International Journal of Communication 19(2025), 698–715 1932–8036/20250005

How, When, and Why to Use AI: Strategic Uses of Professional Perceptions and Industry Lore in the Dubbing Industry

LAURENA BERNABO
University of Georgia, USA

Building on interviews with more than a dozen dubbing professionals, this article centers on dubbing professionals' perspectives on the changing role of artificial intelligence (AI) in the dubbing industry. I trace developments in dubbing from human-centered practices to the current landscape where neural machine translations and AI are used with increasing frequency but not without limits. Using an industry lore approach, I demonstrate how professionals' current predispositions and aversions to new digital technologies shape the contexts in which those technologies are deployed. Although AI raises concerns about automation and job security across media industries, the current dubbing industry lore effectively limits the use of AI technologies, emphasizing dubbing as a culturally rich process requiring a human touch and thus protecting dubbing professionals' jobs.

Keywords: dubbing, artificial intelligence, industry lore, intermediaries, contextual integrity framework

Because of digital technologies and global streaming platforms, such as Netflix, television content flows have increased significantly in the 21st century (Sánchez-Mompeán, 2021). The continued expansion of Netflix's libraries around the globe and the growing availability of free ad-supported television (FAST) channels have produced an increased demand for audiovisual translation (AVT) as content flows throughout the global marketplace. Language plays a key role in global media flows, and English-language U.S. content remains dominant in the global media landscape where it has long been "the nearly universal second choice . . . because it is conveniently available, familiar, and an expensive product at an affordable price" (Tunstall, 1994, p. 19). The United States thus produces this "universal second choice," as the United States "is a very diversified market and produces more culturally 'neutral' media products, which therefore also happen to appeal to global media audiences abroad" (Mast, De Ruiter, & Kuppens, 2017, p. 2563). When translating television and film, companies negotiate among numerous options that require varying amounts of time and money, and machine translations (MTs) using algorithms and artificial intelligence (AI) offer appealing avenues for lessening both. However, dialogue cannot simply be translated from one language to the next; rather, it must go through the process of localization whereby foreign content is made culturally intelligible (Chalaby, 2002). This process includes linguistic translation and the adaptation of cultural references and

Laurena Bernabo: LB78516@uga.edu
Date submitted: 2024-01-15

the like. AVT is one of many media industries that have grown to incorporate AI because it can both accelerate and financially cheapen the process.

Translation is "a tool through which the dominant ideologies are disclosed and questioned" (De Marco, 2009, p. 177), but little scholarship has engaged the shifting perceptions and norms around human versus computer-generated translations in dubbed entertainment media (Jin & Yuan, 2023). Researchers exploring the predispositions of MT technology users in other contexts have found that users appreciate the efficiency of MT for inessential tasks but also desire higher quality translations (Vieira, O'Sullivan, Zhang, & O'Hagan, 2023). In addition, MT technologies are deemed less valuable for translating creative works like literature and audiovisual content than more business-oriented and technical texts (Besacier, 2014; Calvo-Ferrer, 2023; Taivalkoski-Shilov, 2019). Shifting perspectives, scholars who focus on consumers' perceptions find that people show a negative bias toward messages they believe to result from MT rather than humans (Asscher & Glikson, 2023; Calvo-Ferrer, 2023). Even when scholars do consider MT and AI in dubbing, they focus on outlining technological affordances or evaluating the quality of automatic dubs (Baños, 2023; Jin & Yuan, 2023) with little insight into the experiences and perspectives of dubbing professionals or the industry lore (Havens, 2014) surrounding their use of AI.

Drawing on interviews with dubbing professionals, this article explores the evolution of dubbing practices and how MT has been alternately incorporated and resisted, bridging translation studies, media industry studies, and AI studies. This interdisciplinary approach uses theories of intermediaries and industry lore, algorithmic aversion, and the Contextual Integrity framework to make sense of dubbing professionals' perceptions of AI and its growth within the dubbing industry. Although some of the academic approaches to studying machine technologies have "fixated on the machine—its design, actions, and outcomes—almost to the exclusion of humans" (Natale & Guzman, 2022, p. 628), I join Natale and Guzman as well as others to highlight human/machine interconnectedness by applying Havens's (2014) industry lore approach to AI in dubbing. In doing so, I join global media scholars like Bielby (2011) and Mast et al. (2017) who focus on intermediaries entrenched in linguistic issues in global television flows. The current article thus discusses the dubbing industry lore about AI and its power to shape the future of AI in dubbing. Indeed, while I offer an in-depth exploration of industry lore in dubbing, this model offers the opportunity to investigate the impact of AI on other media sectors as well, because, as this study suggests, industry discourse may ultimately play a general role in determining what is and is not accomplished with AI.

Literature Review

Dubbing Intermediaries and Industry Lore

The global media market is fueled not by corporations but by intermediaries: individuals who work across boundaries and understand and promote organizational common sense (Havens, 2014). Havens terms this common sense "industry lore," which sets boundaries around what media professionals believe to be possible and profitable. As Havens notes, industry lore is produced from particular industrial conditions and cultural assumptions, so changes to industrial conditions may encourage shifts in industry lore. Industry lore can be identified or exposed through discussions and interviews with intermediaries; although individual intermediaries' opinions vary and are subjective, they coalesce into a reasonably coherent industry lore that

has material consequences for industrial practices (Havens, 2014, p. 40). For example, media buyers and sellers act as intermediaries who make business decisions based on expectations about what content will do well in any given market, which is, in turn, based on their experience, expertise, and cultural interpretations (Bielby, 2011; Havens, 2006). Industry lore explains and outlines the parameters for standard practices in an industry. Expanding on this, I suggest that lore can be strategically employed by industry professionals whose fields might otherwise be upended by technological developments, not just shaping the choices made but justifying the continued employment of thousands of media workers. In the dubbing industry, dubbing professionals are intermediaries who have developed industry lore that details and prescribes using AI in particular ways—notably, as I show below, in ways that keep them in business, using digital technologies as professional tools rather than as replacements for human workers.

Industry lore shapes the choices global media intermediaries make when choosing how to adapt content, most often through dubbing or subtitling. Chiaro (2009) outlines the practical, political, and sociolinguistic (dis)advantages of each in ways that evoke AVT industry lore, pointing out that dubbing countries are typically motivated by "protectionist reasons . . . to inhibit English and to exalt national languages" (p. 143) or the desire to promote and standardize a minority language. By contrast, many small countries opt for subtitling because dubbing is deemed too expensive, but this should not be mistaken as evidence of these countries' disinterest in dubbed content. Rather, as scholars point out, many countries that previously subtitled foreign-language audiovisual content now demand dubs (Grannell & Chaume, 2023; Sánchez-Mompeán, 2021). Dubbing intermediaries include both individuals who create dubs and executives who promote and maintain industry lore around how dubbed content can and should be created.

Translators act as intermediaries between the countries that buy and sell media texts. A translator adapts a script from the source to the target country linguistically and culturally so that dialogue is both translated into the new language and localized to the target country references (Bernabo, 2022). Dubbing is a complex process because it requires not just translation and localization, as subtitling does, but also adequate synchronization between the source actor's lips and the dubbing actor's dialogue. Early dubbing was quite poor in this regard, leading to a fair amount of derision, but lip sync has improved over time to be nearly seamless, in part because of digital technologies. A dubbing company will typically receive the script in its original language and provide it to the person who will create the foreign-language script. Once the script has been linguistically translated and culturally adapted, it is given to the director and editor for recording. Although the translator is singularly responsible for crafting the dubbing actors' language, the director must command performances that capture the spirit of each character. Dubbing is, therefore, a lengthy and expensive form of AVT that requires work by numerous individuals.

Dubbing and Artificial Intelligence

For decades, dubbing companies worldwide relied almost exclusively on human labor. Examples of traditional, established dubbing companies include VSI, Iyuno, The Kitchen, New Art Dub, Plint, and Transperfect. These remain popular, particularly among larger distributors. For example, Netflix has preferred vendors in each territory, with larger companies like VSI and Iyuno dubbing into many languages and others like Transperfect dubbing into a select few major languages. However, as streaming

platforms like Netflix increase the global flow of television series and films, other digital technologies have also developed that facilitate the localization process, such that foreign-language dubs may now use increasingly sophisticated MTs (Asscher, 2022). Machine translation has "dramatically transformed communication and understanding across languages and cultures worldwide" and "globally changed the very nature of translation phenomena" (Asscher, 2022, p. 1). Machine-learning algorithms can work independently of human intervention to create content (Nah et al., 2024), including linguistic translations. Machine translation might include AI, which refers to "a broad assemblage of technologies that replaces human actions and, increasingly, replaces human decision making in ways that, at least superficially, resemble human thinking" (Ebbe & Kramarae, 2023, p. 195). AI features most prominently through neural MT (Forcada, 2017), which "uses neural networks to learn the statistical relationships between words in different languages, which allows it to generate more accurate and natural-sounding translations than traditional machine translation methods" (Calvo-Ferrer, 2023, p. 1115). Technologies like ChatGPT (Calvo-Ferrer, 2023) and CUBBITT (Popel et al., 2020) similarly offer opportunities to make the dubbing process faster and less expensive. Many traditional dubbing companies increasingly use MT technologies, including AI, whereas newer companies, chief among them Papercup and Deepdub, go further and prioritize AI.

There are numerous reasons a dubbing company might embrace or avoid AI. Oomen, Gonçalves, and Mols (2024) offer insight into the variability of human aversion to, and appreciation for, the kinds of algorithms that are foundational to AI capabilities. People who prefer human methods to AI, even when computer-generated content is superior, are characterized as exhibiting algorithm aversion (Oomen et al., 2024). In the dubbing industry, there is a consensus that computer-generated content is not inherently superior, yet many companies use it as a cost- and time-saving tool that undergoes a human review process. This is true even for more purely AI-driven dubbing companies like Deepdub. There are different degrees to which dubbing professionals communicate an affinity for or aversion to the use of AI in the dubbing process, and the varied responses align well with Nissenbaum's (2009) contextual integrity (CI) framework. CI was developed to understand privacy but has since been applied to other contexts to more generally explain norms of appropriateness for the use of technologies in communication practices (Oomen et al., 2024). The four parameters shaping these norms are contexts (structured social settings), actors (information senders and recipients), attributes (information type), and transmission principles (constraints on information flows). Oomen et al. (2024) articulate the applicability of this model to AI data processing, which is similarly shaped by norms of appropriateness. The CI framework further applies to the dubbing industry, which increasingly uses AI, and many of the dubbing professionals I spoke with indirectly evoked this framework and a degree of algorithmic aversion while making sense of AI's disruption of traditional dubbing processes. The industry lore among dubbing professionals proposes and reinforces norms for appropriate uses of AI in dubbing, including source and target language, market, channel of distribution, and size of audience. For example, industry lore suggests that it is inappropriate to rely meaningfully on AI when translating U.S. blockbuster films for major global media markets. This lore is reinforced by larger discursive communities throughout media industries, but AI can feature prominently in English-language dubs of Portuguese dramas for U.S. audiences to watch on Hulu. Such rules about if, when, and how to use AI in the dubbing process are common among dubbing intermediaries and constitute dubbing industry lore.

Methods

Many media flow scholars have highlighted how countries import and export content, while others have focused on the individual decision-makers who gather at television festivals and conventions and serve as gatekeepers (Havens, 2006). One such fair is MIPCOM, which draws thousands of media professionals worldwide to Cannes, France annually in October. MIPCOM is the largest of the international television trade events, and buyers, sellers, and AVT companies come together not only to conduct business but also to share and learn about global trends and developments. This article builds on the information I learned while conducting field research at MIPCOM in 2022 and 2023, where I interviewed and saw demonstrations by dubbing studios from across North America, South America, and Europe. In addition, I pulled from information I gleaned during previous site visits to individual dubbing studios in Mexico (2015), the United States (2019), and Spain (2022), where I was able to interview dubbing professionals and observe them as they worked. These interviews and observations, spanning eight years, capture the development of industry lore around the use of AI in dubbing; this lore provides unique insider insights into the possibilities afforded by AI technologies, the proscriptions against it, and predictions for the future of the dubbing industry.

I begin with a case study to explain traditional dubbing practices, which were developed long before neural MT. I then turn to the present and explore developments in MT at both traditional and AI-focused dubbing companies through the lens of industry lore, focusing not just on the affordances of new technologies but also on the perspectives and predispositions of the dubbing professionals who might use these new MT technologies. Finally, I look to the future and outline the next stages for advancement in the industry, given the lore shared and perpetuated by dubbing professionals. My focus is on how machines and AI are—and might be—used in dubbing.

AVT: From Human to Hybrid

Traditional Dubbing

To illustrate the complexities of traditional human-driven dubbing, I provide insight into the process through the work of Jesús Vallejo, a script translator in Mexico City. Vallejo is a Mexican man who grew up attending bilingual schools before studying English in the United States and England, earning numerous translation certifications. Vallejo has translated scripts for a dozen U.S. television series and hundreds of films. Vallejo relies on his extensive training and vast knowledge to create Spanish-language scripts for dubbed content. Because of the high costs of dubbing and the linguistic similarities throughout much of Latin America, films and programs dubbed into Spanish for Latin America are typically dubbed just once for the many Spanish-speaking Latin American countries (Fuentes-Luque, 2021). This dub must therefore be pan-Latinx, using generic vocabulary and accents to be equally accessible throughout the region. As a professional film and television translator with many years of experience, Vallejo is frequently able to select the appropriate vocabulary without consulting resources. That said, such resources, including the website Real Academia Española, do exist, and he uses them as needed if only to double-check his instincts.

Beyond the linguistic translation, Vallejo's scripts work to culturally adapt the source material for Latin American audiences. Translators have long been conceived as intercultural mediators (Hatim & Mason, 1990) and experts in intercultural communication (Snell-Hornby, 1999) because of their ability to decode the source material and encode it for new audiences (Mayoral, Kelly, & Gallardo, 1988). The translation process, in fact, relies on the translator's ability to understand multiple sociocultural contexts and interpret across them (Ahmed, Sadiq, Atif, Naseer, & Adnan, 2018). Simply put, a cultural adaptation in AVT requires that translators possess bicultural vision (Santamaria, 2001). Vallejo uses his bicultural vision to localize scripts quickly and efficiently for Latin American viewers, and he is the sole arbiter of what references will be understood, keeping some and adapting others. For example, after hearing a line of dialogue that references the Amish community and apple butter, Vallejo assured me that "Latinos do not know who the Amish are" and thus adapted the references to Mennonites and cheese because Mennonites "are everywhere and are known for selling cheese." While Vallejo may be exceptionally good at his job, he represents the more traditional historical approach to dubbing—an approach that is increasingly disrupted by digital technologies and AI.

Developments in Dubbing

Machine translation plays an ever-growing role in the global marketplace, particularly about linguistic translation. When speaking to dubbing professionals at MIPCOM, many intermediaries from traditional AVT companies described linguistic translation, cultural adaptation, and lip synchronization as three distinct stages in a three-step dubbing process. This is not entirely novel, as Chiaro has previously distinguished between script writing and synchronizing practices (Chiaro, 2009). What was revelatory in our conversations was that these professionals identified the first step (linguistic translation) as an opportunity to save time and money using digital technologies. I described to them the process I had previously observed in Mexico: Jesús Vallejo opened a script for a U.S. television episode for the very first time and translated it sentence-by-sentence, alternately watching lines of the original episode in English, writing an initial translation, and playing the clip again while speaking his Spanish-language translation. In this way, he was able to determine lip synchronicity and equivalence in timing while translating from English to Spanish and localizing U.S. references to pan-Latinx equivalents. Vallejo's choices produced a script with strong lip sync and kinetic synchrony, such that the dialogue matched actors' movements and synchrony, aligning the new dialogue with the opening and closing of actors' mouths, particularly about words containing the letters *m*, *b*, and *p*. My description of Vallejo's skills seemingly caught these dubbing professionals off guard, and they described such a translator as rare and exceptional. One AI expert responded, "Some people are amazing. Like machines." Because dubbing is often understood as a three-step process of linguistic translation, cultural adaptation, and lip sync, digital technologies offer opportunities at different stages, assisting translation professionals who are perhaps not as adroit as Vallejo.

The more traditional dubbing companies are keenly aware of the ways digital technologies and AI can and do change the status quo of their industry and are working to adapt to the changing environment. Indirectly evoking the CI framework, a representative at Transperfect described how needs vary by market, and "the full human process" does not always make sense for smaller audiences, so their company does market research on behalf of clients to help them determine if and how to use AI. As a rule, he explains, AI is appropriate in markets where content needs to be "just good enough," whereas in Japan and other larger

markets with higher standards, "you would obviously want to go in with the best product to start because that's just the expectation of their culture." The Kitchen, a Miami-based dubbing company with dozens of dubbing suites around the world, is actively working with technology companies to develop their software in ways that will best assist the specific needs of the dubbing community rather than eliminating these dubbing companies altogether. Executive Vice President Deeny Kaplan acknowledges that AI can facilitate faster turnaround on linguistic translation, though cultural localization will continue to require human labor. According to Kaplan, the limit seems to be the expression of human emotions, particularly passion, which cannot be adequately captured by AI, but her company's language department has been increasingly happy with the results of their collaborations with tech companies.

Although new technologies may be used in numerous ways through the dubbing process, the CI framework articulates appropriate contexts for their use. These technologies are sometimes used to set a framework for MTs, meaning they provide a model for AI to learn how to translate across languages. Neural MT creates a linguistic translation of a script, which a human translator can then localize, saving considerable time for studios that break down the adaptation process into those three distinct stages; notably, the industry lore circulating at MIPCOM emphasizes the need for human intervention for localizing scripts. Sync can also be improved by digital technologies, and Transperfect engineer Matt Austen demonstrated to me how digital technologies can improve synchronization in the editing process. Using a clip from *Star Wars*, Austen showed both an original clip and its French dub to illustrate how long a line of dialogue should last, explaining, "it would elongate and shrink . . . it's kind of a draft lip sync version and then someone could come in there and adapt it." Discussions among dubbing professionals did not suggest aversion to this form of technological intervention. Another benefit of new technologies is the ability to remotely record dubbing actors, facilitating the "rework" aspect of the editing process, which was similarly deemed acceptable in dubbing industry lore. When there is missing or inadequately dubbed dialogue, it can be (re)recorded remotely by an actor who can be directed to record specific lines, and the turnaround is faster because the actor is not required to return to the studio. With this hybridized remote-dubbing system, actors and directors can digitally connect in the cloud, with directors observing and offering feedback, either through an audiovisual application of their platform or through a text messaging feature. This hybrid approach is also generally more convenient for minor roles that require just a few lines of dialogue, again improving efficiency while cutting down on postproduction costs without replacing human voice actors.

As applied to dubbing, the CI framework articulates norms around the use of AI in two ways: the specific way the technology is used, as described above, and the kinds of projects for which it is used. Plint's Maria D'Alessandro breaks projects into what she terms creative products and essential products. Creative products require human agency at every step because cultural translation and dubbing are understood as crafts, the eloquence of which cannot be replicated by digital technologies. Essential products, on the other hand, can use MT to varying degrees as agreed on by the client, but this option applies only to certain territories because of variations in the quality of MTs. "We only use it where it's valuable and where it makes sense," she states, clarifying that the technology is "quite good" when translating scripts into Latin American Spanish—accurate and efficient—but less so in languages like Hungarian because MT requires extensive human intervention to reach an acceptable standard of quality for localized dubs and subtitles. Machine translation is thus more appropriate in more widely spoken languages and in content that is more likely to be consumed passively or in the background, as opposed

to content intended for appointment viewing. D'Alessandro concludes that she believes there is "a general overestimation of how good it is and how much you can save . . . but it's come a long way and there are some fabulous uses of AI," even if its current utility for dubbing more curated content, such as major films and television programs, remains limited in its current form.

While traditional dubbing companies project cautious optimism about their ability to selectively incorporate digital technologies into their long-established human-driven model, two companies are leaders in AI dubbing. These companies rely heavily on AI across all aspects of the dubbing process, far more than any traditional dubbing company, but they do so while both agreeing with industrial lore's mandate for the human review of MTs and sharing the belief that such translations are not acceptable in all contexts. The London-based company Papercup, established in 2017, provides clients with text-to-speech dubs and specializes in FAST channels with low-end voiceover solutions. The Israeli-based Deepdub was established in 2019 and quickly made a name for itself; it was awarded the Localization Trailblazer EnTech award by the Digital Entertainment Group in 2023. Machine translations of scripts are standard for AI dubbing companies and are a viable, cost-effective option for many of their more traditional colleagues. However, though MT provides the initial script in its new language, industry lore suggests that the quality control process requires a human review of the translated dialogue, addressing issues of measurements, names of places, idioms, jokes, and lip sync. As Deepdub CRO Oz Krakowski points out, "The difference between us and . . . the small startups is that we understand that the technology is not good enough yet; you have to have people in the loop and we have—every process for us includes people." The goal of AI, he explains, is not to replace traditional dubbing professionals, but to carry on the capabilities of exceptional translators like Vallejo, who are highly skilled and "about to be extinct." AI is better able to translate, for example, legal and medical jargon than the average bilingual translator, but the "human in the loop" is still essential for strengthening the final product. Moreover, representatives from both DeebDub and Papercup focus on projects that align with the CI framework for appropriate AVT contexts.

As this conversation indicates, the industry lore among dubbing professionals is that there are limits to the quality of AI-generated dubs, where quality is understood as avoiding obstacles to intelligibility and enabling enjoyable consumption while also dependent on the dub's purpose (Taivalkoski-Shilov, 2019). Part of this is attributable to the fact that narrative textual structure is linguistically fabricated, such that mechanical translations produce distortions in the text. This is why humans remain vital to dubbing processes, even when the scripts are primarily produced by machines. Human translators function as authors who dynamically interpret and creatively re-present the source material for new audiences (Wang & Domínguez, 2016). Papercup's Abhirukt Sapru provides a language for this distinction, dividing media content into different tiers. Broadly speaking, the top tier of high-grade premium content is primarily scripted content that includes blockbusters, tentpoles, and many of the television series in active production, whereas mid-grade content—the largest of the tiers—blends scripted and unscripted content that has a smaller budget but strong brand loyalty among audiences, and low-end content has even smaller budgets. This delineation is not new but established in lore, as Havens (2006) found similar differentiations among dubbing professionals about quality and pricing.

As acknowledged by Plint's D'Alessando, MTs of scripts have been shown to save time and money but at the expense of quality. A text's audience expands into new markets but through the consumption

of inferior products (Besacier, 2014). Representatives for both Papercup and Deepdub volunteer that AI dubs are insufficient for top-tier media products, and that their focus is on mid-grade content. A related element of this is the age of the text, such that Papercup focuses on quick and economical dubs of clients' back catalogs, making classic sitcoms and the like available for global audiences in ways they never were before. AI is also appropriate for low-end content because it can be completely automated (whereas Papercup and Deepdub offer a human/AI hybrid), which further reduces costs for individual content creators who seek the cheapest option available for dubbing services. According to the dubbing industry lore, the top tier of media will never fully embrace AI for two reasons. For one, AI's primary appeal remains its lower cost and quicker turnaround compared with traditional human-based dubbing, but top-tier media comes with a larger budget and fewer time constraints, rendering the benefits achieved through AI less valuable. Second, human voices and performances remain, and will continue to be, the gold standard. "We've done theatrical, and I can tell you, when we compare something we did for one of the studios compared with the original dubbing?" Krakowsi recalls, mentioning a large tent-pole they dubbed experimentally. "The quality of the [original, traditional] dubbing was insane. Phenomenal. I mean . . . like, how did they match the voices? They did the lip sync? . . . And it's only people. How do they do it?" Krakowski concludes that the best people in the industry can create a better dub than AI, but he questions whether AI will ever reach that level while making such high-quality work accessible to everyone. He predicts that AI can and will achieve that level, but that major studios will nevertheless continue to rely on the added value provided by more human involvement in the process. AI dubs are thus more appropriate when dealing with the vast amount of preexisting content that could find new audiences in an increasingly globalized media marketplace.

One final aspect of dubbing where AI shows compelling potential is about voices, and some dubbing professionals at MIPCOM, including representatives from The Kitchen and Transperfect, spoke about their work to help voice actors gain compensation for the work product that AI uses to create dubs. When casting for AI-generated dubs, a company can either match the original voice or use a voice bank to find an algorithmically close fit. To capture a voice for a voice bank, an actor signs away the rights to their voice to a company, which in essence copyrights the voice in perpetuity; one hour of recording an actor's voice is sufficient to capture it so that it may be manipulated and applied to future dubs. Krakowski explains that all the voices they use are synthetic and that they can voice-match the original, though it is rare that they do so. Instead, they create voices that have nuance in ways that align with the character's identity. Thus, while digital technologies now have an established place in the linguistic translation of texts, their potential in voice synthesis offers more novel—albeit growing—opportunities for blending human and computerized efforts to create strong dubs.

The Future of Dubbing

Given the lore pervading MIPCOM, there are a few predictions that scholars can safely make in the dubbing industry. Machine translation and AI are increasingly present in dubbing processes around the world, and their impact will undoubtedly continue to grow in the future. As exceptional translators like Vallejo, who can simultaneously translate language, localize content, and synchronize dialogue with images, grow harder to find, neural MTs will become more appealing. AI-generated dubs will likely remain the cheaper, faster, and easier option, increasingly preferred by distributors with smaller budgets. More

companies will emerge with varying degrees of success to offer low-cost and lesser-quality automated dubbing services, which will be available in more and more languages. Artificially intelligent programs will continue to evolve, incorporating tricks for localizing content, which will cut down on the need for human intervention in what has so far been a tedious editing process (Sakamoto, 2019). Dubbing professionals predict that changes will happen quickly. In pondering the rate of improvement for AI, Krakowski suggests that "every quarter, it's getting a leap forward. It's not like, let's wait a couple of years. A couple years from now, it's gonna be live dubbing, automatic—but this is the speed we're going at."

Beyond linguistic translation and cultural adaptation, what is perhaps most interesting about these ongoing technological developments is their implications with regard to voice. A representative for Transperfect admits, "we know there's people eager to use voice cloning, voice printing, voice morphing, and also speech synthesis." Voice cloning involves using AI to create a synthetic voice based on recordings of human speech, training a machine model to closely replicate that person's voice for any required dialogue. The most famous example of this may be James Earl Jones, the actor who provided the iconic voice of Darth Vader throughout the *Star Wars* (Lucas, 1977) saga. At 91, Jones signed the rights to his voice work archive over to the Ukranian AI startup Respeecher so that the company could recreate his voice for future projects (Breznican, 2022), as it did in the 2022 Disney+ miniseries *Obi-Wan Kenobi* (Harold, 2022). Voice morphing, or voice guide dubbing, is similar in that it can produce an artificially constructed voice, but it does so more directly by converting one voice into another (Ahmed et al., 2018). It would allow, for example, a voice actor to speak a line of dialogue so that the recording sounds like James Earl Jones or any other synthesized voice.

Futuristic practices like voice cloning and voice morphing pose an interesting conundrum for dubbing professionals who strive to save clients' money but also recognize how embedded dubbing norms are among many international audiences. Voice cloning and morphing will offer dubbing directors the ability to re-use the original actor's voice in the dub. For example, a Spanish audience in Madrid could watch a future Tom Cruise film in which they hear a synthesized version of Tom Cruise's voice speaking Spanish. However, this seems to be an unlikely outcome. Beyond the perspective shared by traditional and AI dubbing companies that top-tier content will continue to rely on human dubbing actors who give emotionally captivating performances in a sound booth, dubbing industry lore states that audiences would take issue with voice cloning practices. A representative for Netflix in Madrid argued that Spanish audiences would not like to hear Tom Cruise's voice at all, stating, "We have our own Tom Cruise." Her comment points to the strength of dubbing culture, not just within the industry itself, but among viewers, particularly when it comes to major film stars. In any given country, Tom Cruise and other A-list celebrities are frequently voiced by a single actor time and time again, and audiences come to recognize those voices and associate them with the U.S. actor. Thus, even when the technology is sufficiently advanced to have Tom Cruise seamlessly deliver lines in Spanish, French, and other languages, audiences will reject it because it will cause dissonance with their previous experience. That said, it is possible that voice cloning could be used for new, emerging talents so that the original actors' voice is the only one that is ever known around the globe.

Another developing practice is vubbing, which is sufficiently new to not yet be addressed by industry lore. Vubbing was given a passing mention in French audience data company Glance's 2023 MIPCOM

presentation, "Cracking Audience Trends" (Vaulpré & Rossmanith, 2023) In a section dedicated to AI and emerging technological developments, Glance's Senior Vice President Frédéric Vaulpré made a fleeting reference to vubbing, alongside dubbing, to show that AI can be used to support postproduction processes in the global marketplace. Vubbing essentially takes dubbing one step further by digitally altering the actors' mouth movements to match the dubbed dialogue. Although lip sync has improved noticeably over the past 50 years, vubbing makes dubbing even more seamless, thus diminishing or even eliminating the "uncanny" feeling experienced by audiences of dubbed media (Bosseaux, 2019). When asked about vubbing, Transperfect's Danielle Merrihew exclaimed, "That stuff is so crazy!" Merrihew suggests that it would be easier to take more care in the audio adaptation process to better fit on-screen lip movements and that if vubbing ever takes off as a practice, it will be most popular with countries like the United States that are only recently consuming more dubbed content. Her reasoning is that in territories like France and Italy, where audiences have grown up with dubbed content, viewers experience less dissonance when sync is imperfect. Again, industry lore's reliance on intermediaries' expertise about successes and failures in the global market serves to articulate the potential and limits of AI use in dubbing, and preliminary lore about vubbing indicates that it will not be used by traditional dubbing companies.

Industry lore is less developed when it comes to dubbing for content entering the United States. Though U.S. audiences are not entirely new to consuming translated content, it is still novel for large swathes of the population, and practices here too are in flux. Because so much of the world's media is already in English, there is less demand for dubbing into English, but Netflix and Hulu have made foreign-language content available to U.S. audiences with dubs and subtitles, allowing viewers to select their mode of consumption based on personal preferences. Television series like Spain's *Money Heist* (Pina, 2017–2021) and films like South Korea's *Parasite* (Joon-ho, 2019) have found audiences willing to join the world of international media consumers. Although many countries have been categorized as "dubbing countries" or "subtitling countries," the United States has no such identity, and personal preferences seem to be strong but individualized (Bernabo, 2021; Spiteri Miggiani, 2021). Netflix conducted a study that found that U.S. viewers are more likely to complete a series if they watch it in its dubbed form rather than subtitles (Spiteri Miggiani, 2021), but dubbed content continues to earn criticism for being awkward and unpleasant (Goldsmith, 2019; Sánchez-Mompeán, 2021). As a leading distributor of foreign-language content in the United States, the dubbing norms developed by Netflix will likely shape English-dubbing norms and industry lore. Based on current responses to translated content, particularly the backlash against *Squid Game*'s (Dong-hyuk, 2021–present) machine translated subtitles, we should expect increased human intervention and higher quality translations in the future.

The kinds of foreign content U.S. audiences consume will likely remain top-tier, meaning that distributors will use the traditional human-dependent models of dubbing. Although Deepdub may be capable of creating a more synthetic dub for the U.S. market, Papercup will not even entertain the option as it does not currently offer services for dubbing into English. To do so, suggests Abhirukt Sapru, would be to further saturate an already supersaturated market; because so much of the world's media content is filmed in English, Papercup is instead focused on expanding the content libraries for non-English speakers. For Papercup, it is an issue of access: "if the majority of history is only available to those who speak one language, it's not good for the world generally . . . . Everything should be available for everybody."

Conclusion

This article began with the observation that digital technologies have helped to facilitate an increasingly large and complex global media marketplace, and that digital technologies have, in turn, developed to facilitate dubbing processes. This leads to questions about the precise role of AI in dubbing content for new international markets, the norms developing around the appropriate use of AI technologies, and the role humans might continue to play in the future. These factors inform industry lore, which in turn shapes the material choices dubbing professionals make when using AI in specific ways.

Interviews with more than a dozen dubbing professionals reveal both calm and tension as they navigate a changing industrial environment. Clients interested in cost-cutting technological advancements push to save money, sometimes at the expense of a quality product, and the people interviewed for this article work to steer those clients toward the best solution for their projects. Executives are busy with an endless stream of content to adapt for a fast-growing number of international distribution channels, the needs of which vary by country and platform. Script translators, localization experts, and actors work to demonstrate the superior quality of their work over that of machines, even as they seek and accept compensation for their work to be used to train the AI that is their professional competition.

These interviews reveal industry lore among dubbing intermediaries who share algorithmic aversion, particularly because they do not trust AI dubs to stand on their own without human review for localization and synchronization. However, these dubbing professionals differ in their aversion to other ways AI has functionally intervened in dubbing processes. Furthermore, dubbing industry lore is built on a shared framework for CI when it comes to appropriate uses of human-driven versus AI-driven dubbing; central to this lore is a tiering of content such that AI is deemed more or less appropriate depending on the content being adapted and the new international audience to whom it is being distributed, with higher-tier content requiring more human participation in the dubbing process and lower-tier content deemed acceptable for more reliance on AI.

These findings have a few implications. The first is that language is not just a tool for informational exchange or a potential barrier to understanding, but also an expression of culture (Kilborn, 1993), and that translators, be they human or artificial, thus warrant careful consideration as languages transform and flow across borders. As a result, scholars highlight AI as better suited to certain contexts like business and medicine over the more culturally signified terrain of literature and audiovisual media (Besacier, 2014; Calvo-Ferrer, 2023; Taivalkoski-Shilov, 2019). As Taivalkoski-Shilov (2019) points out, "the dream of creating an immaculate machine translation . . . system that would master literary translation is very old" (p. 689). Though she writes about literary translation specifically, the reality is that dubbing translators have also worked diligently over the course of many decades to improve their craft, and technology companies are increasingly involved in affecting the global media dubbing industry. Translators have long benefitted from digital tools, including online dictionaries, Internet search engines, and time-saving translation memory software (Taivalkoski-Shilov, 2019). Dubbing companies—old and new, large and small—must now navigate shifts in technological affordances, client demands, and global media flows. Technology is both a tool to be used by dubbing professionals and a tool that might be employed in their stead, endangering their livelihoods. This article has begun to shed light on the inherent and interrelated

implications of technological developments, as ethical concerns abound with respect to dubbed products, dubbing as a process, and the dubbing industry (Taivalkoski-Shilov, 2019). However, industry lore may effectively protect much of the dubbing industry by perpetuating a kind of "common sense" about AI usage in ways that protect human intervention and preserve jobs.

A second implication is the potential for AI's cheap and fast translations to assist in democratizing media. While theories of cultural imperialism have long interpreted the global proliferation of U.S. and other English-language media as a threat to local cultures, global media theorists now interpret these flows and counterflows through a more egalitarian lens, and this perspective has been championed by the AI dubbing company, Papercup. As Papercup's Abrhirukt Sapru explained, "right now, we just think the most important thing is that 95% of the world's content is in English. Something like ten, fifteen percent of the world's population speaks it. That's just unacceptable." Sapru's goal is not to Americanize cultures around the globe, but to allow underserved audiences greater access to content that can provide entertainment, effectively overriding the gatekeeping that comes with linguistic barriers. AI dubbing industry lore thus promotes the democratization of access to media content globally, capitalizing on its beneficial affordances but in ways that do not take away human jobs and in fact creates human jobs through AI-dependent dubbing companies.

Finally, neural MTs undoubtedly offer a way to produce dubs more quickly and less expensively; that said, there is little cause for concern that AI technologies will decimate the dubbing industry as we know it. AI has raised concerns throughout various industries that employees will be replaced by cheaper computerized alternatives (Sakamoto, 2019), and postproduction processes like dubbing have been identified as among the media tasks most likely to be impacted by AI (Cho, 2024; Hoover, 2023). However, dubbing companies that specialize in AI are not competing with traditional dubbing companies in the ways we might expect. Deepdub and Papercup are not signing clients who would otherwise hire a traditional dubbing team of human translators, directors, and voice actors. Instead, AI dubs are a financially viable option for companies that would not be able to afford the higher costs of a traditional dubbing company. As Sapru explains, many companies "can never afford to use a human actor or that content would never be dubbed in the first place . . . so the majority of bids we're in are not about us versus traditional media." He continues, "It's about . . . 'we've never considered traditional dubbing because it's so expensive.'" Krakowski expands, "Most of our customers would not dub what they had, what they do with us, if it wasn't *for* us. So, I'm not taking anyone's job. In fact, I'm giving jobs," including voice actors, engineers, and others. The value of companies like these lies in their ability to make content available to underserved audiences very quickly and at affordable prices without sacrificing the higher level of quality, as compared with completely automated dubs that human intervention produces in terms of cultural adaptation, synchronization, and expressive delivery of dialogue.

Although studies of industry lore and professionals' perceptions are inevitably snapshots of moving targets that will evolve alongside technological developments, they nevertheless offer insights into key moments in media history. After all, new technologies are often met with skepticism before becoming common use (Baños, 2023; Jin & Yuan, 2023), and this study tracks varying degrees of comfort with recent and emerging technologies that will continue to enter into dubbing industry practice. My application of the industry lore approach to understanding dubbing professionals' perceptions of AI, and the discourses they develop that prescribe certain uses of AI while proscribing others, demonstrates the

potential of industry lore to shape these industrial practices; each media sector has agency in using AI, and industrial discourses and the lore they produce may concretely affect, to a certain extent, the choices media professionals make across all stages of development, preproduction, production, and postproduction. Havens's (2014) discussion of industry lore posited that lore shapes practices by creating a commonsense basis on which decisions are made when buying and selling media in a global marketplace. This remains true, but my findings demonstrate how lore can shape an industry's evolution in the face of technological advancements. Havens showed how industry lore about audiences' cultural preferences had consequences for the kinds of content made available in different countries, but I find that dubbing industry lore about MT and AI has material consequences for how those resources are used—as tools rather than as replacements. Dubbing industry lore thus plays a part in protecting the dubbing industry from becoming obsolete by highlighting the limits of AI and emphasizing the greater quality that comes with human decision making, which is hard to replicate by machine. Although my focus is on dubbing, it is easy to see how industry lore about AI among media professionals in other areas, as well as professionals in fields outside of media, can function as a line of defense against the potentially devastating influence of AI in their field and work environment.

References

Ahmed, I., Sadiq, A., Atif, M., Naseer, M., & Adnan, M. (2018, February). Voice morphing: An illusion or reality. In *2018 International Conference on Advancements in Computational Sciences* (pp. 1–6). Lahore, Pakistan: IEEE. doi:10.1109/ICACS.2018.8333282

Asscher, O. (2022). The explanatory power of descriptive translation studies in the machine translation era. *Perspectives: Studies in Translation Theory and Practice, 32*(2), 1–17. doi:10.1080/0907676X.2022.2136005

Asscher, O., & Glickson, E. (2023). Human evaluations of machine translation in an ethically charged situation. *New Media & Society, 25*(5), 1087–1107. doi:10.1177/14614448211018833

Baños, R. (2023). Key challenges in using automatic dubbing to translate educational YouTube videos. *Linguistica Antverpiensia, New Series: Themes in Translation Studies, 22*(1), 61–79. doi:10.52034/lans-tts.v22i.763

Bernabo, L. (2021). Whitewashing diverse voices: (De)Constructing race and ethnicity in Spanish-language television dubbing. *Media, Culture & Society, 43*(7), 1297–1310. doi:10.1177/0163443721999932

Bernabo, L. (2022). (De/Re) constructing LGBT characters in Latin America: The implications of Mexican dubbing for translating marginalized identities. *Communication, Culture & Critique, 15*(1), 36–51. doi:10.1093/ccc/tcab045

Besacier, L. (2014). Traduction automatisée d'une oeuvre littéraire: Une étude pilote [Automatic translation of a literary work: A pilot study]. In *21Traitement Automatique des Langues Naturelles* [Natural language processing] (pp. 389–394). Marseille, France: ACL Anthology. Retrieved from https://aclanthology.org/F14-2001.pdf

Bielby, D. D. (2011). Staking claims: Conveying transnational cultural value in a creative industry. *American Behavioral Scientist, 55*(5), 525—540. doi:10.1177/0002764211398077

Bosseaux, C. (2019). Voice in French dubbing: The case of Julianne Moore. *Perspectives, 27*(2), 218–234. doi:10.1080/0907676X.2018.1452275

Breznican, A. (2022, September 23). Darth Vader's voice emanated from war-torn Ukraine. *Vanity Fair*. Retrieved from https://www.vanityfair.com/hollywood/2022/09/darth-vaders-voice-emanated-from-war-torn-ukraine

Calvo-Ferrer, J. R. (2023). Can you tell the difference? A study of human vs. machine-translated subtitles. *Perspectives: Studies in Translation Theory and Practice, 32*(6), 1115–1132. doi:10.1080/0907676X.2023.2268149

Chalaby, J. K. (2002). Transnational television in Europe: The role of pan-European channels. *European Journal of Communication, 17*(2), 183–203. doi:10.1177/0267323102017002692

Chiaro, D. (2009). Issues in audiovisual translation. In J. Munday (Ed.), *The Routledge companion to translation studies* (pp. 141–165). London, UK: Routledge.

Cho, W. (2024, January 30). The Hollywood jobs most at risk from AI. *The Hollywood Reporter.* Retrieved from https://www.hollywoodreporter.com/business/business-news/ai-hollywood-workers-job-cuts-1235811009/

De Marco, M. (2009). Gender portrayal in dubbed and subtitled comedies. In J. Díaz Cintas (Ed.), *New trends in audiovisual translation* (pp. 176–194). Clevedon, UK: Channel View.

Dong-hyuk, H. (Producer). (2021–present). *Squid game* [Television series]. Seoul, South Korea: Siren Pictures.

Ebbe, M., & Kramarae, C. (2023). Smart talk: Feminist communication questions for artificial intelligence. In S. J. Blithe & J. C. Bauer (Eds.), *Badass feminist politics: Exploring radical edges of feminist theory, communication, and activism* (pp. 194–214). New Brunswick, NJ: Rutgers University Press.

Forcada, M. (2017). Making sense of neural machine translation. *Translation Spaces, 6*(2), 291–309. doi:10.1075/ts.6.2.06for

Fuentes-Luque, A. (2021). When Puerto Rico talked to the world: Pioneering dubbing in the Caribbean. *Historical Journal of Film, Radio and Television, 41*(1), 136–151. doi:10.1080/01439685.2020.1766277

Granell, X., & Chaume, F. (2023). Audiovisual translation, translators, and technology: From automation pipe-dream to human–machine convergence. *Linguistica Antverpiensia, New Series: Themes in Translation Studies, 22*(1), 20–40. doi:10.52034/lans-tts.v22i.776

Goldsmith, J. (2019, July 19). Netflix wants to make its dubbed foreign shows less dubby. *The New York Times*. Retrieved from https://www.nytimes.com/2019/07/19/arts/television/netflix-money-heist.html

Harold, J. (Producer). (2022). *Obi-Wan Kenobi* [Television series]. Los Angeles, CA: Disney+.

Hatim, B., & Mason, I. (1990). *Discourse and the translator.* Berkeley, CA: Longman.

Havens, T. J. (2006). Global television marketplace. London, UK: British Film Institute.

Havens, T. J. (2014). Towards a structuration theory of media intermediaries. In D. Johnson, D. Kompare, & A. Santo (Eds.), *Making media work* (pp. 39–62). New York: New York University Press.

Hoover, A. (2023, October 3). Voice actors are bracing to compete with talking AI. *Wired.* Retrieved from https://www.wired.com/story/ai-voice-actors-jobs-threat/

Jin, H., & Yuan, Z. (2023). The application of machine translation in automatic dubbing in China: A case study of the feature film Mulan. *Linguistica Antverpiensia, New Series: Themes in Translation Studies, 22*(1), 80–94. doi:10.52034/lans-tts.v22i.771

Joon-ho, B. (Director). (2019). *Parasite* [Motion picture]. Seoul, South Korea: CJ Entertainment.

Kilborn, R. (1993). "Speak my language": Current attitudes to television subtitling and dubbing. *Media, Culture, & Society, 15*(4), 641–660. doi:10.1177/016344393015004007

Lucas, G. (Director). (1977). *Star wars* [Motion picture]. San Francisco, CA: Lucasfilm.

Mast, J., De Ruiter, K., & Kuppens, A. H. (2017). Linguistic proximity and global flows of television: A study with gatekeepers. *International Journal of Communication, 11*, 2562–2583.

Mayoral, R., Kelly, D., & Gallardo, N. (1988). Concept of constrained translation: Non-linguistic perspectives of translation. *Translators' Journal, 33*(3), 356–367. doi:10.7202/003608ar

Nah, S., Luo, J., Kim, S., Chen, M., Mitson, R., & Joo, J. (2024). Algorithmic bias or algorithmic reconstruction? A comparative analysis between AI news and human news. *International Journal of Communication, 18*, 700–729.

Natale, S., & Guzman, A. L. (2022). Reclaiming the human in machine cultures: Introduction. *Media, Culture, & Society, 44*(4), 627–637. doi:10.1177/01634437221099614

Nissenbaum, H. (2009). *Privacy in context: Technology, policy, and the integrity of social life*. Redwood City, CA: Stanford University Press.

Oomen, T., Gonçalves, J., & Mols, A. (2024). Rage against the artificial intelligence? Understanding contextuality of algorithm aversion and appreciation. *International Journal of Communication, 18,* 609–633.

Pina, A. (Producer). (2017–2021). *Money heist* [Television series]. Madrid, Spain: Atresmedia.

Popel, M., Tomkova, M., Tomek, J., Kaiser, Ł., Uszkoreit, J., Bojar, O., & Žabokrtský, Z. (2020). Transforming machine translation: A deep learning system reaches news translation quality comparable to human professionals. *Nature Communications, 11*(1), 1–15. doi:10.1038/s41467-020-18073-9

Sakamoto, A. (2019). Unintended consequences of translation technologies: From project managers' perspectives. *Perspectives: Studies in Translation Theory and Practice, 27*(1), 58–73. doi:10.1080/0907676X.2018.1473452

Sánchez-Mompeán, S. (2021). Netflix likes it dubbed: Taking on the challenge of dubbing into English. *Language & Communication, 80*(1), 180–190. doi:10.1016/j.langcom.2021.07.001

Santamaria, L. (2001). Función y traducción de los referentes culturales en subtitulación [Function and translation of cultural references in subtitling]. In L. García & P. Rodrigues (Eds.), *Traducción subordinada (II): El subtitulado* [Subordinate translation (II): Subtitling] (pp. 237–248). Vigo, Spain: Universidad de Vigo.

Snell-Hornby, M. (1999). Communicating in the global village: On language, translation and cultural identity. *Current Issues in Language and Society, 6*(2), 103–120. doi:10.1080/13520529909615539

Spiteri Miggiani, G. (2021). Exploring applied strategies for English-language dubbing. *Journal of Audiovisual Translation, 4*(1), 137–156. doi:10.47476/jat.v4i1.2021. 166

Taivalkoski-Shilov, K. (2019). Ethical issues regarding machine(-assisted) translation of literary texts. *Perspectives, 27*(5), 689–703. doi:10.1080/0907676X.2018.1520907

Tunstall, J. (1994). *The media are American.* London, UK: Constable.

Vaulpré, F., & Rossmanith, B. (2023, October). *Cracking audience trends: Who is watching, what, how and why?* Paper presented at MIPCOM, Cannes, France.

Vieira, L. N., O'Sullivan, C., Zhang, X., & O'Hagan, M. (2023). Machine translation in society: Insights from UK users. *Language Resources and Evaluation, 57*(2), 893–914. doi:10.1007/s10579-022-09589-1

Wang, N., & Domínguez, C. (2016). Comparative literature and translation: A cross-cultural and interdisciplinary perspective. In Y. Gambier & L. van Doorslaer (Eds.), *Border crossings: Translation studies and other disciplines* (pp. 287—308). Amsterdam, The Netherlands: John Benjamins.