you again? Mr. Chen?)
没有工作	
我是被包养的	
又是你呀? 陈生	

Visual attributes	Sound
Mid-shot	Light, clarinet music with a moderately quick tempo, playing in the background.
Background: Airport customs	
Middle ground: Yang wearing glasses, in a blue plaid shirt, with a backpack (as the role of Tourist A); Yang in a blonde wig, behind a computer monitor at a customs counter; in professional uniform (playing the role of an immigration officer)	
Position of captions: Close to the center of the screen; next to the person who is speaking.	
Color: The captions are highlighted in yellow.	
Font size: The English subtitles are larger in font size than the Chinese subtitles below.	
Gesture: Yang as Tourist A takes out his cellphone to look up English words	

Explanatory notes
In **role play** Yang performs two problematic ways to answer the airport customs question "What is your occupation?". When Yang as Tourist A expresses that he does not know the English expression for his job and needs to look it up online, he speaks in Mandarin. Then in English, he tells the officer that he is jobless and lives a lavish lifestyle as a sugar baby. Immediately after his responses as Tourist A, again, he is detained by two other immigration officers just like some of the previous role play episodes show. While Tourist A is being taken away, he asks one of the two immigration officers in Cantonese "Is that you again? Mr. Chen?", which is a direct quote from a classic Hong Kong comedy movie in which a woman is surprised that she meets the same taxi driver who repeatedly misinterprets her requests. The intertextual reference to the movie generates humor.

Screen capture
1: 37

On-screen verbal attributes	Off-screen verbal attributes
None	None
Visual attributes	**Sound**
A TV error effect	A censor beep sound

Explanatory notes

A TV error visual effect and a censor beep sound effect are for transition.

Screen capture

1:38

On-screen verbal attributes	Off-screen verbal attributes
Chinese captions: *通常他们问这个问题呢 都想要确认一下你资金是没问题 你最好还是直接迅速地回答这个问题 要不然他们会问更多的问题就麻烦了*	**Narration by Yang:** *通常他们问这个问题呢 都想要确认一下你资金是没问题的· 你最好还是直接迅速地回答这个问题 要不然他们会问更多的问题就麻烦了*
Translated captions in English: They ask this question to see if you are financially equipped So you better just answer it quickly and directly, because if you don't they'll ask more.	(Translation: Usually, the reason why they ask this question is to make sure you have sufficient funds for travel. You'd better give a quick and straightforward answer. If not, they will ask more questions and you will get yourself in trouble.)
Visual attributes	**Sound**
Mid-shot **Background:** Room decor, white board **Middle ground**: Yang talking, direct eye contact with viewers.	Light, clarinet music with a moderately quick tempo, playing in the background.

Position of captions: Close to the bottom of the screen; English translation placed under the Chinese caption.
Font size: Chinese caption in larger font, English translation in smaller font.

Explanatory notes

Explanation is done after a brief role play section using spoken language in a classroom-like setting.

Screen capture

1:48

On-screen verbal attributes	Off-screen verbal attributes
Chinese captions: *正确回应方式* **Translated captions in English:** *The correct way*	**Narration by Yang** (in Chinese)**:** *正确回应方式* (Translation: The correct response)

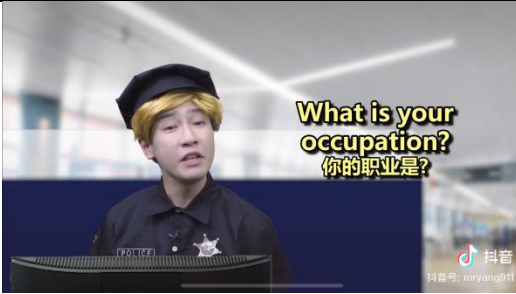| Visual attributes	Sound
Position of captions: Center; Chinese caption above English caption **Color:** Both Chinese and English captions in yellow with a red shadow; black background **Font size:** Chinese caption in larger font, English caption in smaller font	Light, clarinet music with a moderately quick tempo, playing in the background.

Explanatory notes

The scene informs that the next part is about the correct way to answer the second airport customs question discussed in the video, making a **transition** from the explanation to the next role play episode. The colors (the yellow texts with red shadow contrast sharply with the black background) and the position of captions (center) are used to emphasize the textual information.

Screen capture

1:50	1:51

1:52

On-screen verbal attributes	Off-screen verbal attributes
English captions: (Immigration officer): *What is your occupation?* (Tourist B): *I am an English teacher* *I am a fashion designer* **Translated captions in Chinese:** (Immigration officer): *你的职业是?* (tourist B): *我是一名英语老师* *我是一个时尚设计师*	**Role play conversation**: (Immigration officer): *What is your occupation?* (Tourist B): *I am an English teacher.* *I am a fashion designer.*

Visual attributes	Sound
Mid-shot **Background:** Airport customs **Middle ground:** Yang in red plaid shirt, with a backpack (as the role of Tourist B); Yang and another person in a blonde wig and professional uniform (playing the role of an immigration officer) **Position of captions:** Close to the speaker and the center of the screen during the role play; captions of explanations are close to the bottom of the screen. **Color:** The captions are highlighted in yellow; green check mark, the repeated scene in black and white. **Font size:** The English subtitles are larger in font size than the Chinese subtitles below.	Light, clarinet music with a moderately quick tempo, playing in the background.

Role play and Explanation (1:53-2:42): how to answer the airport customs question "Do you have anything to declare?"

Screen capture	
1:53	

On-screen verbal attributes	Off-screen verbal attributes
None	None
Visual attributes	**Sound**
TV error effect	TV static sound
Explanatory notes	

A TV error effect and a TV static sound effect are for **transition**.

Screen capture	
1:54	

On-screen verbal attributes	Off-screen verbal attributes
English captions: *Top 5* *Do you have anything to declare?* **Translated caption in Chinese:** 你有要申报的东西吗？	**Narration by Yang:** *Top 5. Do you have anything to declare?*

| Visual attributes	Sound
Position of captions: Center; English caption above Chinese caption **Color:** English caption in white; Chinese caption in yellow; black background **Font size:** about the same size of English and Chinese	Light, clarinet music with a moderately quick tempo, playing in the background.

Explanatory notes

The airport customs question is highlighted using color (the white and yellow texts contrasts sharply with the black background) and position of captions (center).
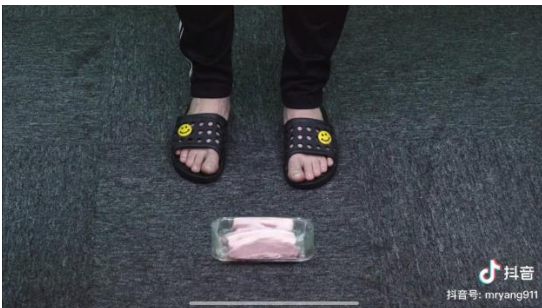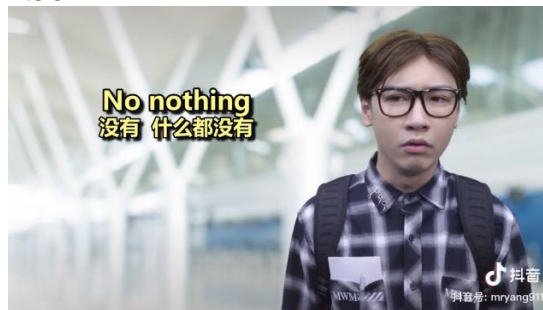
Screen capture

1:57

1:58

2:00

2:04

2:09

2:12

2:19
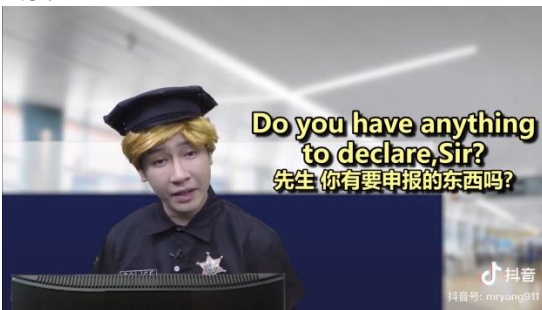
2:23

On-screen verbal attributes	Off-screen verbal attributes
English captions: (Immigration officer): *Do you have anything to declare, Sir?* (Tourist A): *No nothing* *I'm a car mechanic* *So have a spanner. Very normal* *In the night time, I am a pig killer Very normal* *Because I drink…I kill…I…* *My slippers* **Translated caption in Chinese:** (Immigration officer): 先生, 你有要申报的东西么? (Tourist A): 没有, 什么都没有 我是一个汽车维修员, 所以有一个"士巴拿" 非常正常 晚上我是个杀猪的 非常正常 因为我喝…我杀…我…哎呀我的拖鞋啊喂	**Role play conversation**: (Immigration officer): *Do you have anything to declare, Sir?* (Tourist A): [in English] *No nothing* *I'm a car mechanic So have a wrench Very normal.* *In the nighttime, I am a pig killer (mimicking pig squeal sound) Very normal* *Because I drink…I kill…I…* [in Cantonese] 哎呀呀我的拖鞋啊喂! (Translation: Ahh I dropped my slides!)

Visual attributes	Sound
Mid-shot **Background:** Airport customs **Middle ground:** Yang wearing glasses, in a blue plaid shirt, with a backpack (as the role of Tourist A); Yang in a blonde wig, behind a computer monitor at a customs counter; in professional uniform (playing the role of an immigration officer) **Position of captions:** Close to the center of the screen; next to the person who is speaking. **Color:** The captions are highlighted in yellow. **Font size:** The English subtitles are larger in font size than the Chinese subtitles below.	Light, clarinet music with a moderately quick tempo, playing in the background. When Tourist A drops a wrench and a pack of raw pork on the ground, a surprise sound effect is added.

Gesture: Yang as Tourist A repeatedly drops his personal items on the ground and picks them up.

Explanatory notes

In **role play,** Yang as Tourist A fails to declare some of his personal items that could be prohibited or restricted from entering a foreign country.

Screen capture

2:26

On-screen verbal attributes	Off-screen verbal attributes
Chinese captions: (Yang as the broadcaster):*肉类水果跟危险物品都不能带的！* **Translated caption in English**: *Meat, fruits and dangerous items are strictly prohibited*	**Narration by Yang** (in Chinese)**:** *肉类水果跟危险物品都不能带的！* (Translation: Meats, fruits, and dangerous items must not be brought across the border!)

Visual attributes	Sound
Mid-shot **Background:** Detention room **Middle ground:** Tourist A's right hand chained with handcuffs to a window **Position of captions:** Upper left and lower right of the screen **Color:** White captions; the still frame in black and white. **Font size:** The Chinese subtitles are much larger in font size than the English subtitles below. **Gesture:** Yang seated	None

---

Explanatory notes

Explanations this time are not presented in the classroom-like setting. Instead, a black and white still frame is used to show Yang as Tourist A gets detained in a room as a result of failing to declare prohibited or restricted items and attempting to bring those goods across the border. The explanations of what items are forbidden to bring into a foreign country are delivered through the narration by Yang as the broadcaster while the still frame is displaying.

The transition between the role play and the explanation after the role play is realized through the switch of linguistic codes (from English during the role play to Mandarin Chinese for instructional language) as well as visual resources including the different positions of captions, colors, and font sizes.

Screen capture

2:28

On-screen verbal attributes	Off-screen verbal attributes
Chinese captions: *正确回应方式* **Translated captions in English:** *The correct way*	**Narration by Yang** (in Chinese)**:** *正确回应方式* (Translation: The correct response)

Visual attributes	Sound
Position of captions: Center; Chinese caption above English caption **Color:** Both Chinese and English captions in yellow with red shadow; black background **Font size:** Chinese caption in larger font, English caption in smaller font	Light, clarinet music with a moderately quick tempo, playing in the background.

Explanatory notes

The scene informs that the next part is about the correct way to answer the second airport customs question discussed in the video, making a **transition** from the explanation to the next role play episode. The colors (the yellow texts with red shadow contrast sharply with the black background) and the position of captions (center) are used to emphasize the textual information.

Screen capture

2:30

2:32

2:34

On-screen verbal attributes	Off-screen verbal attributes
English captions: (Immigration officer): *Do you have anything to declare, Sir?* (Tourist B): *I brought a bottle of alcohol and some snacks* **Translated captions in Chinese:** (Immigration officer): *先生 你有要申报的东西么?* (Tourist B): *我带了一瓶酒和一些零食*	**Role play conversation**: (Immigration officer): *Do you have anything to declare, Sir?* (Tourist B): *I brought a bottle of alcohol and some snacks.*

Visual attributes	Sound
Mid-shot **Background:** Airport customs **Middle ground:** Yang in red plaid shirt, with a backpack (as the role of Tourist B); Yang and another person in a blonde wig and professional uniform (playing the role of an immigration officer) **Position of captions:** Close to the speaker and the center of the screen during the role play; captions of explanations are close to the bottom of the screen. **Color:** The captions are highlighted in yellow; green check mark; the still frame in black and white. **Font size:** The English subtitles are larger in font size than the Chinese subtitles below.	Light, clarinet music with a moderately quick tempo, playing in the background. At 2:34, the video presents a black and white, freeze frame of Yang after the Tourist B saying "I will be staying at this hotel". A 'ting' sound effect occurs with the green check mark and that still frame.

Explanatory notes

In this **role play** episode, the example response to the airport customs question is highlighted using a still frame, a big green check mark and a 'ting' sound effect. This also includes the colors (the big green check mark under the captions directs attention to the captions; the black and white scene accentuates the yellow captions) and sound (the "ting" sound effect). Additionally, explanations are provided to clarify how to respond to the airport customs question "Where will you be staying?" appropriately and why the example response in the role play is effective to answer the question.

Screen capture

2:35

On-screen verbal attributes	Off-screen verbal attributes
Chinese captions: (Yang as the broadcaster): 任何贵重的东西都需要申报的呦 **Translated captions in English:** *Anything expensive needs to be declared*	**Narration by Yang:** 任何贵重的东西都需要申报的呦 (Translation: Expensive items need to be declared.)

Visual attributes	Sound
Position of captions: Captions of explanations are close to the bottom of the screen.	Light, clarinet music with a moderately quick tempo, playing in the background.

Explanatory notes

Explanations are inserted into this role play episode through the narration by Yang as the broadcaster while the still frame is showing. Again, the transition between the role and the explanation is implicitly realized through the switch of linguistic codes and various visual resources.

Screen capture

2:37

2:38

2:40
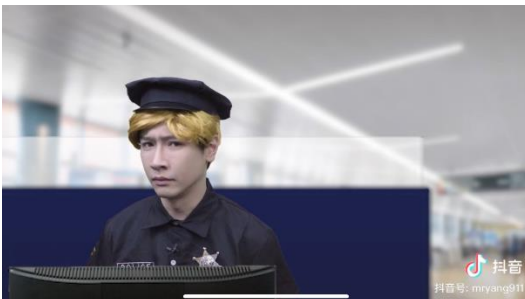
2:42

On-screen verbal attributes	Off-screen verbal attributes
English captions: (Tourist B): *I am here to shop* *(Two immigration officers): Welcome* **Translated captions in Chinese:** (Tourist B): *我是来消费的* (Tourist B): *欢迎欢迎*	**Role play conversation**: (Tourist B): *I am here to shop.* (Two immigration officers): *Ah! Welcome! Welcome! Welcome! Welcome!*

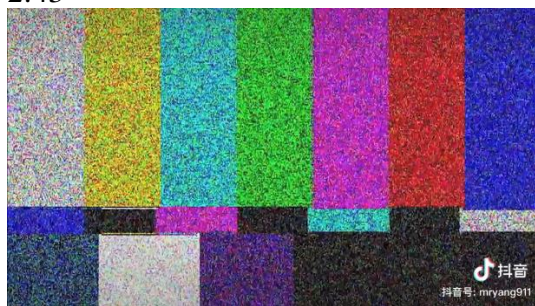| Visual attributes	Sound
Mid-shot **Background:** Airport customs **Middle ground:** Yang in red plaid shirt, with a backpack (as the role of tourist B); Yang and another person in a blonde wig and professional uniform (playing the role of immigration officers) **Gesture:** Yang as an immigration officer rubbing two fingers and his thumb together while welcoming Tourist B. **Position of captions:** Close to the speaker **Color:** The captions are highlighted in yellow. **Font size:** The English subtitles are larger in font size than the Chinese subtitles below. **Body language:** At first, Yang as an immigration officer frowns, moves his head slightly to the right side, and looks sideways at Tourist B. However, after Tourist B states that his purpose of visit is to shop, the two immigration officers smile widely, making a beckoning gesture with their palms facing up. Yang as one of the immigration officers also rubs two fingers and his thumb together.	Light, clarinet music with a moderately quick tempo, playing in the background until Yang as Tourist B finishes saying "I'm here to shop". Then, the Spring Festival Overture, a tune widely used on various festive occasions in China, is playing when the two immigration officers welcome Tourist B.

Explanatory notes
The **role play** resumes after the previous explanation. After Tourist B declared a bottle of alcohol and snacks he brought, the immigration officer was still suspicious of him. To convince the immigration officer that he should be allowed entry, Tourist B states that his

310

purpose of visit is to shop. Upon hearing that, the two immigration officers extend an immediate, warm welcome to Tourist B.

Concluding remarks (2:43-2:47): wraps up the video.

Screen capture

2:43

On-screen verbal attributes	Off-screen verbal attributes
None	None
Visual attributes	**Sound**
A TV error effect	A TV static sound
Explanatory notes	

A TV error effect and a TV static sound effect are for **transition**.

Screen capture

2:44	2:47

On-screen verbal attributes	Off-screen verbal attributes
Chinese captions: *跟警官 聊天不是闹着玩的 所以出国前一定要做好准备 学会了吗？* **Translated caption in English:** *Planning to speak with an officer is never fun So be sure to prepare well before you go*	**Narration by Yang:** *跟警官 聊天不是闹着玩的 所以出国前一定要做好准备 学会了吗？* (Translation: Talking to immigration officers is a serious matter, so make sure you prepare appropriately before going abroad. Got it?)
Visual attributes	**Sound**

311

|	
Mid-shot **Background:** Room decor, white board **Middle ground**: Yang talking, direct eye contact with viewers. **Position of captions:** Close to the bottom of the screen; English translation placed under the Chinese caption **Font size:** Chinese caption in larger font, English translation in smaller font.	Light, clarinet music with a moderately quick tempo, playing in the background

Explanatory notes
Yang concludes the video by addressing the importance of answering the airport customs questions appropriately.

Appeal to following (2:47-2:51)

Screen capture	
2:48	2:51
On-screen verbal attributes	**Off-screen verbal attributes**
Chinese captions: 我是明明可以靠脸的 但偏偏不要脸的 MrYang 飞~ **Translated caption in English:** I'm handsome but I am also an idiot sometimes Flying~	**Narration by Yang:** 我是明明可以靠脸的 但偏偏不要脸的 Mr. Yang 飞~ (Translation: I'm Mr. Yang. My face could be the moneymaker, but I insist on not saving it. See you.)
Visual attributes	**Sound**
Mid-shot **Background:** Room decor, white board **Middle ground**: Yang talking, direct eye contact with viewers. **Position of captions:** Close to the bottom of the screen; English translation placed under the Chinese caption **Font size:** Chinese caption in larger font, English translation in smaller font.	Light, clarinet music with a moderately quick tempo, playing in the background

Post-credit scene (2:52-3:00): What happens to tourist A after he was detained?

Screen capture

2:52
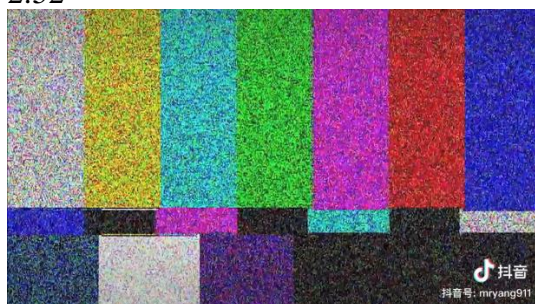
On-screen verbal attributes	Off-screen verbal attributes
None	None
Visual attributes	**Sound**
A TV error effect	A TV static sound

Explanatory notes

A TV error effect and a TV static sound effect are for **transition**.

Screen capture

2:52	2:57
2:59	3:00

On-screen verbal attributes	Off-screen verbal attributes
Chinese captions: 过关 过关这辈子都不能过关啦 是不可能的啦 又不会英文 啪! 可以过关了	**Role play conversation**: (Tourist B): 过关? 过关这辈子都不能过关啦, 是不可能的啦 ,又不会英文. 啊, 可以过关了. (Translation: To cross the border? I will never be able to cross the border in my entire life. It is a mission impossible because I don't speak English. [The handcuffs open suddenly, making a "pop" sound.] Well, now I can get through the customs.)

Visual attributes	Sound
Mid-shot **Background:** Detention room **Middle ground:** Yang in blue plaid shirt, with a backpack (as the role of Tourist A); his right hand chained with handcuffs. **Position of captions:** Close to the bottom of the screen **Color:** White captions **Gesture:** Yang seated when his right hand is chained with handcuffs; He stands up and goes out of camera sight after the handcuffs open.	The handcuffs open suddenly, making a "pop" sound. Someone laughing in the background after Tourist A claims that he can get through the customs.

Explanatory notes

The video ends with a post-credit scene, which is another **role play** episode telling viewers what happens to Tourist A after he was detained. In this scene, Tourist A claims that he will never get through immigration in his entire life because he does not speak English. This scenario constructed by Yang originates from a viral video interview on Chinese social media, in which a thief named Liqi Zhou was seen detained at a police station, with one hand chained to a window, proudly proclaiming that he would never work for others in his entire life and that he enjoyed the time in prison very much. This intertextual reference adds a humorous effect to the video. Meanwhile, by associating the ability to speak fluent English with the possibility to connect to the world, Yang highlights the significance of learning English. What happens next in the role play—the handcuffs open suddenly (possibly due to an issue with the

locking mechanism) and Tourist A mistakenly takes it as an official notice that he is allowed entry to the foreign county, adds an additional layer of humor.

Appendix B: A Multimodal Transcription of Dabai's Video of Vocabulary

Introduction (0:00): To introduce the topic and provide a preview.

Screen capture	**On-screen verbal attributes**	**Off-screen verbal attributes**
0:00 	**Chinese captions:** 加不加 s 单词意思大不同 (Translation: adding -s ending may completely change the meaning of a word)	**None**

Visual attributes	**Sound**
Mid-shot **Video layout:** A split screen on the left and the right side **Background:** Dark gray curtain **Middle ground:** On the left side of the screen, Dabai is in an azure blue shirt, facing the camera; on the right side of the screen, Dabai is wearing a shirt of the same style but in turquoise blue, facing the camera. On both sides of the screen, Dabai remains in the same posture and is wearing a white baseball cap. **Captions (position/color/font size):** Near the bottom of the screen; yellow font color with white text background.	**None**

Explanatory notes
The introduction to the topic of the video is presented through a still frame using textual and visual resources. Verbal resources are not used in this part. The split screen shown in this introductory still frame also provide a preview of the layout of the video.

Explanation (0:01-0:16)

Screen capture

0:01	0:02	0:04	0:06

0:08　　　　0:09　　　　0:13　　　　0:15

On-screen verbal attributes	Off-screen verbal attributes
Captions (on the top of the screen): *加不加 s 单词意思大不同* (Translation: adding -s ending may completely change the meaning of a word) **Captions** (near the bottom): *good 好的　　goods 商品　 green 绿色　 greens 蔬菜　 new 新的　 news 新闻　 glass 玻璃　 glasses 眼镜*	**Narration by Dabai**: *good, 好的; goods, 商品; green 绿色; greens 蔬菜; new 新的; news 新闻; glass 玻璃; glasses 眼镜*

Visual attributes	Sound
Mid-shot	**None**

A Split screen: On the left and the right side
Background: Dark gray curtain on both sides of the screen
Middle ground: On the left side of the screen, Dabai is in an azure blue shirt, facing the camera; on the right side of the screen, Dabai is wearing a shirt of the same style but in turquoise blue, facing the camera. On both sides of the screen, Dabai remains in the same posture and wears a white cap.
Captions (position/color/font size): The descriptive words/title of the video is moved (from the bottom of the screen) appearing on the still frame at the beginning of this video, to the top of the screen with a black background. The caption stays in yellow, but the font size is

317

slightly smaller. The English words and their Chinese translation are presented in white text mainly, placed on the lower part of the screen. The difference in the spelling between each pair of the words is displayed on the split-screen, which is the added -s or -es suffix , is marked in yellow in the word on the right side of the screen.

Explanatory notes
For each pair of words that differ in their endings, Dabai presents one that does not end with an -s or -es on the left of the screen by reading the word and the Chinese translation out loud first. During this time, the right side of the screen is frozen. The left side of the screen freezes immediately upon Dabai finishes uttering the word and its meaning in Mandarin, and the captions remain. Then the right side of the screen shows Dabai presenting the second word that ends with an -s or -es and its Chinese translation. As soon as Dabai reads the word and its Chinese meaning, the right side of the screen freezes again. The comparison of the two words are indicated by the differences in the captions, which include the -s or -es suffix highlighted in yellow and the different Chinese meanings. It is also suggested by the split screen in which the shirts Dabai wears are different colors. Once both of the words in a pair are presented, the captions are removed, which suggests the end of the comparison and a new pair of words will be introduced. Then the video moves on to show the next pair of words following the same structure: the word without an -s or -es suffix is introduced first on the left side of the screen, while the right side of the screen is frozen. Then the left side of the screen freezes, directing viewers' attention to the right side where the second word that ends with an -s or -es appears.

Concluding remarks (0:17)

Screen capture	On-screen verbal attributes	Off-screen verbal attributes
0:17	**Chinese captions:** 学会了吗	**Narration by Dabai:** 学会了吗? (Translation: Got it?)

Visual attributes		Sound
Mid-shot **Background:** Dark gray curtain		A popular *Douyin* sound effect often used when a task is

Middle ground: Dabai in an azure blue shirt, sitting in a green chair, facing the camera **Captions (position/color/font size):** In the middle of the screen**,** white text contrasts Dabai's blue shirt.		completed, or a goal is achieved
Explanatory notes		
Dabai wraps up the video lesson by asking the viewers if they have learned what he taught, which is a question that many *Douyin* vloggers use at the end of their videos.		

Appeal to likes (0:18-0:19)

Screen capture	On-screen verbal attributes	Off-screen verbal attributes
0:18 	**Chinese captions:** *记得点赞～*	**Narration by Dabai:** *记得点赞* (Translation: Don't forget to leave a like)

| Visual attributes	Sound
Mid-shot **Background:** Dark gray curtain **Middle ground:** Dabai in an azure blue shirt, sitting in a green chair, facing the camera **Captions (position/color/font size):** In the middle of the screen**,** white text that contrasts Dabai's blue shirt. **Gesture:** Dabai puts his right hand close to the camera, making a "finger heart" gesture by crossing his thumb and index finger	A popular *Douyin* sound effect "打卡成功" (mission complete) often used when a task is completed, or a goal is achieved.
Explanatory notes	
Daibai uses a "finger heart" gesture, making his fingers look like a heart shape, to appeal to likes and show his positive feeling to viewers.	

Appendix C: A Multimodal Transcription of Scott's Video of English Prepositions

---

Introduction (0:00-0:03): Introduces the topic of discussion

Screen capture

0:00
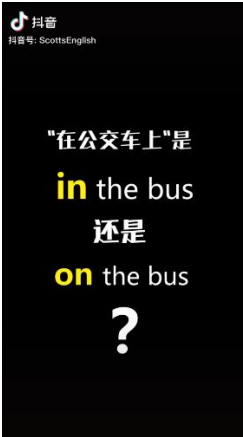
On-screen verbal attributes	Off-screen verbal attributes
Captions: *"在公交车上"是 in the bus 还是 on the bus?* (Translation: Which is correct: 'in the bus' or 'on the bus'?)	**Narration by Scott:** *"在公交车上"是 in the bus 还是 on the bus?* (Translation: Which is correct: 'in the bus' or 'on the bus'?)

Visual attributes	Sound
Black **Background** **Captions (position/color/font size):** Text filling the center of the screen; the English word "in" and "on" are in yellow, the rest of the text (both Chinese and English) are in white.	a short clip of bluegrass music played with a quick tempo in the background

Explanatory notes
The introduction frames the video by raising a question that draws viewers' attention to the English preposition usage. The emphasis on the choice of prepositions as the video's focus is realized through color: the two prepositions "in" and "on" in the text are highlighted in yellow, which contrast with the black ground and the rest of the text in white. In addition, Scott reads these two prepositions in narration with emphasis to further direct audiences' attention.

Explanation and Examples (0:03-0:20)

Screen capture

0:03	0:05

On-screen verbal attributes	Off-screen verbal attributes
Captions: *简单得很嘛 一条规则 你必须坐在里面的就是 in* (Translation: It's simple. One rule. You use 'in' before the object that requires you to sit inside)	**Narration by Scott:** *简单得很嘛 一条规则 你必须坐在里面的就是 in* (Translation: It's very simple. You just need to remember one rule. You use 'in' before the object that requires you to sit inside)

Visual attributes	Sound
Video layout: Top half screen features Scott wearing a white collared shirt; texts placed right under the top half of the screen **Medium close-up shots** **Camera angle:** Slightly low angle (camera points slightly upwards) **Background:** <of the video>: in grey <of the top half screen featuring Scott>: The background where Scott is featured in appears to be Scott's home. As Scott keeps moving while talking to the camera, things appear in the background, including a cat and a patio. **Captions (position/color/font size):** Placed right under the top half of the screen; large front size; Chinese texts are in yellow while the English preposition "in" mentioned in the video is in white. **Gesture:** Scott raises his index finger when says "One rule".	a short clip of bluegrass music played with a quick tempo in the background

Explanatory notes

The emphasis on the discussed preposition "in" is again, realized through color. The difference in color between "in" in white and other texts in yellow creates a highlight on the preposition "in". The background where Scott is featured in changes from time to time, as Scott moves while talking to the camera. What appears in the background, including a cat and a patio, indicates that Scott made the video at home. A sense of intimacy and ordinariness is conveyed through close-up shots, the home settings, and the pet. In addition, the slightly low camera angle, which helps

321

associate power with Scott, as well as the formal white collared shirt Scott wears indexes professionalism, serves to establish Scott's identity as an English expert.

Screen capture

0:08	0:09	0:10	0:11

On-screen verbal attributes	Off-screen verbal attributes
English Captions: *in a car* *in a taxi* *in a helicopter* **Translated Captions in English:** *在轿车里 在出租车里 在直升机里* **Chinese Captions:** *这些交通工具你几乎没有办法站立起来* (Translation: You can barely stand in those transportation vehicles.)	**Narration by Scott:** *in a car, in a taxi, in a helicopter.* 这些交通工具你几乎没有办法站立起来 (Translation: You can barely stand in those transportation vehicles.)

Visual attributes	Sound
Video layout: Top half screen features Scott wearing a white collared shirt. Texts are placed on the center of the screen if there are images showing on the bottom half of the screen. Captions are placed right under the top half of the screen when there is no image showing on the bottom half of the screen. **Medium close-up shots** **Camera angle:** Slightly low angle (camera points slightly upwards) **Background:** <of the video>: in grey <of the top half screen featuring Scott>: The background where Scott is featured in appears to be Scott's home. As Scott keeps moving while talking to the camera, more things appear in the background. In addition to the cat and the patio showing up in the previous screenshots, some home decorations and storage boxes appear in the background as well. **Captions (position/color/font size):** Captions that are associated with examples are placed on the center of the screen if there are images showing on the bottom half of the screen, with	a short clip of bluegrass music played with a quick tempo in the background

English captions above the Chinese texts. English and Chinese texts are in the same size, both in white.

Captions that are associated with explanations are placed right under the top half of the screen when there is no image showing on the bottom half of the screen. The explanatory words are in yellow.

Gesture: Scott raises his index finger almost each time when he provides a new example.

Explanatory notes

Scott provides three different phrases in which the preposition "in" is used to attach to a transportation vehicle. Each sample phrase is juxtaposed with an image on the bottom half of the screen showing people sitting in the transportation vehicle mentioned in the sample phrase, illustrating and supporting the rule Scott previously mentioned: "in" is used before the object that requires you to sit inside. The translated caption in Chinese for each English expression uses "在…里" (to be … inside) instead of "在…上"(to be…on), although both are acceptable in Chinese. The images showing people sitting in the transportation vehicles and the word choice in translated caption in Chinese are combined to associate "里" (meaning: inside) with "in" and objects that require people to sit inside. The Chinese and English texts placed in the center of the screen slightly overlap with the top half of the screen featuring Scott and the images displayed on the bottom half of the screen, through which the information provided from these three parts unify as a coherent piece. After examples, Scott further summarizes that as people can barely stand in a "car", "taxi", and "helicopter", "in" is the right preposition to use to describe someone's presence in those transportation vehicles.

Screen capture

0:14

On-screen verbal attributes	Off-screen verbal attributes
Captions: *身体可以立起来的就是 on* (Translation: You use 'on' before the object that you can stand inside.)	**Narration by Scott:** *身体可以立起来的就是 on* (Translation: You use 'on' before the object that you can stand inside.)

Visual attributes	Sound
Video layout: Top half screen features Scott wearing a white collared shirt; texts placed right under the top half of the screen	a short clip of bluegrass music

Medium close-up shots	played with a quick
Camera angle: Slightly low angle (camera points slightly upwards)	tempo in the
Captions (position/color/font size): Placed right under the top half of the	background
screen; large front size; Chinese texts are in yellow while the English	
preposition "on" mentioned in the video is in white.	
Captions (position/color/font size): Placed right under the top half of the	
screen; large front size; Chinese texts are in yellow while the English	
preposition "on" mentioned in the video is in white.	
Gesture: Scott raises his index finger from time to time	

Explanatory notes

Scott moves on to explain when the preposition "on" needs to be used.

Screen capture

0:16	0:17	0:18	0:19

On-screen verbal attributes	Off-screen verbal attributes
English Captions: *on a ship*	**Narration by Scott:** *on a ship*
on a bus	*on a bus*
on a bike	*on a bike*
on a skateboard	*on a skateboard*
Translated Captions in English: *在船上 在*	
公交车上 在自行车上 在滑板上	

Visual attributes	Sound
Video layout: Top half screen features Scott wearing a white collared	a short clip of
shirt. Images illustrating examples of English phrases are showing on the	bluegrass music
bottom half of the screen.	played with a quick
Medium close-up shots:	tempo in the
Camera angle: Slightly low angle (camera points slightly upwards)	background
Background:	
<of the video>: in grey	
<of the top half screen featuring Scott>: The background where Scott is	
featured in appears to be Scott's home. As Scott keeps moving while	

324

talking to the camera, things such as furniture and home appliances appear in the background.

Captions (position/color/font size):
Captions that are associated with examples are placed on the center of the screen, with English captions above the Chinese texts. English and Chinese texts are in the same size, both in white.

Gesture: Scott raises his index finger or puts his thumb and index finger as a circle when giving examples.

Explanatory notes

Scott provides four different phrases in which the preposition "on" is used to attach to a mode of transportation. Each sample phrase is juxtaposed with an image on the bottom half of the screen showing people standing on the transportation vehicle mentioned in the sample phrase, illustrating and supporting the rule Scott previously mentioned: "on" is used before the object that people can stand on. The translated caption in Chinese for each English expression uses "在…上" (to be … on). The images showing people standing a transportation tool and the word choice in translated caption in Chinese are combined to associate "上" (meaning: on) with "on" and objects that people can stand on.

Appeal to likes (0:20-0:21)

Screen capture

0:20

On-screen verbal attributes	Off-screen verbal attributes
Chinese Captions: *点赞* (Translation: give me a like)	**Narration by Scott:** (with a Sichuan accent) *点赞!* (Translation: Give me a like!)

Visual attributes	Sound
Video layout: Top half screen features Scott wearing a white collared shirt; texts placed right under the top half of the screen **Medium close-up shots** **Camera angle:** Slightly low angle (camera points slightly upwards) **Captions (position/color/font size):** Chinese caption in yellow, placed right under the top half of the screen; large front size.	a short clip of bluegrass music played with a quick tempo in the background

Gesture: Scott raises his hand with his palm facing the camera and quickly waves at the camera vertically.

Explanatory notes

Scotts makes a quick, vertical, hand waving gesture at the camera to get more viewers' attention and says "give me a like" with a demanding tone of voice and a Sichuan accent to appeal to likes. The demanding tone of voice makes Scotts request for likes sounds like a direct order from someone in power, which corresponds to the public identity Scotts intends to establish: an English language teacher/expert. However, the directness of the request is mitigated by the Sichuan accent Scotts adopts to convey a sense of humor.

Appendix D: A Multimodal Transcription of Jiadecang's English Learning Vlog

---

<div align="center">

Vlog (0:00-0:13)

</div>

Screen capture

0:00   0:01

On-screen verbal attributes	Off-screen verbal attributes
Chinese captions: *自学英语第 959 天* (Translation: Day 959 of learning English by myself)	**None**

Visual attributes	Sound
Long shots with camera gradually zooming in **Background:** Blue sky, a yard full of dirt and small rocks,(traditional?) Tibetan style residential architectures with flat, red roofs, square-shaped overhang decorations, many windows, built on an elevated site; a tractor **Middle ground:** Jiadecang sitting on the ground, facing a campfire made of wood and stone bricks. He is wearing a winter jacket, a shirt with big, white English letters on it, and sneakers. A large black pot is placed on the campfire. A red tractor is behind him. **Captions (position/color/font size):** On the top of the screen; yellow font color **Gesture:** Jiadecang is sitting on the ground, putting more wood into the fire and blowing on the fire to help it burn better	A Tibetan folk song called "Aku Pema" is playing in the background.

Explanatory notes

The text superimposed on the top of the video is used to record the number of the vlogs of Jiadecang's daily English language learning routine. It serves as both a record of Jiadecang's English language learning progress and the title of the vlog. A Tibetan rural area and a Tibetan folk song are featured to construct Jiadecang's authenticity as a Tibetan man.

Screen capture

0:03   0:05

On-screen verbal attributes	Off-screen verbal attributes
Captions (On the top of the screen): *自学英语第 959 天* (Translation: the 959th day of learning English by myself) **Captions** (Slightly near the bottom): <English captions>: *Life is a kind of enjoyment, don't complain about anything* <Chinese captions>: 生活是一种享受 不要抱怨任何事	**Narration by Jiadecang** (in Chinese)**:** *Life is a kind of enjoyment, so don't complain about anything.* (Meaning: Enjoy your life and stop complaining)

Visual attributes	Sound
Long shot **Background:** Tibetan style residential architectures with flat, red roofs, square-shaped overhang decorations, many windows, and painted black frames around doors, built on an elevated site **Middle ground:** Jiadecang is stewing many big chunks of bone-in yak meat in a large pot on a campfire. **Gesture:** Jiadecang removes the lid of the pot. **Captions (position/color/font size):** The text indicating the number of days learning English remains unchanged, staying on the top of the video. The English-speaking Jiadechang narrates and its Chinese translation is superimposed near slight near the bottom. The font size of the English text and the Chinese translation is much smaller than the top text. All texts are in yellow front color.	A Tibetan folk song called "Aku Pema" is playing in the background.

Explanatory notes

Jiadecang shares his reflection on life in English through the voice-over narration. The fashionable, athletic clothing of Jiadecang and his use of English stands in a stark contrast to the open fire cooking and the underdeveloped Tibetan rural region in the video.

Screen capture

0:08

On-screen verbal attributes	Off-screen verbal attributes
Captions (On the top of the screen*): 自学英语第959天* (Translation: the 959th day of learning English by myself) **Captions** (slightly near the bottom): <English captions>: Cook the yak meat <Chinese captions>: *煮牦牛肉*	**Narration by Jiadecang:** *Cook the yak meat* (Meaning: I'm making a yak stew)

Visual attributes	Sound
Close-up shot **Foreground:** Big chunks of bone-in yak meat covered by water in a large pot. **Captions (position/color/font size):** The text indicating the number of days learning English remains unchanged, staying on the top of the video. The English-speaking Jiadechang narrates and its Chinese translation is superimposed near slight near the bottom. The font size of the English text and the Chinese translation is much smaller than the top text. All texts are in yellow front color.	A Tibetan folk song called "Aku Pema" is playing in the background.

Explanatory notes
Jiadecang explains in English that he is making a yak stew. The yak stew is a traditional Tibetan dish that contributes to Jiadecang's authenticity as a Tibetan.

Screen capture	
0:11	0:12

On-screen verbal attributes	Off-screen verbal attributes
Captions (On the top of the screen*): 自学英语第 959 天* (Translation: the 959th day of learning English by myself)	**Narration by Jiadecang:** *The yak meat*

Visual attributes	Sound
Mid-shot **Gesture:** Jiadecang puts the lid back to the pot, then he winks and clicks his tongue at the camera. **Captions (position/color/font size):** The text indicating the number of days learning English remains unchanged, staying on the top of the video.	A Tibetan folk song called "Aku Pema" is playing in the background.

Explanatory notes
Jiadecang repeats "the yak meat" in lowered voice after saying "cooking the yak meat", directing viewers' attention to this English expression of "the yak meat", perhaps particularly to the word "yak" as it is not a common word that learners would come across. However, as Jiadecang repeats "the yak meat" in lowered voice and this expression he repeats is not presented as text on the video, it makes the information he repeats seemingly less audience oriented. Rather, the action of repeating could be the way Jiadecang uses to practice English and memorize new words, including "yak", which then contributes to Jiadecang's leaner identity. He winks and clicks his tongue at the camera to make an asynchronous interaction with audiences.

Appendix E: Copyright Permission

CCC Marketplace™

This is a License Agreement between Yaqian Jiang ("User") and Copyright Clearance Center, Inc. ("CCC") on behalf of the Rightsholder identified in the order details below. The license consists of the order details, the Marketplace Order General Terms and Conditions below, and any Rightsholder Terms and Conditions which are included below.

All payments must be made in full to CCC in accordance with the Marketplace Order General Terms and Conditions below.

Order Date	20-Jun-2022	Type of Use	Republish in a thesis/dissertation
Order License ID	1238385-1		
ISBN-13	9781032268279	Publisher	Taylor and Francis (Books) Limited UK
		Portion	Page

LICENSED CONTENT

Publication Title	English in China	Rightsholder	Taylor & Francis Informa UK Ltd - Books
Date	06/10/2022		
Language	English	Publication Type	Book

REQUEST DETAILS

Portion Type	Page	Rights Requested	Main product
Page range(s)	97-100	Distribution	Worldwide
Total number of pages	4	Translation	Original language of publication
Format (select all that apply)	Electronic		
		Copies for the disabled?	No
Who will republish the content?	Academic institution	Minor editing privileges?	Yes
Duration of Use	Current edition and up to 5 years	Incidental promotional use?	No
		Currency	USD
Lifetime Unit Quantity	Up to 499		

NEW WORK DETAILS

Title	The Making of Microcelebrities on Douyin: Multimodal Design and Online Identities in Informal English Instruction Videos	Institution name	University of South Florida
		Expected presentation date	2022-07-01
Instructor name	Camilla Vasquez		

ADDITIONAL DETAILS

Order reference number	N/A	The requesting person / organization to appear on the license	Yaqian Jiang

REUSE CONTENT DETAILS

Title, description or numeric reference of the portion(s)	Teaching English on Douyin: Language commodification and translanguaging	Title of the article/chapter the portion is from	Teaching English on Douyin: Language commodification and translanguaging
Editor of portion(s)	Songqing Li	Author of portion(s)	Yaqian Jiang
Volume of serial or monograph	N/A	Issue, if republishing an article from a serial	N/A
Page or page range of portion	97-100	Publication date of portion	2022-06-15

Appendix F: USF Fair Use Worksheet

INSTRUCTIONS

Check all boxes that apply, and keep a copy of this form for your records. If you have questions, please contact the USF General Counsel or your USF Tampa Library Copyright Librarian.

Name: Yaqian Jiang   Date: 7/15/2022

Class or Project: Dissertation

Title of Copyrighted Work: The making of microcelebrities on Douyin: Multimodal design and online identities in informal English instruction videos

PURPOSE AND CHARACTER OF THE USE

Likely Supports Fair Use	Likely Does Not Support Fair Use
☑ Educational	☐ Commercial
☐ Teaching (including multiple copies for classroom use)	☐ Entertainment
☑ Research or Scholarship	☐ Bad-faith behavior
☑ Criticism, Parody, News Reporting or Comment	☐ Denying credit to original author
☐ Transformative Use (your new work relies on and adds new expression, meaning, or message to the original work)	☐ Non-transformative or exact copy
☐ Restricted Access (to students or other appropriate group)	☐ Made accessible on Web or to public
☑ Nonprofit	☐ Profit-generating use

Overall, the purpose and character of your use ☑ supports fair use or ☐ does not support fair use.

NATURE OF THE COPYRIGHTED MATERIAL

Likely Supports Fair Use	Likely Does Not Support Fair Use
☑ Factual or nonfiction	☑ Creative or fiction
☑ Important to favored educational objectives	☐ Consumable (workbooks, tests)
☑ Published work	☐ Unpublished

Overall, the nature of the copyrighted material ☑ supports fair use or ☐ does not support fair use.

AMOUNT AND SUBSTANTIALITY OF MATERIAL USED IN RELATION TO WHOLE

Likely Supports Fair Use	Likely Does Not Support Fair Use
☑ Small amount (using only the amount necessary to accomplish the purpose)	☐ Large portion or whole work
☑ Amount is important to favored socially beneficial objective (i.e. educational objectives)	☐ Portion used is qualitatively substantial (i.e. it is the 'heart of the work')
☑ Lower quality from original (ex. Lower resolution or bitrate photos, video, and audio)	☐ Similar or exact quality of original work

LeEtta Schmidt, lmschmidt@usf.edu and Drew Smith dsmith@usf.edu
Reviewed by USF General Counsel 08/11/2015