

Engineering a “perfect” Mandarin: Language standardization as a means of creative agency in Taiwan’s dubbing recording

Spencer C. Chen^{}

Department of Anthropology, Hamilton College, 198 College Street, Clinton, NY, 13323, United States

*Corresponding author. Department of Anthropology, Hamilton College, 198 College Street, Clinton, NY, 13323, United States. E-mail: scchen@hamilton.edu

This article examines the production of Mandarin dubbing to explore how language workers in the movie execute and negotiate their creative decisions in ways that both manifest their idiosyncratic language ideologies and (re)affirm the dominant cultural beliefs about appropriate language expressions. Using video/audio recordings collected through a multilocal recording technique and off-record metapragmatic commentaries, I closely analyze the temporally unfolding moments in a recording session documented at a prominent studio in Taiwan in 2016. I argue that language work is a robust, collective process in which dubbing experts continuously shape and reshape their performance to align with the production and linguistic standards while simultaneously and discursively affirming and reaffirming the dominant cultural values represented by their work. Through this analysis, I also demonstrate that collateral and residual activities generated during the work production provide insights as significant as practitioner discursivity into the nature of linguistic creativity in language work and how it is executed behind the scenes.

Introduction

In October 2016, the Mandarin dubbing of a preview trailer for the Japanese anime *Pokémon the Series* reignited the decades-long moral panic over the “corrupted” Mandarin use in Taiwanese media. For weeks, anonymous commentators passionately spoke on nationally televised interviews and wrote on social media that Taiwanese people are failing academically and culturally because they are exposed to such “corrupted Mandarin dubbing (*lan pei yin*).” However, during this heightened moral panic, few were willing to admit that, although the trailer showed a misplacement of the standard voiceless sibilant *sh-* [ʃ] with its nonstandard variant, the voiceless alveolar sibilant *s-* [s], in a few dubbing lines, the actual episode in question was dubbed exclusively in Standard Mandarin. Intriguingly, despite the frustration directed at the behind-the-scenes language workers responsible for scripting, voicing, recording, and packaging the “corrupted dubbing,” no attempts were made to understand or explore their creative and practical reasons for

these linguistic choices. Could the misplacement of [s] be an intentional choice by the dubbing experts to challenge the authority of Standard Mandarin? Might it be a deliberate move to linguistically distinguish a character from its accent-free counterparts? Or could it simply be a recording error where the mispronounced sounds went unnoticed?

The moral panic, including the accompanying metapragmatic debates on- and offline, over language discussed here is nothing new, nor is it culturally unique by any measure. Indeed, applied linguists and sociolinguists have documented such language ideological reactions best described as “verbal hygiene” (Cameron 1995) towards incidences that disrupt the validity of a standard language in disparate locales and venues (Kroskrity 2000). Underneath these heated discourses of language impurity lie the ideological underpinnings of both language experts and the general public concerning how languages are socioculturally and political-economically imagined, construed, and manifested as indispensable of one’s moral self (Inoue 2004; Zhang 2017; Everhart 2021). Work in the interdisciplinary field of language standardization, as much as it abounds with accounts from language experts in domains such as school (Hubers et al. 2023), workplace (Woydack 2019), print media (Errington 1991), and broadcasting (Coupland 2001), urgently demands attention to the “invisible” behind-the-scenes language workers in media to understand how and whether they are surrogates of the state in “making and enforcing rules for language-use with the intention of reducing optional variation in performance” (Cameron 2000: 324; see also Milroy and Milroy 1998). Similarly, the scholarship on mediated language tends to conveniently, if not unfairly, regard the nonstandard expressions as intentional creative choices by the workers to evoke stereotypes associated with these linguistic features (Queen 2004; Lippi-Green 2011) or to exaggerate characterization in the dubbed text (Ferrari 2009; Iaia 2015), despite limited empirical evidence about the behind-the-scenes creative process these workers must navigate. What sets the Taiwanese case analyzed in this article apart is how language experts, in ways similar and yet more sophisticated than the nonexpert reactions reported above, proactively engage with, rather than mindlessly regurgitate, tenets of a language ideological apparatus I call Standard Mandarin Culture (SMC) that shapes how the Taiwanese populace (re)imagines what constitutes linguistic standardness.

Using video/audio recordings collected through a multilocal recording technique and ethnographic interviews conducted at a dubbing studio in Taipei in 2016, this article strives to fill the gaps in the scholarship on two levels based on the premise that language workers in media exercise their linguistic decisions with greater agency than the current scholarship acknowledges. Theoretically, it argues that language workers do not mindlessly reproduce the dominant language ideology but instead actively construct, debate, and negotiate what counts as “standard” and discard what does not. Their creativity is neither a mere reaffirmation of dominant cultural beliefs nor an uninhibited expression of personal creativity. Rather, the analysis below will demonstrate that it is a robust, collective process through which dubbing experts continuously shape and reshape their performance to align with the production and linguistic standards, while also affirming and reaffirming the dominant cultural values represented by their work. Methodologically, this article argues that research on language work can benefit not only from the primary language product, such as dubbing, but also, more importantly, from the collateral and residual activities generated during its creation. Typically regarded as disposable “waste, a useless byproduct” by both researchers and practitioners (Condry 2013: 13), off-record comments and informal conversations during production reveal workers’ creative, ideological, and other motivations that they might otherwise consider too trivial to report and reflect on during traditional interviews. As the analysis below will show, when juxtaposed with practitioner reflexivity (Thurlow 2020), these comments provide significant insights into what linguistic creativity entails and how it is executed behind the scenes.

Thus, in the following sections, I begin by reviewing the scholarship on language standardization, situating it within Taiwan’s complicated history with Standard Mandarin and its implementation as the official language since 1946, during which dubbing experts played significant roles. I then provide an ethnographic description of contemporary dubbing production to contextualize

the data collected during a recording session in Taipei in 2016. Next, I analyze the temporally unfolding processes of the recording session, focusing on how three dubbing professionals attempt to merge segments from four unusable retakes into a linguistic composite that reaffirms their version(s) of “perfect” Mandarin. Finally, I summarize the major findings and extrapolate the theoretical and practical implications of this research for understanding the roles of behind-the-scenes language workers in the movie.

Theorizing “SMC” in Taiwan

Standard(ized) language, as a (by)product of centralized nation-states, results from a series of language regimentation through state institutions and enables its speakers to connect themselves concretely and abstractly to their affiliated language community (Silverstein 1996). Sociolinguist James Milroy further abstracts these belief systems and social mechanisms as “standard language culture” and writes,

“Certain languages, including widely used ones such as English, French, and Spanish, are believed by their speakers to exist in standardized forms, and this kind of belief affects the way in which speakers think about their own language and about ‘language’ in general. We may say that speakers of these languages live in **standard language cultures**” (2001: 530, emphasis original).

The scholarship on language standardization adopts a *diachronic* approach to various “standard language cultures” in order to understand how the standardizing process serves as social engineering through language (e.g. Anderson 1991; Bauman and Briggs 2003; Zhang 2017). Such state-supported regimentation through language regulates specific modes of communication such that only users of the rectified discursive practices are granted authority in cultural, social, and political-economic domains (Woolard 1985; Bourdieu 1991). While the existing literature offers an instructive model to approach language standardization diachronically, we still have much to understand regarding the ways in which subjects of standardist hegemony respond to its top-down forces from the ground. To be clear, there are abundant accounts of how the linguistically oppressed navigate standardization (Rosa 2019). For instance, in his ethnography of Taiwanese multigenerational families, Sandel (2003) provides an ethnographic glimpse into how Mandarin-dubbed programs exacerbated children’s linguistic shift from other Sinitic languages to Standard Mandarin.

What I see as lacking is a *synchronic* account of how the linguistic majority—people who are “by definition” native to the standard variety and are thus “by assumption” the least affected—participates in the standardist hegemony on an everyday basis. From a Gramscian perspective (Gramsci 1971), top-down forces alone are insufficient for any “standard language culture” to subsist as effective and relevant. It also requires the involved social actors to participate in the hegemonic cultural logic of standardization. Under these conditions, language experts, with authority bestowed directly from their linguistic prowess and indirectly from the state recognition, exercise great influence over the public on what linguistic expressions constitute what is morally acceptable. Of all effective practices to perpetuate the standardist culture, “commoditization” (Silverstein 1996: 290) of the standard language has been widely practiced because it materializes the selected linguistic form as “the object of a brisk commerce in goods-and-services for which experts make themselves available (authorities or connoisseurs in the humanistic social fields, therapists/doctors or technicians in the scientific ones, etc.) and advertisers become focal personnel of personal decision-making to acquire the desired commodity for a price” (ibid: 291). Thus, in the same way that compiling physical dictionaries for Standard Indonesian has strengthened the sense of cultural belonging among diverse ethnic groups in Indonesia (Errington 1991), codifying Received Pronunciation through print and broadcasting media also has allowed the British to (re) construct a cultural imagination of nationhood as well as moral citizenship at different historical moments (Agha 2007).

Taiwan is no exception to such language standardist logic and its commoditized forms. With a colonial history dating back to the sixteenth century, Taiwan's most recent transition from Japanese rule to the Chinese Nationalist regime in the post-WWII era has yet again significantly reconfigured her sociolinguistic landscape. In particular, Standard Mandarin (known as *Guoyu*), which was essentially a foreign language to the then Taiwanese populace, replaced Japanese as the official and the High variety (Kubler 1981). To facilitate the transition, the state enforced a series of language policies and programs, mandating Standard Mandarin as the sole legitimate medium of communication from 1946 to 1987 (Wei 2008). Meanwhile, it also enlisted private sectors such as the radio and the newly established televised broadcasting networks to further the language regimentation (Sandel 2003). Thus, despite the relaxed state regulations and changing public opinions favoring multilingualism in recent decades, the public continues to unconsciously regard Standard Mandarin as the default medium of communication (Chun 2017: 57–74). This complex web of state and private forces has created a long-lasting sociolinguistic consciousness that treats Standard Mandarin as hierarchically most prestigious over nonstandard forms of Mandarin and other non-Mandarin languages.

The sociolinguistic work on Taiwan holds a critical stance towards such ideological apparatus, which I dub “SMC,” as the direct result of the state enforcement and examines the policy's socio-cultural and political-economic impacts on people from various linguistic backgrounds (Kubler 1981; Her 2012). Coined to capture the pervasive influence of Standard Mandarin in Taiwan, the term Standard Mandarin Culture describes the (un)conscious orientation towards language and cultural politics, where one's self, identity, disposition, status, and role are contingently measured by their displayed competence in Standard Mandarin at any given interactional moment. The SMC indexes not only linguistic competence but also broader sociopolitical alignments, where language proficiency is (in)directly intertwined with cultural identity and social mobility. Similar to Received Pronunciation and its symbolic status in the UK (Agha 2007), Standard Mandarin (sometimes only selected features of it) has been utilized as positive/negative identity markers by different groups in interactions (e.g. Wei 2008; Lin 2009; Su 2019). At moments when the spotlight is placed on language, people would (re)voke the discourse of standardness, as we have seen above regarding the dubbing for *Pokémon*. However, very few could clearly define Standard Mandarin. At best, they may identify the “popular culprits” of nonstandard Mandarin, among which the absence of retroflex sibilants receives the most attention (e.g. Hsu and Tse 2007; Lin and Chan 2021). Inasmuch as one is pressed to specifically define standardness, they would rely heavily on stripping off layers of features that they ideologically deem nonstandard before reaching an ambiguous, highly ideologized notion of Standard Mandarin. This semiotic process, or rather, exclusionary ideologization of standardness, robustly sustains the SMC. In this regard, for the Taiwanese, Standard Mandarin represents more of an ideological phantasm than a linguistic actuality. In particular, based on corpus and cross-generation sociophonetic surveys of Mandarin speakers, linguists have postulated that Standard Mandarin has never truly been spoken as a native language by the Taiwanese populace from the start, irrespective of their ethnonational origin (Her 2012, 2014; Hsu 2014). Instead, Standard Mandarin as a concept serves as a discursive means to transcend temporal and spatial specificities and delineate what is considered lingua-culturally Taiwanese as opposed to Chinese (e.g. Chen 2018; Su 2019). This conflicting stance towards Standard Mandarin and its associated indexicalities underscores the significant influence the SMC continues to exert on the Taiwanese populace's perception of language in media. Generally, the prevailing view does not critically question why or how Standard Mandarin remains the default media register in the contemporary mediascape, unlike the scrutiny applied to linguistic minorities (Lin 2009; Friedman 2018).

In the Taiwanese context, dubbing, or *peiying* (lit. “to add voice”) as locally known, is a media practice and profession specialized in the (re-)recording of human voices, music, and sound effects, and (re-)mixing of them into soundtracks for audiovisual entertainment products such as cartoons, films, television programs, video/mobile games, and public announcements. The making of Mandarin dubbing of all communicative events exemplifies both figuratively

and literally the very processes through which Standard Mandarin becomes not only a means of communication (i.e. Mandarin for content comprehension) but also a form of commodity (i.e. dubbing for media consumption) within Taiwan's SMC. In part due to the industry's inception in the 1960s as a state-supported media postproduction sector to promote Standard Mandarin outside schools and state institutions in less intrusive ways (Chen 2022), language workers in this profession—including but not limited to translators, script adapters, voice actors, sound engineers, quality controllers, and market sales representatives—have not only taken the responsibility themselves but have also been held responsible by the state and the public for language expressions in media. Unpacking the dubbing production thus provides an interesting point of entry for understanding both diachronic and synchronic processes of language standardization. For one, metapragmatic discussions on set between dubbing professionals over what constitutes Standard Mandarin underscore the ideological salience and relevance of the SMC through historical transmission. On the other hand, dubbing is a spontaneous exclusionary process at the very moment of its production that erases unwanted parts from the soundtracks, which in turn further consolidates and perpetuates the effects of the SMC. In what follows, I first describe several context- and profession-specific practices of dubbing language workers and subsequently situate these off-screen practices within the broader context of the SMC.

The data: Mandarin dubbing recording in Taiwan

The data for this article comes from my longitudinal ethnographic research in Taipei since 2016. In this research, I have worked both as an ethnographer and practitioner, closely examining first-hand various production sectors of the dubbing industry and their collaborative creative practices. In addition to theorizing collaborative creativity in linguistic work, I consciously adopted an application-committed approach to understanding “the people doing the work of translation—speaking the languages, reading and writing the texts, and making the interpretive choices that create a basis for the way the texts are understood” (Pritzker 2012: 344).

Dubbing, like other forms of performance mediated through language, has often been regarded as the linguistic endeavor and decision of a single individual who decides what and how to translate the source text, be that person a translator, voice actor, director, or creative executive (Lippi-Green 2011). Much against such impression, however, dubbing is a distributed, collaborative, and multiparty performance (also known as “voice acting”) and production (also known as “voiceover” and “voice-over”). For a film to be dubbed into the target language, the source film must undergo a lengthy production process that requires the involved professionals to contribute respectively and partially to the final commodified linguistic expressions. Within academe, dubbing has conventionally and extensively been examined as “texts” to address issues of translatability between the source and target texts and the linguistics of dubbed dialogs (see Díaz Cintas and Anderman 2009). Without a doubt, dubbing involves intersemiotic translation, particularly during the initial stage when translators and script adapters produce dubbing scripts based on the source material. However, when the dubbing recording begins, concerns over translatability and textual fidelity are often greatly, if not completely, overshadowed by the practical needs to ensure intelligibility, vocal expression, and lip-synchronization. As my research and a growing number of interdisciplinary studies have shown (e.g. Sherouse 2015), attending to forms of language work other than interlingual transfer contributes to a much-needed “cultural turn” that would shed light on the “[i]deology, and concepts of power, otherness, patronage, resistance, as well as gender and post-colonial issues” involved in dubbing (Chaume 2014: 297). This cultural turn allows us to move beyond the translational microcosm to discern more broadly “the ‘macrocosm’ of broad-based cultural processes and social ‘facts’” that language workers must navigate (Pritzker 2012: 344).

For the purpose of this article, I narrowed my scope to focus on voice recording at dubbing studios, an institutional setting where professionals turn text into soundtracks in a collaborative manner (as opposed to the recording conducted by individual professionals at home). Nowadays, as Fig. 1

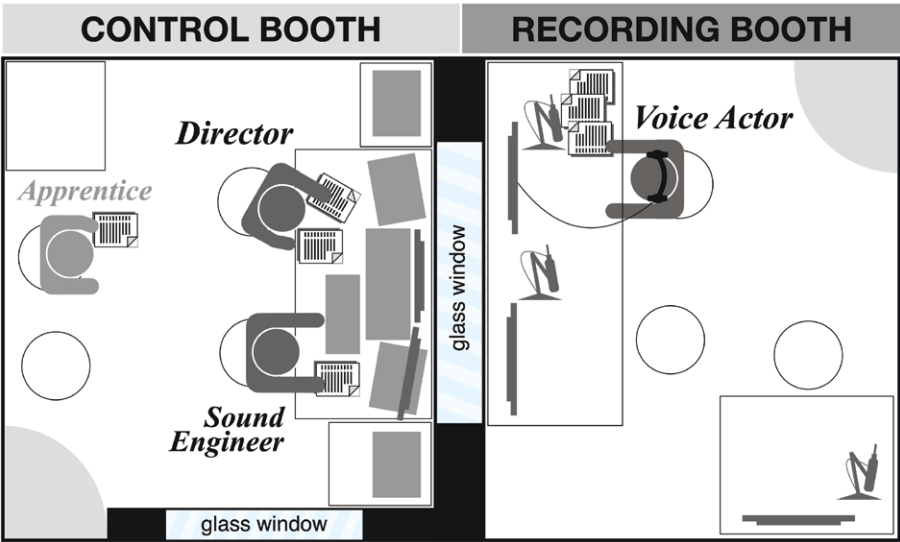
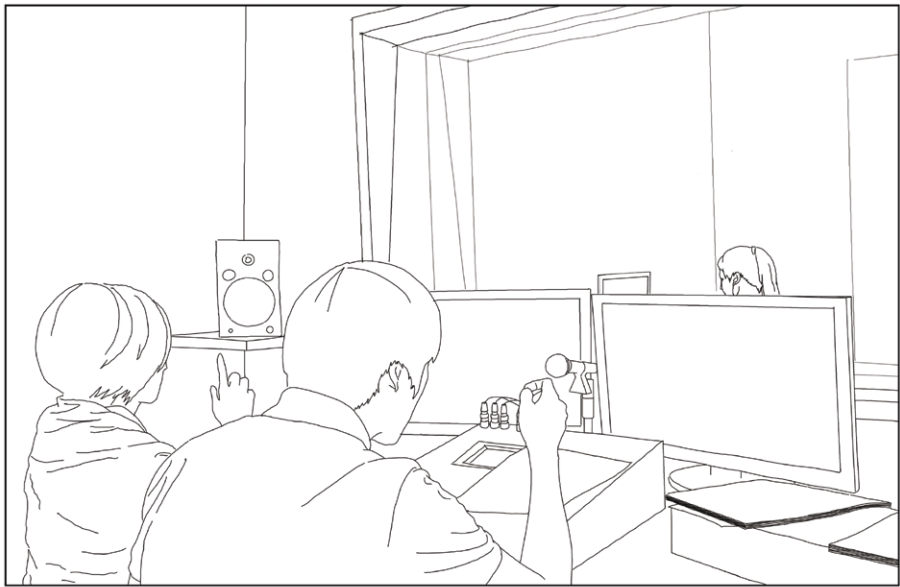


Figure 1. A typical scene of an “individual recording” session in a recording studio. The drawing on the top captures the perspective of an apprentice on how the recording unfolds in front of them, whereas the layout below indicates the built environment with which these professionals interact.

demonstrates, a dubbing recording session in Taiwan typically involves only a handful of people occupying two small enclosures, with a voice actor in the recording booth and a director (who is also a voice actor on the project), a sound engineer, and occasionally an aspiring voice-acting apprentice (who may serve as the director’s additional pair of eyes and ears) in the control booth. The control booth has full access to everything generated in the recording booth (such as speech and noise) through loudspeakers and can freely communicate with the recording booth through the intercom. In contrast, the recording booth must rely on the microphones and headphones to communicate

with the control booth. The control booth often unilaterally disables the intercom to either avoid interrupting the recording or make comments on the actor’s performance. Commonly known as “individual recording (*danshou*)” within the industry, this small-group format allows a voice actor to come in individually to the studio and focus only on their parts without the need to coordinate with their coactors. It also allows the sound engineer to concentrate entirely on recording and editing this actor’s soundtrack before mixing it with the soundtracks of other actors.

At any point during the recording, the involved dubbing professionals can edit details as small as a word to ensure the final result sounds “normal (*zhengchang*)” and “clear (*qingchu*).” On the surface, notions of “clarity (*qingchu*)” and “normalness (*zhengchang*)” in dubbing seem much less about standardness than about sound quality and auditory perception. A question about clarity explicitly problematizes one’s dubbing attempt. It immediately impels them to produce a “clearer” version in place of the “muffed” one. Yet, in many cases, such questions are in fact intended as indirect requests to remedy mistakes in pronunciation or syntax without overtly challenging one’s Mandarin proficiency. The discourse of linguistic normativity (Woolard 1986; Cameron 1995) prompts the wrongdoers to abide by the socially acceptable rules of Mandarin and to rid their dubbing expressions of nonstandard elements. It pulls these professionals into a metadiscursive sphere where everyone present on set is confronted by and must adhere to the notion of standardness.

Below, I concentrate on a four-hour recording session for a Japanese anime, which I observed as an ethnographer and aspiring apprentice in September 2016 at a prominent studio in Taipei. This session involved three dubbing professionals occupying two booths (see Fig. 1 for visualization): a voice actor in the recording booth and a voice director and sound engineer in the control booth. To reliably capture the concurrent discursive and nondiscursive actions within the studio setting, I set up separate video and audio recorders in the control and recording booths. I then transcribed the recorded data using a two-column transcription system developed specifically for this setting (see Note 1). By cross-examining excerpts from postrecording interviews and unstructured discussions with the involved professionals, I analyze the temporally unfolding moments in the recorded interaction where the actor, director, and sound engineer adhere to and operate within the logic of the SMC. Throughout this process, they metadiscursively appraise, negotiate, and refine their versions of “Standard Mandarin” until they transform discarded audio files of “nonstandard” sentences into satisfactorily rendered sentences that are both linguistically and ideologically “standard.”

Erasing dubbing mistakes through rerecording

In this four-hour recording session to be analyzed below, the dubbing crew sets forth to record six episodes (thirty minutes each) in which the voice actor dubs various characters. Roughly 45 min into the recording, as the actor concentrates on his main character “Fuqin (lit. father)” without much interruption, the voice director abruptly stops him with a request to fix a line he dubbed several tens of minutes earlier. The line in question reads, “Women buneng pian ta. Pian name chunzhen de ren shi hui zao tianqian de (We cannot fool him. We will get cursed if we trick such an innocent guy).” Fig. 2 reproduces the original printed script—with my analytic descriptors on the top and Pinyin transliteration and English translation at the bottom—that the voice actor holds during the recording.

ROLE	TIMECODE	SCRIPT LINE
focus of attention		
父親	1832	我們不能騙他 因為騙那麼純真的人是會遭天譴的
Fuqin	1832	women buneng pian ta. ~~因為~~ pian name chunzhen de ren shi hui zao tianqian de
Father		We cannot fool him. <i>Because</i> we will get cursed if we trick such an innocent guy.

Figure 2. Reproduction of the printed script line handed to the voice actor.

This line by itself is a simple sentence so common in dubbing scripts that would not normally warrant dedicated attention. Or rather, it was not supposed to, as the voice actor would reveal to me in the postrecording interview after the four-hour recording session. The only noticeable “oddity” (if anything) is the crossing-out of the word “yinwei (because)” with a black-color marker made by the director’s apprentice prior to the recording session. The omission of the subordinating conjunction “because” breaks the original script line, “We cannot fool him, because we will get curses if we trick such an innocent guy,” into two separate lines that are still pragmatically relevant while syntactically independent. In other words, this is more of an editorial preference than a linguistic necessity, let alone a standardist move to regiment linguistic expressions. With that said, this line not only became the object of verbal hygienic preoccupations (Cameron 1995: 27) but served as the plane through which the three involved professionals negotiated their standardist beliefs on what makes a “perfect” Mandarin.

Transcript 1 captures the complications that this line engenders (See [Appendix](#) for the transcription key) during the recording.¹ The director (coded as “Dir.”) does not clarify what she particularly sees as problematic in the actor’s original dubbing attempt in her initial request for redub. Since they are a couple of ten pages down from the problematized line when she initiates this request exchange, the actor (coded as “VA”) asks for clarification (01–02). Rather than spelling it out for the actor, the director simply proposes a remedied version in which she tacitly removes what she flags as an error (05–06). Furthermore, she gives the impression that the “real” problem lies in the actor’s inaccurate pronunciation of the word “ren (person)” (07–08). Subsequently, the director signals the sound engineer (coded as “SE”) to rewind and play his original dubbing recording (coded as “VA®”) for them to jointly inspect “the problem”

CONTROL BOOTH	RECORDING BOOTH
01 Dir. : ni xian bang wo gai qianmian na yi ju? <i>Could you redub an earlier line first?</i>	02 VA : nali? <i>Where?</i>
03 Dir. : 'WOMEN BUNENG PIAN TA' hougian <i>After the line 'WOMEN BUNENG PIAN TA,'</i> <i>[we cannot fool him]</i>	03 Dir. : 'WOMEN BUNENG PIAN TA' hougian <i>After the line 'WOMEN BUNENG PIAN TA,'</i> <i>[we cannot fool him]</i>
04 : bang wo jiang, <i>please say,</i>	04 : bang wo jiang, <i>please say,</i>
05 : 'PIAN NAME CHUNZHEN DE REN,' <i>'PIAN NAME CHUNZHEN DE REN,'</i> <i>[if we trick such an innocent guy]</i>	05 : 'PIAN NAME CHUNZHEN DE REN,' <i>'PIAN NAME CHUNZHEN DE REN,'</i> <i>[if we trick such an innocent guy]</i>
06 : 'HUI ZAO TIANQIAN DE!' <i>'HUI ZAO TIANQIAN DE!'</i> <i>[we will get cursed]</i>	06 : 'HUI ZAO TIANQIAN DE!' <i>'HUI ZAO TIANQIAN DE!'</i> <i>[we will get cursed]</i>
07 : Ni nage [zen ³⁵] zi shao- <i>And your [zen³⁵] ('person') is a bit....</i>	07 : Ni nage [zen ³⁵] zi shao- <i>And your [zen³⁵] ('person') is a bit....</i>
08 : biancheng [ren ³⁵] le. <i>It's become [ren³⁵].</i>	08 : biancheng [ren ³⁵] le. <i>It's become [ren³⁵].</i>
09 VA : hao a <i>Okay</i>	
10 Dir. : hou, 'WOMEN BUNENG PIAN TA, YINWEI,' <i>Hmm, 'WOMEN BUNENG PIAN TA, YINWEI,'</i> <i>[we cannot fool him, because]</i>	10 Dir. : hou, 'WOMEN BUNENG PIAN TA, YINWEI,' <i>Hmm, 'WOMEN BUNENG PIAN TA, YINWEI,'</i> <i>[we cannot fool him, because]</i>
11 : hougian na yi ju jiu hao. <i>After this line, redo the next line.</i>	11 : hougian na yi ju jiu hao. <i>After this line, redo the next line.</i>
12 : 18:32 nabian. <i>It's at the timestamp 18:32 there.</i>	12 : 18:32 nabian. <i>It's at the timestamp 18:32 there.</i>
13 : ((points at 18:30 on SE's monitor))	

Transcript 1. A gloss of abbreviations: Dir. (director); SE (sound engineer); VA (voice actor); VA® (recorded soundtrack of VA's dubbing performance).

CONTROL BOOTH	RECORDING BOOTH
14 SE : ((rewinds to 18:30))	
15 ((plays rec. of VA's first attempt))	
16 VA® : { women buneng pian ta. } { <i>We cannot fool him.</i> }	16 VA® : { women buneng pian ta. } { <i>We cannot fool him.</i> }
17 { pian name chunzhen de [lən ³⁵], } { <i>If we trick such an innocent guy.</i> }	17 { pian name chunzhen de [lən ³⁵], } { <i>If we trick such an innocent guy.</i> }
18 { shi hui zao tianqian de! } { <i>we will get cursed</i> }	18 { shi hui zao tianqian de! } { <i>we will get cursed</i> }
19 Dir. : ((stops SE and presses the intercom))	
20 dui, zhe ju. Hou. Women cong, <i>Yes, this line. Alright. Let's start,</i>	20 Dir. : dui, zhe ju. Hou. Women cong, <i>Yes, this line. Alright. Let's start,</i>
21 'BUNENG PIAN TA' houmian jie. <i>after the line 'BUNENG PIAN TA.'</i> [cannot fool him]	21 'BUNENG PIAN TA' houmian jie. <i>after the line 'BUNENG PIAN TA.'</i> [cannot fool him]
22 'PIAN NAME CHUNZHEN DE [zən ³⁵],' <i>Just say 'PIAN NAME CHUNZHEN DE [zən³⁵],'</i> [If we trick such an innocent guy.]	22 'PIAN NAME CHUNZHEN DE [zən ³⁵],' <i>Just say 'PIAN NAME CHUNZHEN DE [zən³⁵],'</i> [If we trick such an innocent guy.]
23 'HUI ZHAO TIANQIAN DE!' jiu hao le <i>'HUI ZHAO TIANQIAN DE!' That's all.</i> [we will get cursed!]	23 'HUI ZHAO TIANQIAN DE!' jiu hao le <i>'HUI ZHAO TIANQIAN DE!' That's all.</i> [we will get cursed!]
24 bu yao jiang 'SHI' le. Xiexie. <i>Don't add 'SHI' anymore. Thanks.</i> [be]	24 bu yao jiang 'SHI' le. Xiexie. <i>Don't add 'SHI' anymore. Thanks.</i> [be]
25 SE : ((rewinds to 18:30 and hits RECORDING))	

Transcript 1. Continued

(12–24). Upon listening to the actor's original take, she finally identifies what she is essentially problematizing throughout the entire exchange, i.e. the copula “shi (be)” (20–24).

Up until the point in line 24, where the director finally reveals her true request about dropping the copula “shi (be),” she has been direct yet simultaneously indirect about how she desires the word “ren (person)” to be pronounced. At first, she does not discursively comment on it; instead, she offers an unfinished assessment of the actor's original pronunciation of *ren* (07–08). Subsequently, in a matter-of-fact tone, she juxtaposes her demonstratively standardist pronunciation of *ren* [zən³⁵] with his nonstandard version of *ren* [rən³⁵] to pinpoint what the issue is. While common in Taiwan Mandarin (Chen 1999), the use of the alveolar lateral approximant [l] (or [z]) in place of the retroflex /ɾ/ (realized at any point on the spectrum between the approximate [ɹ] and fricative [ʐ]) is often metadiscursively marked as nonstandard and highly discouraged in the dubbing industry.

The way in which she constructs this request sequence is of great interest to understanding how influential and effective the SMC is at alerting one to be hyper-critical of their own (as well as others') linguistic expressions. Here, by withholding her full assessment of the trouble source (07), the director initiates the interactive framework that Jefferson (1987) calls “exposed correction,” a format that demands coparticipants to treat the problem as “a matter of collaborative, step by step construction that a correction will be an interactional business in its own right, with attendant activities addressing issues of competence and/or conduct” (1987: 99). The director's tacit replacement of the actor's highly stigmatized nonstandard use of the lateral approximant [l] with a less stigmatized alveolar tap [ɾ] (07–08) does not simply make her request less face-threatening. In fact, coupled with the joint review of the actor's original recording (16–18), which serves as evidence to hold him accountable for his “wrongdoing,” as well as the director's subsequent emphatic demonstration of the “correct” pronunciation (22), her imitation metapragmatically delimits the collaborative framework of correction. That is, at this point, the director has successfully established that her request to redub the line is entirely motivated by the standardist construct rather than creative concerns. Unlike in a regular interactional sequence where the



Figure 3. Reproduction of the actor's scribbles on the printed script line.

problematizee can either accept or reject the correction, the actor is interactively tasked to remedy the exposed form to avert challenges to his Mandarin competence.

As we see below in [Transcript 2a](#), the actor is now so preoccupied with his rendition of *ren* for the upcoming retake that he is caught off guard when the director abruptly makes her second exposed correction to remove the word *shi* for syntactic brevity (24). As [Fig. 3](#) shows, the actor only has enough time to quickly circle and underline *ren* and put an X mark above *shi* before the sound engineer rewinds to the starting point to begin the rerecording.

The director may consider it a relatively simple fix that merely requires both the actor to swiftly revoice the line in question and the sound engineer to laminate the newly recorded part on the existing soundtrack. The actor is supposed to voice the newly revised line exactly as the director instructs, namely, “*pian name chunzhen de ren hui zao tianqian de* (If we trick such an innocent guy, we will get cursed).” As it turns out, however, this supposedly easy fix is not as simple for the actor to accomplish as the director initially thought. As soon as the rerecording begins, the director's ambiguous directive regarding what needs to be amended immediately causes the actor to stumble. In each of the four consecutive retakes, the actor continues to fail to produce *ren* and remove *shi* as satisfactorily and precisely as the director expects. Seeing that the actor cannot complete both requests after four retakes, the director eventually decides to ignore the line and moves on. From a practical point of view, these retakes are nothing but “production waste” that generates zero value of any kind. Nevertheless, from an applied linguistic viewpoint, these retakes—and the accompanying discursive and nondiscursive practices—elucidate the ways in which the involved experts adeptly balance their linguistic creativity with the standardist logic. [Transcript 2a](#) presents how the actor struggles to dub the newly revised line in his first two attempts.

In his first retake (coded as VA¹ in 27–30), as soon as he finishes the first clause and begins to vocalize the second one, the actor realizes that he unconsciously retains *shi* (28). He hesitates a little at first and then starts to practice saying the second half without *shi* (29–30). What he does not seem to realize is that his rendition of the phonology of “*ren*” with an alveolar flap [r] does not exhibit the retroflexion that the director is after. Noticing his struggle, the sound engineer pauses to wait for him before the director instructs. Interestingly, notwithstanding his early stumble with *shi*, the actor practices pronouncing *ren* [ʐən³⁵] repeatedly (32). Utilizing the actor's two self-initiated repairs (28–30 and 32, respectively) as an interactional solicitation for corrective suggestions ([Jefferson 1987](#)), the director first acknowledges his noticing *shi* with a simple “yes (*dui*)” and subsequently cautions him about his *ren* by issuing another overly exaggerated, “proper” version [ʐən³⁵] (34).

As the sound engineer begins to rerecord for the second time, it becomes clear that the actor is too preoccupied with dropping *shi* to pronounce *ren* precisely as the director demands. In this second voiceover attempt (coded as VA² in 37–38), he successfully removes *shi* but unsuccessfully reads *ren* with an alveolar flap. This time, the director adopts another corrective strategy to approach the unresolved issue. Instead of openly commenting on his retroflexion, she justifies yet another retake with the claim that the second dubbing attempt fails to synchronize well with the character's mouth movements (40–41). Although the delay of accusation evidenced in the director's approach may productively create “an interactional environment which minimizes the production of arguments” ([Garcia 2019: 60](#)), it may also unproductively delay the resolution of the dispute, particularly in such a case where the task is to resolve a “disputed” linguistic expression as revealed below in [Transcript 2b](#). Here, quite

CONTROL BOOTH	RECORDING BOOTH
25 SE : ((rewinds to 18:30 and hits RECORDING))	26 VA® : { women buneng pian ta. } { We cannot fool him. }
26 VA® : { women buneng pian ta. } { We cannot fool him. }	27 VA ^{v.1} : ≈ pian name chunzhen de [rən ³⁵], ≈ ≈ pian name chunzhen de [rən ³⁵], [If we trick such an innocent guy.]
	28 ≈ shi hui tian-≈ ≈ shi hui tian-≈ [we will get cur-]
	29 ei::, shenme? a, o! Uh, what? ah, oh!
	30 ≈ Hui- hui- hui zao tianqian. ≈ ≈ hui, hui, hui zao tianqian. ≈ [will, will, will get cursed]
31 SE : ((hits pause w/o making comments))	32 VA : ((pauses; starts to practice “{zən ³⁵ }”))
33 Dir. : ((turns on the intercom))	34 Dir. : dui. nage ^[zən ³⁵], xiaoxin o Yes. And watch out your ‘{zən ³⁵ }’
34 dui. nage ^[zən ³⁵], xiaoxin o Yes. And watch out your ‘{zən ³⁵ }’	36 VA® : { women buneng pian ta. } { We cannot fool him. }
35 SE : ((rewinds to 18:30 and hits RECORDING))	37 VA ^{v.2} : ≈ pian name chunzhen de [lən ³⁵], ≈ ≈ pian name chunzhen de [lən ³⁵], ≈ [If we trick such an innocent guy.]
36 VA® : { women buneng pian ta. } { We cannot fool him. }	38 ≈ hui zao tianqian de! ≈ ≈ hui zao tianqian de! ≈ [we will get cursed!]
39 Dir. : ((stops SE and turns on the intercom))	40 Dir. : keneng yao zai kuai yidiandian. You might need to speak faster.
40 keneng yao zai kuai yidiandian. You might need to speak faster.	41 You dian guo le. It’s a bit off [i.e., a failed lip synchronization].
41 You dian guo le. It’s a bit off [i.e., a failed lip synchronization].	

First
Retake

Second
Retake

Transcript 2a . A gloss of abbreviations: Dir. (director); SE (sound engineer); VA (voice actor); VA® (recorded soundtrack of VA’s dubbing performance); VA^{v.1} (first attempt); VA^{v.2} (second attempt).

intriguingly, after rewinding to the starting point, the sound engineer silently creates a new separate track to record rather than erasing the second attempt (42–43) as he did with the first take (35 in [Transcript 2a](#) above). As it turns out, this act is, in fact, the engineer’s tacit assessment of and participation in the linguistic ideological dispute over linguistic standardness. In his third reattempt (coded as VA^{v.3} in 45–46), the actor becomes very careful with the retroflex [ʒ] in *ren* but still flounders with the *shi* clause. In the process of voicing the clause, he not only retains the copula *shi* but also accidentally omits the main predicate “*zao* (encounter),” resulting in an ungrammatical clause “*shi hui tianqian* (*[we] will curse [noun])” (46). Aware of the syntactic mistake, the actor quickly halts and begins to rehearse the *shi* clause several times (48). Meanwhile, as observable in line 47, the sound engineer pauses the recording before the actor and the director notice the mistake. Interestingly enough, similar to what he did at the beginning of the third retake (43), the sound engineer again opened a new soundtrack to record the upcoming fourth retake (49). In the actor’s fourth and final take (coded as VA^{v.4} in 51–52), he generates the standard retroflex fricative [ʒ] as requested (51) and smoothly voices the second clause without any hesitation. In this version, the copula *shi* is retained in his final version (52). For some reason, however, the actor does not initiate self-repair or seem interested in “fixing” it as he was in the previous three retakes. Frustrated with the unsuccessful results from the invested

CONTROL BOOTH	RECORDING BOOTH
42 SE : ((rewinds to 18:30))	
43 ((opens a new track to record))	
44 VA® : { women buneng pian ta. } { <i>We cannot fool him.</i> }	44 VA® : { women buneng pian ta. } { <i>We cannot fool him.</i> }
	45 VA ^{v3} : ≈ pian name chunzhen de [^] [ʒən ³⁵] ≈ ≈ <i>pian name chunzhen de</i> [ʒən ³⁵] ≈ [<i>If we trick such an innocent guy</i>]
	46 ≈ shi hui tianqian- ≈ ≈ <i>shi hui tianqian-</i> ≈ [<i>we will curse-</i>]
47 SE : ((pauses and rewinds back to 18:30))	
	48 VA : ei:: shi- shi- shi hui. um. <i>Eh::, shi, shi, shi, hui... Um.</i>
49 SE : ((opens another new track to record))	
50 VA® : { women buneng pian ta } { <i>We cannot fool him</i> }	50 VA® : { women buneng pian ta } { <i>We cannot fool him</i> }
	51 VA ^{v4} : ≈ pian name chunzhen de [ʒən ³⁵], ≈ ≈ <i>pian name chunzhen de</i> [ʒən ³⁵], ≈ [<i>If we trick such an innocent guy.</i>]
	52 ≈ shi hui zao tianqian de! ≈ ≈ <i>shi hui zao tianqian de!</i> ≈ [<i>we will get cursed!</i>]
53 Dir. : ((shakes heads; turns off the intercom)) ((stops SE and speaks to SE directly))	
54 hao ba. xian zheyang, xian guoqu. <i>Alright. Let it be. Let's pass it for now.</i>	
55 dengdeng zai yi. <i>We'll edit this later.</i>	
56 SE : ((nods; rewinds to another timestamp))	

Transcript 2b. A gloss of abbreviations: Dir. (director); SE (sound engineer); VA (voice actor); VA® (recorded soundtrack); VA^{v3} (third attempt); VA^{v4} (final attempt).

time and efforts, the director turns off the intercom and reluctantly tells the sound engineer to forgo the issue and proceed to finish the actor's remaining lines.

The entire rerecording session may appear as nothing but a series of corrections that demonstrates the director's linguistic authority on standardness as well as the actor's and the engineer's willing submission to the SMC. I maintain that, however, practitioner exchanges such as the rerecording documented here—all its encompassed synchronically unfolding actions (and inactions), whether verbal and nonverbal—presents an invaluable perspective on *what and how* language workers in media do with language that is too often erased in the scholarship on mediated language use, due perhaps to the accessibility issue in the first place (Thurlow 2020). By detailing the work conditions and the invested efforts of the three experts to adjust a sentence that viewers may simply overlook, I demonstrate that media workers *do* make conscious decisions—be they linguistic, aesthetic, or practical—and are very self-aware of the potential effects their choices may bring forth. More importantly, while the scholarship on language workers prioritizes discursivity (such as work discourses or self-reports) (Duchêne 2009), my closed-up analysis of the rerecording also indicates that decisions may be made *without* discursive reflexivity. In the next section, I attend to the moment that immediately follows the rerecording session to discuss the engineer's intriguing act of creating a separate audio recording (lines 43 and 49) instead of erasing the selected part from the existing master soundtrack. The two seemingly “trivial” instances substantively speak volumes of the engineer's important—often quiet and unnoticed—contributions to the making of Mandarin dubbing, despite his absence in the discursive commentary on language expressions throughout the entire recording session.

Crafting the perfect dubbing from scrapped recordings

Thirty minutes later, when the actor finished recording his characters in the first three episodes, he requested a smoke break before proceeding to finish the other three episodes. As he walked out to the balcony to enjoy his ritual postrecording cigarette, I joined him for a casual small talk as usual. Out of the blue, he surprised me with an unprompted commentary about the *ren*-and-*shi* rerecording.

“Sure, that (first) [ʒən³⁵] was indeed poorly pronounced. It was *unclear* (*bu qingchu*). But honestly, there was actually no need to change the next clause. It’s basically the same as saying either ‘*shi hui zao tianqian de*’ or ‘*hui zao tianqian de*’! They all sound very *normal* (*zhengchang*)! It’s not like there were too many words that we had to omit some [i.e., for better lip-synchronization]. I could finish the line in time. So why bother insisting on cutting the word? That’s why I didn’t change that line in the final take.”

He acknowledged the necessity of paying attention to *ren* since it was “unclear (*bu qingchu*).” But his frustration came from the director’s insistence on dropping *shi*. For him, the difference between the original script line with *shi* and the director’s version without *shi* has nothing to do with grammaticality (see Fig. 4).

Both versions convey the exact same message, “If we fool the person, we will get cursed.” The original script adopts the emphatic construction “*shi... de*” to stress the conditional consequence, namely, “*hui zao tianqian* (we will get cursed).” However, *shi* is syntactically optional in the “*shi... de*” construction. In other words, the director’s version is essentially identical to the original script line, not only denotationally but also structurally speaking. Omitting the word from the line is more of a personal preference than a grammatical necessity, let alone a standardist ideological move.

Of course, as the actor pointed out, there are cases when such omission, or for that matter, the addition of words, is a simple but efficient strategy to achieve proper lip-synchronization. Such is the case that, for instance, when a character on the screen is shown to open and close its mouth five times, its dubbing would prefer a five-word Mandarin line over, say, a seven-word one. Thus, at first glance, this could have been the unstated practical reason why the director asked for the retake in the first place (lines 01–08). Yet, in practice, the actor had no trouble matching his dubbing performance with the character’s mouth movements. Excluding all unfinished retakes (the first one, lines 27–30, and the third one, lines 45–46), he was able to fit the “longer” version in both his original take (lines 17–18) and the fourth retake (lines 51–52). His second retake (lines 37–38), which is a word shorter than the script line, has also achieved successful lip-synchronization. Still, the director described it as a failed mouth-matching and thus demanded yet another retake. As I have described above, her intention was not to ensure better lip-matching but to eliminate the “mispronounced” *ren* without any explicit commentary on the standardness.

After my quick chat with the actor on the balcony, I returned to the control booth to chat with the sound engineer as we waited for others to return. The director shortly joined us in the control booth. Knowing that the engineer had saved all the retakes from the last session as separate soundtracks, she promptly asked if they could jointly inspect the audio files right away. The engineer quickly played the second retake and the fourth one. Upon listening through two versions

Original Ver.	pian name	chunzhen	de	ren,	shi	hui	zao	tianqian	de!
	fool	that	innocent	de-PTCL	person	be	will	encounter	curse-PN
Director’s Ver.	pian name	chunzhen	de	ren,	ø	hui	zao	tianqian	de!
	fool	that	innocent	de-PTCL	person	ø	will	encounter	curse-PN
“If we trick such an innocent guy, we will get cursed!”									

Figure 4. Two versions of the script line to be “engineered.” A gloss of linguistic terms: PTCL (particle); PN (proper noun); ø (null).

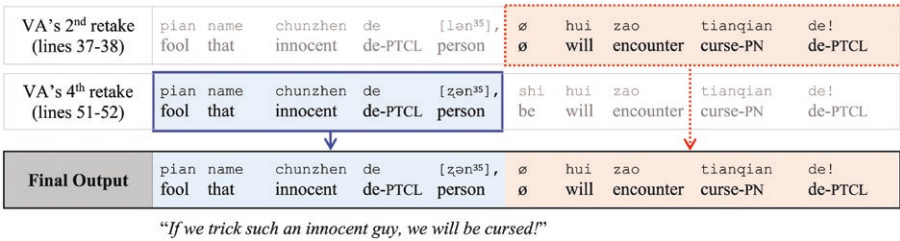


Figure 5. Creating the “perfect” dubbing line by copy-pasting the selected segments from the two retake versions. A gloss of linguistic terms: PTCL (particle); PN (proper noun); ø (null).

several times, the director suggested, “Let’s do a small surgery (*xiao shoushu*).” By “small surgery,” she was referring to another common strategy used in the industry—sound (re)mixing. While most dubbing professionals prefer redubbing to overwrite unwanted parts over sound mixing, some would rely on this strategy when the “problems” can be easily fixed by simple editing. In cases such as this particular one, where the voice actors cannot produce the desired result after many tries, they would make use of the mixing technique to produce the desired result. Fig. 5 shows their process of crafting the “perfect” dubbing line from the two selected retakes.

As Fig. 5 indicates, the second and the fourth retakes both contain a segment that the director wished to remove, i.e. the nonstandard *ren* [lən³⁵] and the inclusion of *shi*. Thus, after locating the short pause between the two clauses in each version, the sound engineer cut the first clause from the fourth retake (i.e. line 51 in the blue square) and the second clause from the second retake (i.e. line 38 in the red dotted square). On yet another separate soundtrack, he pasted the two “pieces” in the correct order and started “treating the surgical cut” by adjusting the values of their respective spectrograms. In a couple of seconds, he presented us with the result of this operation with a proud hit on the play button. “*Pian name chunzhen de [ʒən³⁵], hui zao tianqian de!*” The final output was a take that beautifully synchronized with the character’s mouth movement, and most importantly, had the correctly pronounced *ren* and no *shi* in the sentence precisely as the director wanted. Satisfied with the result, the engineer proceeded to perform another “surgery” to officially merge this newly created “perfect” recording with the master soundtrack.

As we observed the engineer’s swift actions, the director began to reflect on the rerecording process we had spent in the previous session and the “little surgery” just now to “perfect” this four-second line:

“We spent so much time on this to ensure we maintain a good *standard* speech that sounds *pleasant*. That [lən³⁵] was so unclear (*bu qingchu*) that we could not keep it. If we cannot do a good job at overdubbing even one single line, people would think when they hear it, ‘Hmm, something is weird. It doesn’t sound like how people talk *normally* (*zhengchang*).’ Then they wouldn’t want to watch the show. So it’s worth the effort. We must cut out the *impure* Mandarin. It’s just like the saying, ‘A straight tree is pleasant to the eyes (*shu zhi cai haokan*)!’”

She ended the comment with a tree-planting metaphor that I sometimes heard in the metadis-cursive discourse of these dubbing professionals. The director drew an analogy between their effort to “straighten” this line and bonsai cultivation. If a bonsai provides visual aesthetics for its spectators, dubbing likewise has to offer auditory pleasure for the viewers. To achieve this, just like a bonsai artist trims and prunes unwanted leaves and branches from a specimen, they, as dubbing professionals, must “cut out the *impure* Mandarin” to cultivate “a good standard speech that sounds *pleasant*.” In the same vein as the professional verbal hygienists in the United Kingdom (Cameron 1995) and elsewhere (Silverstein 1996; Agha 2007), the director deemed “lin-guistic impureness” as the key contributing factor that would alienate viewers due to the “weird” auditory experience it brings. Thus, as a responsible director, she sweated to remove all differently

enacted versions of *ren* that are “unclear (*bu qingchu*)” and do not resemble “how people normally talk (*zhengchan shuohua*).”

What she did not mention, however, is why she had been concerned with the copula *shi* so much that she was willing to put extra effort into mixing a “perfect” line out of two recordings. If linguistic purity was the goal of all these edits, the fourth retake with its correctly pronounced [ʒən³⁵] essentially met the criterion. There was nothing—both in terms of standardness and linguistic structure—in that very enactment to warrant her annoyance or desire to conduct the “small surgery.” In fact, much later that night, when I walked with her to a nearby metro station, I followed up with her on the *shi* reduction. I brought up what the actor said to me during the smoke break, to which the director reactively responded as follows:

“The word *shi* is a redundant word here. It shouldn’t be kept! He made that claim [i.e., *shi* is optional in this sentence] only because he doesn’t know what Standard Mandarin (*Biaozhun Guoyu*) really is. A good voice actor should speak the standard (*Yige hao de peiyin-yuan jiu yinggai yao biaozhun*)!”

From her reaction, it is evident that the director knew *shi* in this sentence is structurally “optional” as far as standardness is concerned. Nonetheless, to her, this was not a harmless personal preference. She took this grammatical optionality for a grammatical mistake, an indication of the actor’s unskilled use of Standard Mandarin. Getting rid of it is the way to ensure the line stays “standard” in the way she believes in.

Indeed, as the actor implicitly complained during the smoke break, the director’s insistence was irrational. These reactionary behaviors, beyond their affective affordances for a speaker to reaffirm how they feel about (non)standardness in ways that reflect the larger ideological construct, also serve as the means for the speaker to strengthen that felt sociolinguistic reality both ideationally and synchronically (Kroskrity 2000). In other words, the redub process analyzed here is not entirely about whether the actual result conforms to the SMC (however much it may seem to be); rather, it is very process of operating on the logic of the SMC that matters (Cameron 1995). These three dubbing practitioners are not merely surrogates of the ideological apparatus of the SMC and blindly promote and perpetuate its influences. They reaffirm its ideological relevance and, in the process, potentially alter and challenge it (although not in this case) through a lengthy negotiation over how *ren* should be pronounced and whether *shi* should be retained in a simple sentence. Just as the director had her specific vision, the actor and the engineer both had theirs. The three took different courses of action—including insisting on retakes, quietly creating extra soundtracks for later editing, and violating the instruction by retaining the problematized word—that all contributed towards one professional goal: to make “perfect” Mandarin dubbing for the Taiwanese audience.

Conclusion

In this article, I examined the production of Mandarin dubbing in a studio in Taipei to understand how language workers in media execute and negotiate their creative decisions in ways that not only manifest their idiosyncratic language ideologies but also (re)affirm the dominant cultural beliefs about what constitutes “standard.” Situated against the cultural context I dub “SMC,” in which Standard Mandarin is favored over its nonstandard counterparts, I sought a cultural explanation for why dubbing experts in this research were willing to dedicate tremendous time and effort to correct linguistic expressions as small as a particle or syllable. In a profession where every second counts, any excessive investment made to regiment language expressions is costly. From a practical perspective, the stupendous effort I documented in this article for a four-second-long dubbing line is so economically irrational that it demands consideration of the lingua-cultural factors that these workers must navigate (Woolard 1985). From an ideological perspective, the back-and-forth over what could go into the final recording requires critical attention to the work conditions of dubbing professionals rather than an oversimplified statement of it being a continuing process of language standardization (Milroy and Milroy 1998).

Thus, in response to the call in Crispin Thurlow's edited volume (2020) for a nuanced approach to language workers that takes into account both their work discourses and work conditions, I closely examined the complex web of discursive and nondiscursive actions during and immediately after a recording session to illustrate that dubbing professionals do not blindly conform to and voluntarily promote the SMC as past studies on Taiwan suggest (Jan 2017). Certainly, practitioner reflexivity offers tremendous theoretical and applied importance (Duchêne 2009). These dubbing experts showed a strong metapragmatic awareness of standardness and proactively engaged in linguistic intervention during recording, as evidenced in their frequent comments about someone or the targeted features for being "unclear (*bu qingchu*)" and "abnormal (*bu zhengchang*). That being said, too often, such practitioner reflexivity is collected "out of context" in that the researcher (re)constructs the work conditions and ideological operations based on practitioners' retroactive self-reports *outside* the very work context in which these decisions and actions are made. Thus, the ethnographic approach that I adopted in this research underscores its applied implications that would complement the discourse-centered approach in applied linguistics (Sherouse 2015; Ganti 2021) and the text-centered approach in translation studies (Díaz Cintas and Anderman 2009). Rather than treating language workers' discursive accounts as passive reflections of the dominant belief about language standardness (or lack thereof), I ethnographically situated these statements about the dubbing text within the very concurrent moments of production in which these workers debate, negotiate, and compromise on what defines acceptable language expressions in media. Since dubbing recording is never a linear process of audiovisual text production, where voice actors simply transform written dialogs into moldable sound materials for further treatment by sound engineers, mediated language such as dubbing cannot be examined solely as a reflection of dominant language ideology that sustains the cultural majority (Lippi-Green 2011) or perpetuates stereotypes at the expense of linguistic minorities (Queen 2004). Instead, I contend it is imperative for scholars to approach language work in media as a multiparty, multimodal, and nonlinear production process that mediates the complex interactions between the local production networks of language experts and the broader political-economic and ideological apparatus of language work.

Notes on contributors

Spencer Chen is an Assistant Professor of anthropology and linguistics at Hamilton College, USA. His research investigates language ideologies, cultural production, and what happens when they intersect. With a regional focus on East Asia, Chen studies how people deploy semiotic resources such as language and visual cues to navigate disparate forms of sociocultural and political complexities in various settings, including media production, everyday interactions, and social media.

Acknowledgements

I am deeply grateful to Paul Kroskrity, Elinor Ochs, Crispin Thurlow, Hongyin Tao, Erin Debenport, Edwin Everhart, Sonya Rao, Rosalie Edmond, and three anonymous reviewers for their invaluable feedback. I would also like to extend my thanks to the dubbing professionals in Taipei for their generosity and fellowship during my fieldwork since 2016. Any errors, omissions, and deficiencies in this work are entirely my own responsibility.

Funding

The research on which this article is based was made possible with generous support from the U.S. National Science Foundation (DDIG-1824089), Institute of Ethnology of Academia Sinica, the Taiwanese Overseas Pioneers Grants (TOP Grants), and the Asia Pacific Center at the University of California, Los Angeles.

Notes

¹In order to capture the unique complexities of studio interactions, I developed the two-column transcription system that would not only visualize but also account for the built environment of studio recording. The left column transcribes all the interactional turns that originated within the control booth, whereas the right column focuses on the turns that originated in the recording booth. The line between the two columns signifies the physical barrier of the soundproof wall as well as the technology-mediated nature of the documented interactions. Additionally, the text is presented in either black or gray ink to capture the participants' uneven access to linguistic and paralinguistic information. Text in black ink indicates all the verbal and nonverbal cues (e.g., speech, noise, nonverbal actions, and pre-recorded soundtracks) by anybody during a recording session that are entirely accessible to and controlled by the participants in the control booth. In contrast, text in gray indicates verbal information that is available only through the headphones to the participants in the recording booth. Thus, as evidenced in lines 53–56 in [Transcript 2b](#), the director or the sound engineer would sometimes disable the intercom to comment on or complain about the actor's performance without the latter's knowledge.

References

- Agha, A. (2007) *Language and Social Relations*. Cambridge: Cambridge University Press.
- Anderson, B. (1991) *Imagined Communities: Reflections on the Origin and Spread of Nationalism*. London: Verso.
- Bauman, R., and Briggs, C. (2003) *Voices of Modernity: Language Ideologies and Politics of Inequality*. Cambridge: Cambridge University Press.
- Bourdieu, P. (1991) *Language and Symbolic Power*. Cambridge: Polity Press.
- Cameron, D. (2000) 'Styling the Worker: Gender and the Commodification of Language in the Globalized Service Economy', *Journal of Sociolinguistics*, 4: 323–47. <https://doi.org/10.1111/1467-9481.00119>
- Cameron, D. (1995) *Verbal Hygiene: The Politics of Language*. London: Routledge.
- Chaume, F. (2014) 'Research Paths in Audiovisual Translation: The Case of Dubbing', in C. Millán and F. Bartrina (eds) *The Routledge Handbook of Translation Studies*, pp. 288–302. London: Routledge.
- Chen, P. (1999) *Modern Chinese: History and Sociolinguistics*. New York, NY: Cambridge University Press.
- Chen, S. (2022) 'From Mandarin Monopoly to Sinitic Polypoly: The Story of Dubbing Peiyin in Postwar Taiwan', *International Journal of Taiwan Studies*, 5: 308–31.
- Chen, S. (2018) 'Visualizing Language Ideologies and Verbalizing Perceived Linguistic Boundaries: The Case of Mandarin Chinese in Contemporary Taiwan', *Global Chinese*, 4: 11–36.
- Chun, A. (2017) *Forget Chineseness: On the Geopolitics of Cultural Identification*. Albany, NY: State University of New York Press.
- Condry, I. (2013) *The Soul of Anime: Collaborative Creativity and Japan's Media Success Story*. Durham, NC: Duke University Press.
- Coupland, N. (2001) 'Dialect Stylization in Radio Talk', *Language in Society*, 30: 345–75. <https://doi.org/10.1017/s0047404501003013>
- Díaz Cintas, J., and Anderman, G. eds (2009) *Audiovisual Translation Language Transfer on Screen*. New York, NY: Palgrave Macmillan.
- Duchêne, A. (2009) 'Marketing, Management and Performance: Multilingualism as Commodity in a Tourism Call Centre', *Language Policy*, 8: 27–50.
- Errington, J. (1991) 'On the Ideology of Indonesian Language Development: The State of a Language of State', *Pragmatics*, 2: 417–27.
- Everhart, E. (2021) 'Crafting (Non)accented Linguistic Identity in Northeastern Japan', *Journal of Anthropological Research*, 77: 35–51.
- Ferrari. (2009) 'Dubbing The Simpsons: Or How Groundskeeper Willie Lost His Kilt in Sardinia', *Journal of Film and Video*, 61: 19–37.
- Friedman, P. (2018) 'The Hegemony of the Local: Taiwanese Multiculturalism and Indigenous Identity Politics', *Boundary 2*, 45: 79–105.
- Ganti, T. (2021) 'English is So Precise. And Hindi can be so Heavy!': Language Ideologies and Audience Imaginaries in a Dubbing Studio in Mumbai', in R. Rossoukh and S. Caton (eds) *Anthropology, Film Industries, Modularity*, pp. 41–62. Durham, NC: Duke University Press.

- Garcia, A. C. (2019) *How Mediation Works: Resolving Conflict Through Talk*. New York, NY: Cambridge University Press.
- Gramsci, A. (1971) *Selections from the Prison Notebooks*. New York, NY: International Publishers.
- Her, O. (2012) 'Language and Group Identity: On Taiwan Mainlanders' Mother Tongues and Taiwan Mandarin', *Language and Linguistics*, 10: 375–419.
- Her, O. (2014) 'Taiwan Mandarin, a Mandarin Variety Formed under the Social and Language Contact between Various Chinese Dialects and Their Speakers', *Language and Linguistics*, 15: 635–62.
- Hsu, H., and Tse, J. (2007) 'Syllable-Final Nasal Mergers in Taiwan Mandarin – Leveled but Puzzling', *Concentric: Studies in Linguistics*, 33: 1–18.
- Hubers, F. et al. (2023) 'Hypercorrection as a By-product of Education', *Applied Linguistics*, 41: 552–74. <https://doi.org/10.1093/applin/amz001>
- Iaia, P. (2015) *The Dubbing Translation of Humorous Audiovisual Texts*. Newcastle upon Tyne: Cambridge Scholars Publishing.
- Inoue, M. (2004) 'What Does Language Remember?: Indexical Inversion and the Naturalized History of Japanese Women', *Journal of Linguistic Anthropology*, 14: 39–56. <https://doi.org/10.1525/jlin.2004.14.1.39>
- Jan, O. (2017) 'The Myth of "Standard National Language"? Phonetic and Sociolinguistic Analyses on Dubbing Performance in Taiwan', Master's thesis, National Taiwan University.
- Jefferson, G. (1987) 'On Exposed and Embedded Correction in Conversation', in G. Button and J. Lee (eds) *Talk and Social Organization*, pp. 86–100. Clevedon: Multilingual Matters.
- Kroskrity, P. ed. (2000) *Regimes of Language: Ideologies, Politics, and Identities*. Santa Fe, NM: School of American Research.
- Kubler, C. (1981) 'The Development of Mandarin in Taiwan: A Case Study of Language Contact', Doctoral Dissertation, Cornell University.
- Lin, S. (2009) 'How Listening Is Silenced: A Monolingual Taiwanese Elder Constructs Identity through Television Viewing', *Language in Society*, 38: 311–37. <https://doi.org/10.1017/S0047404509090514>
- Lin, Y., and Chan, M. (2021) 'Linguistic Constraint, Social Meaning, and Multi-Modal Stylistic Construction: Case Studies from Mandarin Pop Songs', *Language in Society*, 51: 603–26.
- Lippi-Green, R. (2011) *English with an Accent: Language, Ideology, and Discrimination in the United States*. London: Routledge.
- Milroy, J. (2001) 'Language Ideologies and the Consequences of Standardization', *Journal of Sociolinguistics*, 5: 530–55. <https://doi.org/10.1111/1467-9481.00163>
- Milroy, J., and Milroy, L. (1998) *Authority in Language: Investigating Standard English*. New York, NY: Blackwell.
- Pritzker, S. (2012) 'Living Translation in US Chinese Medicine', *Language in Society*, 41: 343–63.
- Queen, R. (2004) 'Du Hast Jar Keene Ahnung': African American English Dubbed into German', *Journal of Sociolinguistics*, 8: 515–37. <https://doi.org/10.1111/j.1467-9841.2004.00272.x>
- Rosa, J. (2019) *Looking Like a Language, Sounding Like a Race: Raciolinguistic Ideologies and the Learning of Latinidad*. New York, NY: Oxford University Press.
- Sandel, T. (2003) 'Linguistic Capital in Taiwan: The KMT's Mandarin Language Policy and its Perceived Impact on Language Practices of Bilingual Mandarin and Tai-gi Speakers', *Language in Society*, 32: 523–51.
- Sherouse, P. (2015) 'Russian Presence in Georgian Film Dubbing: Scales of Inferiority', *Journal of Linguistic Anthropology*, 25: 215–29. <https://doi.org/10.1111/jola.12090>
- Silverstein, M. (1996) 'Monoglot a Standard in America: Standardization and Metaphors of Linguistic Hegemony', in D. Brenneis and R. Macaulay (eds) *The Matrix of Language: Contemporary Linguistic Anthropology*, pp. 284–306. Boulder, CO: Westview Press.
- Su, H. (2019) 'The Metapragmatics of Taiwanese (Im)Politeness: Conceptualization and Evaluation of Limao', *Journal of Pragmatics*, 148: 26–43. <https://doi.org/10.1016/j.pragma.2019.05.018>
- Thurlow, C. ed. (2020) *The Business of Words: Wordsmiths, Linguists, and Other Language Workers*. New York, NY: Routledge.
- Wei, J. (2008) *Language Choice and Identity Politics in Taiwan*. Lanham: Lexington Books.

- Woolard, K. (1985) ‘Language Variation and Cultural Hegemony: Towards an Integration of Sociolinguistic and Social Theory’, *American Ethnologist*, 12: 738–48.
- Woolard, K. (1986) ‘The Politics of Language Status Planning: ‘Normalization’ in Catalonia’, in N. Schweda-Nicholson (ed.) *Language in the International Perspective*, pp. 91–102. Norwood, NJ: Ablex.
- Woydack, J. (2019) *Linguistic Ethnography of a Multilingual Call Center: London Calling*. Basingstoke: Palgrave Macmillan.
- Zhang, Q. (2017) *Language and Social Change in China: Undoing Commonness through Cosmopolitan Mandarin*. New York, NY: Routledge.

Appendix: Transcription notations

Symbol	Meaning
text	original speech in Mandarin
gray text	speech played through headphones
Italics	English translation
((text))	nonverbal action
{text}	recordings, prerecorded soundtracks
≈ text ≈	performance (voice-acting; noticeable voice alternation)
[ˈtekst]	phonetic transcription
[	overlapping actions
^	stress
-	cutoff
.	falling intonation
!	rising intonation
?	questioning intonation
::	lengthening
=	latching
°	low volume
CAPS	emphasis