

6

Languages

Introduction

Aquí Se Habla Español

On October 30, 2020, Bad Bunny and Jhay Cortez release *DÁKITI*, a reggaeton song about the sexual tension of a relationship carried out in secret. Named after a beach in Puerto Rico, *DÁKITI* breaks audience records in record time. In less than a month after its release, it becomes the first song in history to reach the top of the Hot Latin Songs chart and the top ten of Billboard's Hot 100 at the same time. The video clip has more than one billion views on YouTube at the time of writing this chapter in April 2022.

On TikTok, the song is popular for dance challenges and viral lipsyncs. In one of the videos, with more than one million views, user @ralphlarenzo translates *DÁKITI*. The bio of the account states, "I sing Spanish songs in English/Yo canto canciones en español a inglés!" and this is accompanied by the Puerto Rican flag emoji. Many comments about the video revolve around the merits of listening to the song in Spanish versus in English. A user states, "Heck no, Spanish sounds better!!!" Another writes, "I'm glad its in Spanish

lol.” In response someone posts, “He’s Puerto Rican—adding a skull emoji—it’s not even Spanish it’s Latin.”

That popular translation of *DÁKITI* on TikTok and the ensuing dialogue is indicative of the tension between the global success of Benito Antonio Martínez Ocasio, a.k.a. Bad Bunny, and his decision to compose and sing all of his songs in Spanish—with the exception of his single *Yonaguni*, released in 2021, which has a coda in Japanese.¹ In his own words: “I write my songs, it’s my ideas, my production, and I’m not going to have ideas and lyrics come to me in English. I’ve said it from the onset” (Cobo 2020). Whether intentional or not, this artist’s position vis-à-vis English language can be read as an act of postcolonial resistance. Latinx and Latin American artists in the mainstream of the global music industry have been, at least until recently, expected to release versions of their songs in both Spanish and English. Bad Bunny notes that this linguistic practice was perhaps “necessary and they [the artists] opened doors to this Latin boom, but that moment for me is over. I am very proud to get to the level where we speak in Spanish, and not only in Spanish, but in the Spanish we speak in Puerto Rico. Without changing the accent” (Mars 2021; translation from Spanish).

The act of resistance and the vindication of identity, expressed with intensity in his reference to the issue of accent, is also evident in the artist’s social media activity. Whereas many celebrities and influencers express themselves in English on their social media accounts to increase their appeal to a global audience, Bad Bunny—who has more than 38 million subscribers on YouTube, a similar amount of followers on Instagram, and 3.8 million followers on Twitter at the time of writing this chapter—writes all of his posts in Spanish. Yet, when it comes to how social media companies address this content, they sometimes translate it into English. In a video produced by the Twitter company and uploaded to its official YouTube channel, Martínez Ocasio is recorded saying, “Hey people, I’m Bad Bunny and these are the stories of some of my tweets” (translation from

Spanish).² Both Bad Bunny's words as well as the tweets he analyzes in the video are written in Spanish. However, Twitter adds English subtitles, thus positioning itself as an English-speaking platform. What language communities does a platform like Twitter imagine for its users to assume that a translation of a non-English video made by an artist who chooses to express himself in Spanish across his social media accounts is required and that such translation must be into English?

In a February 2020 interview with *Billboard Magazine*, conducted in English, Bad Bunny is asked about the name of his album *YHLQMDLG*. This album would end up becoming the first full-length Spanish-language album in history to top Billboard's all-genre chart. The dialogue that ensues is as follows:

Interviewer: Do you have a title [for the new album]?

Bad Bunny: Yeah, *Yo Hago Lo Que Me Da La Gana*.

Interviewer: Okay.

Bad Bunny: Okay? You don't even know what I said [laughing].

Interviewer: [laughing] I know, I cannot repeat that back, so that's why I had you say it for me, so that I didn't have to.³

In the context of this chapter, the dialogue is noteworthy for the power dynamics associated with the role of language, the tensions between English and Spanish, and the issue of translations. Who translates whom in a world in which English is often presented as the lingua franca of digital culture? When and why do these translations happen or not? How are they disputed? How does the assumption of English as the lingua franca of digital culture contrast with the multiplicity of linguistic experiences of the billions of social media users for whom English is not their native language or who do not understand this language? What does it mean, for instance, that Twitter exercises power in translating some, but not all, posts from some languages and into a few others? Paraphrasing Langdon Winner's (1980) seminal article, do social media translations have politics?

Like and Amen

In March 2013, Jorge Mario Bergoglio, born and raised in Argentina, is named the first non-European pope in history. After his appointment, Pope Francis takes over the Twitter presence of the papacy inaugurated by his predecessor, Benedict XVI. The number of followers rises dramatically since then—from three million in 2013 to more than fifty-two million at the time of this writing—to the point that Jorge Carrión calls Pope Francis “the first influencer appointed directly by God” (Solaris 2020). The millions of followers are distributed across papal Twitter accounts in nine languages, since Pope Francis tweets, almost simultaneously, in Arabic, English, French, German, Italian, Latin, Polish, Portuguese, and Spanish. The papal communication practices on Twitter end up reflecting part of the multilingualism inherent to the lived experiences of people using social media.

In 2016, Pope Francis’s social media exposure expands to Instagram through the @franciscus account shortly after he meets with Kevin Systrom, then CEO of the platform; the text of each post is written in nine languages, in the form of a list. In November 2020, the account comes to the forefront of a public image scandal. Some followers notice that @franciscus has given, along with more than 132,000 other accounts, a “like” to a photo of Brazilian model Natalia Garibotto posing in a swimsuit. The event prompts the opening of an investigation inside the Vatican to trace back what many considered a serious error. However, beyond the content of the liked image there is another aspect especially salient for the current chapter: the signifier of the *like*. Besides the oddity of this particular like, since Pope Francis’s official accounts never react to the content of other accounts, how is one to interpret the meaning of a nontextual signifier across multiple languages? Is it possible to imagine a universal interpretation of this type of signifier? Or, far from it, are we facing an iconography that is deeply polysemic and liable to generating misunderstanding within a millenary institution such as the Catholic Church?

In his message for the LIII World Communications Day, in 2019, Pope Francis ended up contrasting the significance of the like button in digital culture with the amen of the religious dogma: “This is the network we want. A net made not to trap, but to liberate, to guard a communion of free people. The Church itself is a net woven by Eucharistic communion, in which union is not founded on ‘like’ but on truth, on the ‘amen’ with which each one adheres to the body of Christ by welcoming others” (Pope Francis 2019; translation from Spanish).⁴

Just as the interpretation of sacred scriptures has led to a series of semiotic conflicts throughout the history of Catholicism, the appearance of new signifiers in digital culture has ushered social media practices into a terrain fraught with misunderstanding as the norm rather than the exception, making “the uncertainty of meaning” (Furedi 2016, 525) a topic of everyday conversation. When someone likes a post, is the post being liked at the level of the enunciation or at the level of the person who creates it? Is the like a sign of agreement, sympathy, irony, or of other intentions? Moreover, if the content being liked appears in different languages, is the meaning of the like also transformed by virtue of what liking might mean in different languages and within various national and regional contexts? Is the use and interpretation of this social media iconography universal, or is it a matter of global signifiers that are contextualized locally, mediated by each context’s linguistic singularities and their cultural, institutional, and political associations?

Why Linguistic Comparisons Matter

One of the key constitutive elements of both subjectivity and social life, language is the second pathway we propose in this book to programmatically develop a comparative perspective in the study of social media. As with histories, its role is to make more robust the analysis of platforms in their cross-national and regional, cross-media, and cross-platform dimensions. This leads to denaturalizing

an English-language bias that works, in many cases, in tandem with the present-day bias analyzed in the previous chapter. It also implies acknowledging the complex dynamics tied to novel visual signifiers that have become increasingly popular in social media in particular and in digital culture in general.

These two epistemic operations are aimed at countering two dominant scholarly practices that have marked accounts of language in social media within the field of communication studies. The first is the assumption of English as the lingua franca of digital culture (Mullaney 2017; Cheruiyot 2021; Mitchelstein and Boczkowski 2021; Suzina 2021). As Gerard Goggin and Jack McLelland argued more than a decade ago, “there has been little attempt to generate a discussion between scholars working on different language cultures or to develop modes of analysis that do not take Anglophone models as their starting point” (Goggin and McLelland 2009, 6). The second practice is the dominance of the textual dimension of platform use and the relatively much less attention given to visual signifiers whose polysemy resists the computational tools increasingly deployed to make sense of language as an aggregation of words and word frequencies (Highfield and Leaver 2016; Pearce et al. 2020).

The vignettes presented in the introduction to this chapter address these two topics and highlight some key challenges that emerge with the incorporation of a comparative linguistic focus into the study of platforms. The Bad Bunny vignette tackles the English-language bias and, in doing so, points to dynamics of oppression, resistance, and identity in the intertwinement of popular culture, social media, and politics. The Pope Francis vignette expresses, along with multilingualism, the complex place taken by novel visual signifiers—from likes to favs and from emoji to stickers. To counter the limitations posed by the English-language bias and by the dominance of text-only analyses, in this chapter we explore future paths of research in cross-national and regional, cross-media, and cross-platform scholarship that reveal the generative place of multilingual

and multimodal communication in social media. Doing so will help us provide accounts with greater descriptive fit and heuristic power about platforms and their relationship with language.

Language in Cross-National Comparisons

Antecedents

Asaf Nissenbaum and Limor Shifman (2022) study the ways in which satire on social media in reaction to global events works across local cultures. Their goal is to “probe the multifaceted interactions between the global-local and entertaining-disruptive dimensions of contemporary digital satire” (Nissenbaum and Shifman 2022, 937). To do so, they compare a sample of humorous posts across Twitter and Weibo originally written in one of five languages—Arabic, Chinese, English, German, and Spanish—during November 2016 in relation to the election of Donald Trump as the forty-fifth president of the United States. Nissenbaum and Shifman (2022) find few themes and issues shared across satiric posts written in any one of the five languages under study. These were “references to physical appearance, personal relationships, and competitive political dynamics, none of which offered substantial criticisms” (937). Overall, global humor coincided in emphasizing entertainment, whereas locally oriented humor embraced the rather disruptive elements of satire.

Nissenbaum and Shifman also create a typology of local humorous responses to global events, distinguishing between “inbound,” “transitional,” and “outbound” satire across the geographic and cultural regions where the sample of humorous posts originated from. Inbound satire is characterized by analyzing global issues to compare them to local scenarios; this was central for the case of posts written in German. In transitional satire, which was frequent in posts written either in Arabic or Spanish, what matters is the symbolic position of the local audience and its relationship with global

dynamics. In outbound satire, the local audience takes a detached position to comment about global events; this was noted to be mostly present in posts written in Chinese. In conclusion, their inquiry into language as used on social media across countries and regions allowed Nissenbaum and Shifman to illuminate dynamics of the interplay between globalization and political commentary.

Another study that demonstrates the potential of a language focus in comparative cross-national work is by Marco Toledo Bastos, Cornelius Puschmann, and Rodrigo Travitzki (2013), who center on transnational activism organized around specific causes. They ask whether language operates as a barrier or as a bridge when it comes to coordinating efforts through hashtags on social media. To do so, they examine a corpus of 8.4 million tweets and find that out of 455 hashtags, 53 percent were in English and the remaining 47 percent were predominantly in Portuguese or Spanish. The authors show that “linguistic division plays an important role in structuring the network communities” (Bastos, Puschmann, and Travitzki 2013, 166) and observe little overlap between groups that used different languages. However, they note that those hashtags linked to political activism, particularly around the *Indignados* and Occupy protests, had the highest level of network degree, suggesting that “political campaigns based on social media are driven by highly-active, politically engaged users that tweet across different hashtags and are immune to language barriers. Thus, political hashtags transcend linguistic communities, grouping together users and messages produced in a number of different languages” (Bastos, Puschmann, and Travitzki 2013, 168).

Drawing upon different theoretical and methodological frameworks, these two studies show how the incorporation of a linguistic perspective can highlight cross-national dynamics that are key to the lived experience of using social media: cross-cultural convergence and divergence (Nissenbaum and Shifman 2022), and barriers

or bridges across linguistic communities (Bastos, Puschmann, and Travitzki 2013). However, despite their contributions, both studies lack an explicit cross-national agenda on language issues, which we begin to articulate next.

Future Developments

We propose two avenues for future research on language in cross-national social media studies. The first one concerns situations of territorial displacement, and the second one engages with the role of national and regional contexts in the production, circulation, and reception of new signifiers.

Approximately 3.5 percent of the world population reside in a country in which they were not born, exceeding predictions that were made almost two decades ago for the year 2050 (World Migration Report 2020). Furthermore, it is estimated that more than 1 percent of the global population is currently displaced either because of forced migration due to persecution or conflict, or because of statelessness.⁵ For decades now, digital media have played a substantial role in articulating family and work communication across distances (Uy-Tioco 2007; Madianou and Miller 2013; Madianou 2016; Nedelcu and Wyss 2016; Gillespie, Osseiran, and Cheesman 2018; Madianou 2019), and in operating as instruments of surveillance and border control (Latonero and Kift 2018; Leurs and Smets 2018; Sánchez-Querubín and Rogers 2018). To address how this unfolds comparatively in social media, our proposed direction of research partly draws on work in the field of digital migration studies (Brinkerhoff 2009; Alonso and Oiarzabal 2010; Hegde 2016; Alencar 2018; Leurs and Prabhakar 2018). This field is dedicated to understanding the link between processes of territorial displacement and uses of information and communication technologies. Thus, it converges with prior scholarship that emphasized the central place of traditional media during diasporic experiences (Appadurai 1996; Karim

2003; Kraidy 2005), and also with the transnational turn of migration studies that focuses on studying ties across countries or regions (Faist 2004; Nedelcu 2012; Leurs 2019).

Situations of territorial displacement provide a unique window to explore dynamics related to the usual loss of the imagined or assumed correspondence between inhabited place and spoken language, thus challenging the notion that any language has a default status in social media practices, including English. That loss of correspondence is often tied to processes of power asymmetry that tend to be linked to already existing dynamics of economic and political inequalities. Therefore, these situations represent a significant terrain to examine the interconnections between multilingual practices and political dynamics as expressed on social media across national and regional settings. In addition, they highlight the issue of language diversity to which an individual who must move from one country to another is exposed.

Some possible topics to inquire within this future path of research would include how users with different national origins and in situations of geographic displacement manage switching languages when communicating through platforms with at least three different social groups: those remaining in their homeland, those who belong to their diasporic communities in other locations, and those colocated within the local territory. How do these dynamics play over time in cases in which groups of forced migrants are gathered in refugee camps that produce encounters of speakers of different languages? How do they differ in situations of voluntary migration? Furthermore, how do these dynamics vary in both cases in relation to the type of tie at stake, such as familial, romantic, friendship, and work, among others? What are the variations that relate to the kinds of topics discussed among speakers—namely, politics, religion, everyday life, sports, and work, among others? Finally, how are ties and topics influenced by the different power relations enacted in the various situations of territorial displacement across these three possible groups?

The second future avenue for research we propose centers on the role of national and regional contexts in the production, circulation, and reception of new signifiers that have become increasingly popular on social media. As discussed in chapters 1 and 2, a central tenet of common approaches to cross-national and regional studies has to do with assuming that national and regional contexts correspond to specific communication styles and practices (Hofstede 1983, 1991, 1998). Taking this into consideration, how have different contexts of origin shaped the ideation, design, production, and initial evolution of key constellations of new signifiers? For instance, what was the role of Japanese culture for the case of emoji (Gottlieb 2009), or South Korean culture for the development of stickers (Steinberg 2020), or culture in the United States for the design of reactions such as the “like” button (Bucher 2021)? To what extent have national variations in key occupational cultures—such as copyists, illustrators, graphic designers, engineers, and marketing specialists, among others—shaped the construction of these signifiers and the different options developed? In addition, how do these variations relate to broader visual cultural patterns such as the influence of manga in everyday life in Japan (Ito, Okabe, and Matsuda 2005) and the historical role of button technologies in communication interfaces in the United States (Plotnick 2018)?

One element that has made these new signifiers so powerful in digital culture is their cross-national circulation. That is, had they stayed popular only in their country of origin their influence in that country’s everyday communication would have been great, but not so much at the global level—except perhaps within the practices of diasporic communities, as we noted earlier. Yet, these new signifiers have become a mainstay of social media partly because of their uptake in different parts of the world (Gómez-Cruz and Siles 2021). This opens a host of issues related to how they circulate in communication practices across national and regional borders. For example, what could we learn from the use and experience of WhatsApp

groups if we considered the dynamics of sticker sharing among members situated in multiple countries with markedly different communication cultures? In addition, how do language choices in a smartphone's configuration shape which emoji are algorithmically suggested to users? Finally, does an apparently simple signifier such as the thumbs up emoji mean the same in national contexts with divergent cultures of interpretation? In other words, and going back to the second vignette we used at the start of this chapter, did the like by the @franciscus account mean the same to followers and parishioners in different parts of the world?

These last two questions point to the complex dynamics of reception in the case of the new visual signifiers, many times used to disambiguate the lack of tone that is associated with written language (Kavanagh 2016). Is a smile equally interpreted across national and regional contexts? The question exceeds the realm of social media, leading us to inquiring into whether the feeling of happiness and its gestural expression are universal or, in contrast, are decisively shaped by broader cultural configurations of everyday life different in various parts of the world. Cross-national and cross-cultural differences certainly apply when it comes to nonverbal communication (Lim 2002). Similarly, when the hashtag symbol is used to aggregate conversations on Twitter or Instagram, does belonging to these ad hoc communities of discourse mean the same to their participants located in different parts of the world? Finally, how are we to understand the various interpretations that users might have of emoji skin tones across places and their experiences and practices regarding issues of race and ethnicity?

In this section we began to develop the pathway of language by first drawing on studies that show the value of cross-national comparisons focused on linguistic variation, and second by outlining two concrete avenues for future programmatic scholarship: one focused on situations of territorial displacement and the other on the role of national and regional contexts in the production, circulation, and

reception of new signifiers. However, language practices on social media often connect with those related to traditional media. It is to the examination of cross-media issues we turn to next.

Language in Cross-Media Comparisons

Antecedents

Rachelle Vessey's (2015) account of the coverage of the "Pastagate" shows the generative place of language in cross-media dynamics. This affair unfolded in Canada in 2013 when the Quebec Board of the French Language warned a local Italian restaurant to stop using terms such as "pasta" and "calamari" on its menu and to use their French equivalents instead to preserve this language in Quebec. Since the event produced a great deal of media coverage, both locally and internationally, Vessey analyzes how French and English languages were represented by traditional media in four countries—Canada, France, the United Kingdom, and the United States—and by user comments on Twitter. To do so, the author examines the language in which news stories and tweets were written, and their impact shaping the linguistic representations at stake.

Among other results, Vessey finds that both traditional English- and French-speaking media—except for Canadian French-speaking traditional media—and users' tweets largely produced negative representations of Pastagate, depicting "English as a humanised, international language that is necessary for business and French as a marginalised, overly policed language" (Vessey 2015, 268). The author notes that in a context of coexistence of traditional and social media, "the 'barometer' effect of the media reveals the intensification of pressure exerted on minoritised groups to translate linguistic cultures into English and globalised, market-driven contexts" (Vessey 2015, 269).

Another fruitful antecedent is that by Anna S. Smoliarova, Tamara M. Gromova, and Natalia A. Pavlushkina (2018). This study deals

with an emotional aspect of news consumption—the use of Facebook reactions, a functionality that debuted in 2016 to accompany the “like” button, until then the only possible reaction offered per the platform’s design. To do so, the authors focus on news consumption practices undertaken by the Russian immigrant community in Israel. They examine whether there is a correlation between type of reaction and the behaviors of either commenting or sharing a news story on Facebook.

Smoliarova and colleagues find significant correlations between the type of reaction and the kind of engagement with the news article. For example, posts with “angry” and “laughing” reactions tended to be more commented than shared, while those with the highest number of “likes” were not associated with any particular action. The authors warn that “the localization of verbal analogues may question comparative research of Facebook reaction usage across the world. For example, [the] Russian version of reactions that is studied in the paper includes ‘outrageous’ instead of ‘angry,’ ‘super’ instead of ‘love,’ and ‘sympathize with/am sorry’ instead of sad” (Smoliarova, Gromova, and Pavlushkina 2018, 251).

These two studies show the potential of looking at the role of language in cross-media comparative work. In Vessey (2015) we find tensions between English- and French-language representations that signal a degree of convergence across traditional and social media. The account by Smoliarova, Gromova, and Pavlushkina (2018) highlights how the use of new signifiers, which might vary across languages, tie to different types of interaction with traditional media content. Although both studies address to some extent issues of language across media, neither do so as part of a comparative programmatic agenda. We next continue developing it.

Future Developments

We outline two future paths of research in this subsection: first, we address the dynamics of translation across traditional and social

media; second, we tackle the interactions across these two media regarding the incorporation of new signifiers in their respective language practices.

Traditional media with global reach, including leading news outlets and the film industry, have historically developed sophisticated translation processes to make their products available to consumers living in various parts of the world and communicating in different languages (Snell-Hornby 1999; Morley and Robins 2002; Straubhaar 2007; Lobato 2018). These processes have so far followed a relatively slow industrial and one-to-many logic, which has been recently disrupted by the relatively faster and many-to-many counterparts that are paramount on social media (Lacour et al. 2013; Lenihan 2014; Salameh, Mohammad, and Kiritchenko 2015; Desjardins 2016). Such disruption provides a fertile window to further our understanding of the role of language in cross-media dynamics.

Concerning production matters, it would be important to compare the human and technological resources devoted by traditional and social media to their translation efforts and the degree to which they are combined. While news and film companies have tended to rely mostly on human labor, social media companies have primarily resorted to algorithmic translation due to a combination of the volume of content available and the speed at which platforms operate. What happens when content originating in traditional media makes it to social media? How do translations happen in this process? How does technology shape it? Furthermore, what is the role of users-as-translators as opposed to that of professional translators hired by traditional media companies?

The last question leads into issues of distribution and circulation. For instance, although the design of platforms such as YouTube allows for a space to share song lyrics and user comments, the technology behind cable television channels such as MTV offers an environment, at least a priori, more resistant to multidirectional flows of information. Thus, when media products circulate in ways that cross

boundaries between traditional and social media, they problematize these different stances regarding participation from the audience. The song *Yonaguni*, mentioned in the introduction of this chapter, surprised many Bad Bunny fans with its coda in Japanese. Within hours of its release, thousands of comments on the YouTube official video began to offer translations. As has become typical in the genre of “reaction videos,” in which users share their takes on their first encounter with a media product, hundreds of people uploaded clips of themselves reacting to Bad Bunny’s Japanese lyrics and sharing their interpretation of the coda. How do translations circulate in traditional versus social media? What are the various implications of such processes for the content’s reach and reception?

Continuing with the reception of translations across media, how do consumers of traditional media interpret and engage with translated content versus the comparable processes undertaken by users of social media? Whereas the former have limited opportunities to voice concerns if they are unsatisfied with the translated content, the latter have ample avenues to not only express their dissatisfaction but also to propose alternative translations and make them available to other users. What are the implications of these divergent interactive capabilities regarding power dynamics between media production and consumption? Furthermore, the greater translation agency at the disposal of social media users has the potential to foster polysemy regarding the content that circulates, further illustrating the salience of the age-old trope of *traduttore traditore*⁶ in the digital age.

The second path of future research we propose centers on how new signifiers so germane to platform communication have been incorporated into the language of traditional media and how the semiosis characteristic of traditional media has shaped language practices on social media. In which ways is the iconography embodied by elements like emoji, hashtags, stickers, and reactions represented in traditional media? Conversely, how do social media represent, in their digital environments, the visual and aural repertoire linked

to traditional media, as can be seen, for instance, in disparate elements such as cinema billboards, radio jingles, television ads, and newspaper pagination?

In many cases, there seems to be a process of visual mimesis by which traditional media depict what they observe on platforms and further stabilize their meaning in popular culture. Stemming from the world of fiction, the series *Emily in Paris* (2020) tells the journey, from Chicago to Paris, of an American digital marketing expert primarily through the visualization on the television screen of the protagonist's Instagram account. In doing so, it presents an almost exact replica of the visual aspect of the platform, representing likes and reactions to signal the account's success in the character's life. In the domain of nonfiction, gossip and entertainment television programs and newsprint tabloids often draw upon the new signifiers as a source of scoops—for instance, by assuming from an exchange of likes the existence of a romantic relationship between two celebrities.

The comparative question about the representation of new signifiers in traditional media is particularly complex and far from settled. This is because of the strong polysemy associated with the repertoire of these signifiers on social media. The film *Searching* (2018) sets out to narrate a father's desperate search for the whereabouts of his teenage daughter entirely through screens in the digital environment. In its mimetic visual representation of platforms, it features scenes in which the meaning of a single emoji—for example, in Venmo, a social and mobile payment platform where emoji circulate to name banking transactions and socially smooth money exchanges—has the power to twist the course of a police investigation. How do traditional media manage this proliferation of new signifiers and their potential polysemy? How do they combine their own long-standing formats and languages with the recent but powerful emergence of social media iconography? How are the experiences of users on a given platform shaped by the presentation of what they see as emanating from traditional media?

In this section we continued developing the language perspective by building upon two studies that demonstrate the value of looking at language dynamics between traditional and social media, and then proposing two avenues for programmatic work: one focused on the dynamics of translation and the other on the incorporation of new signifiers. However, as we have argued repeatedly, social media are not uniform because there is significant heterogeneity across platforms. In the next section we delve into what this means for accounts of language practices.

Language in Cross-Platform Comparisons

Antecedents

A study by Michele Corazza and colleagues (2019) constitutes a useful instance of cross-platform comparisons focused on linguistic matters. The authors design and test a natural language processing methodology for detecting hate speech on social media able to operate in Italian across multiple platforms. Their motivation is that most of the existing data sets and approaches used to detect hate speech on social media are written in English and focus on one platform at a time, usually Twitter. To this end, Corazza and colleagues draw upon data in Italian from Facebook, Twitter, Instagram, and WhatsApp. They find, for example, that “learning to detect hate speech on the short length interactions that happen on Twitter does not benefit from using data from other platforms” (Corazza et al. 2019, 5). They also find that the emoji detection and transcription system is not as useful for this platform as for the others, probably because of the relatively lower use of these signifiers on Twitter. The authors conclude that “this shows that the language used on social platforms has peculiarities that might not be present in generic corpora, and that it is therefore advisable to use domain-specific resources” (Corazza et al. 2019, 5).

Noting that no previous studies have compared emoji use across platforms, Khyati Mahajan and Samira Shaikh (2019) examine this matter on Twitter and Gab, a platform heavily used by the alt-right community, especially in the United States. They analyze how emoji were used in content produced on these two platforms regarding the Charlottesville massacre, which occurred in the United States in 2017. The authors find that on Gab the sentiment was more negative, and that emoji were used in greater quantity than on Twitter. Furthermore, on Gab certain emoji linked to the political movement of former president Donald Trump (the frog face emoji) prevailed. In contrast, certain emoji that would indicate empathy, such as the broken heart or the peace sign, were more present on Twitter. Finally, Mahajan and Shaikh (2019) note Gab's use of positive emoji in a context of negative connotation, observing that "Gab users tend to use the emoji more in a sarcastic tone, whereas Twitter users tend to use the emoji more to express their disbelief during the event" (2).

Taken together, these two studies highlight the importance of undertaking cross-platform comparisons to understand language matters and show how far from settled content interpretations can be. The work of Corazza and colleagues (2019) showed that relying solely on English in a single platform would miss detection of hate speech in other languages and platforms, with important implications for the regulation of social media content. The study by Mahajan and Shaikh (2019) illustrated the variance in emoji use across two platforms. However, despite their significant contributions, none of these antecedents are part of a larger cross-platform comparative agenda. In the next subsection we continue the process of developing it.

Future Developments

We propose two possible research directions to advance a comparative agenda centered on exploring patterns of variation in written language and in the new signifiers across platforms.

The first direction inquires about the prevalence of various languages on different platforms. A common aspect throughout the design of platforms is that they tend to combine various degrees of personalization of the user experience, including the ability to configure preferred languages, together with different options of algorithmic translations into one or more additional languages. This opens the possibility of variation regarding the language or languages in which content is presented and also how it is received by users. The presence or absence of this variation, in turn, enables the analyst to probe a range of dynamics regarding culture, power, social structure, race and ethnicity, and gender, among others.

Does the country of origin of a platform affect the language considered official by the platform or the range of languages available to users to configure their settings? Furthermore, are there any recognizable patterns of variation by platform in this regard? Besides often having a default language and additional ones available in the user settings, some platforms offer the possibility of automatic translations of content posted in a particular language that an algorithm supposes the user does not understand. In this case, what are the criteria that influence algorithmic decision-making regarding translations of posts originally made in a language into another one, and how does this vary across platforms? Moreover, are these translations made visible (labeled as such), or do they remain opaque and therefore made invisible to users? Are these translations imposed or do they allow for a degree of customization by the user? How does the translation rating system shape the service offered? Again, how do these variations across platforms affect the dynamics of content production, circulation, and reception? What happens when the same company owns a constellation of platforms, such as the case of Meta's ownership of Facebook, Instagram, and WhatsApp? How do linguistic and translation policies vary across them? In the Tower of Babel of social media (Mocanu et al. 2013), questions such as these can help illuminate variations in language production and

distribution across platforms that can in turn help analysts address broader cultural, social, and political matters.

In addition to these issues of language variation in production and distribution dynamics, future research could also inquire into language use and interpretation. Are there major patterns of variation in terms of the languages used across different platforms? Furthermore, do different language communities form within platforms? It is not uncommon, for example, to find YouTube or TikTok comments from users who wish to gather around their own imagined linguistic community, through messages such as “where are the ones who speak [insert language]?” How does this vary, if at all, by social media platform, and why? Beyond the language or languages commonly used by platform, there are issues of interpretation. How do users who speak different languages make sense of a post originally made in a language different from theirs, and how does this vary by platform? Do they share their interpretations publicly on platforms? Moreover, what do they think of the aforementioned algorithmic translations—specifically, those that platforms offer them often in the absence of users requesting them—and also regarding their resulting quality?

The second future path of research proposes inquiries into the cross-platform variation surrounding the repertoire of new signifiers prevalent on platforms. This iconography is expressed linguistically on at least two levels: textual and visual. Textually, a series of questions arises regarding the ways in which the new signifiers are translated into different languages and the implications that such translations have on the production, circulation, and reception of the content. For example, when it comes to reactions, how are their official names translated into different languages on various platforms? As the work of Smoliarova, Gromova, and Pavlushkina (2018) already suggested, these differences lead to important questions: To what extent is it possible to compare the use of Facebook reactions across languages in which these signifiers have different linguistic value? For example, in Spanish there are at least two expressions to communicate

love for someone or something—*te quiero* and *te amo*; in English there is only one that dominates—I love you. The possible range of Facebook reactions in Spanish does not seem to contemplate this difference, opting for a third option that is placed as the equivalent of love in English and that switches from noun (object) to verb (action): I love it [*me encanta*].

Visually, social media have been the seedbed of a new iconography of signifiers, from vernacular signs original to a particular platform, such as the at sign or the hashtag on Twitter, to elements shared across multiple platforms, such as emoji or stickers. In terms of their production, do these repertoires vary by platform? How do representations of the same signifiers change visually depending on the platform and its interaction with the operating system of the device in which users access the content? When it comes to their circulation, in what ways do certain signifiers travel from one platform to another? How are their uses stabilized or contested, in a context of potential polysemy? On Twitter the hashtag often groups a content within a series of discussions, but on Instagram the same symbol is generally used to increase the visibility of a post. Finally, in imagining their reception, how does the use of new signifiers vary across platforms? Why are certain social media, as Mahajan and Shaikh's (2019) work demonstrated, more prone to frequent emoji use than others? In what ways does the sharing of new signifiers across platforms transform over time?

The exploration of language patterns across platforms leads to a number of questions and avenues for research since it enables the analyst to explain and understand aspects of social media use that remain, in many cases, invisible or relatively little discussed. In this chapter we have suggested several possible directions of research in cross-national and regional, cross-media, and cross-platform dimensions aimed at building a programmatic agenda of future work. Next, we close this chapter by bringing these various strands together.

Conclusions

We have argued that just as the pathway of histories aimed to counter the present-day bias that runs through much of the literature on platforms, the pathway of language intends to offer an alternative that complements the twin tendencies to take English and textual communication as the default modes of symbolic praxis in scholarship on social media. The vignettes presented in the introduction sought to illustrate these tendencies. Bad Bunny and his use of social media in Spanish, contrary to the translations imposed by different platforms and the global music industry, underscored the intersection among language, politics, and popular culture. Pope Francis and his platform practices not only showed a strong multilingualism but also highlighted the polysemic conflict provoked by a novel signifier such as the like button. Moreover, throughout the middle sections of this chapter we sought to broaden the spectrum of languages and visual signifiers that would be helpful to study on platforms.

We proposed three avenues of research to counter the English-language bias—one for each of the three dimensions of comparative work we address in this book. Regarding cross-national and regional studies, we suggested examining the use of social media in contexts of territorial displacement, both voluntary and forced, in which a high degree of linguistic diversity tends to be present. As Sirpa Leppänen and Ari Häkkinen (2013) have argued, “Within them [social media], communication and interaction are often multimodal and linguistically and discursively heterogeneous, such heterogeneity serving participants as a means for indexing identifications which are not organized on the basis of local, ethnic, national or regional categories only, but which are increasingly translocal. In social media practices, participants are thus orienting not only to their local affiliations but also to groups and cultures which can be distant but with which they share interests, causes or projects” (2013, 18).

Moreover, for cross-media scholarship we outlined a research direction focused on the processes whereby translations are produced, circulated, and received in interactions between traditional and social media. Finally, in terms of cross-platform studies we argued that accounts of variance of languages used in the different platforms could constitute a particularly fruitful terrain to explore larger societal issues at play in both continuities and discontinuities of experience across the ever-growing array of platforms that constitute the social media landscape.

As Barton and Lee (2013) argue, “instead of examining CMC [computer-mediated communication] from a solely monolingual, usually English, perspective, a growing body of research is interested in how speakers of various languages have adopted such new forms of writing to different extents” (6). Thus, we also suggested three avenues of research to complement the dominant textual focus of most social media scholarship and built on contributions from the domain of digital discourse studies (Herring 1996; Thurlow and Mroczek 2011; Thurlow 2018; Bou-Franch and Garcés-Conejos Blitvich 2019; Sumner et al. 2020). Regarding the cross-national and regional dimension, we outlined a series of strategies to examine the production, circulation, and reception of novel signifiers in different geographic locales. In addition, to further develop cross-media comparative work, we offered alternatives to inquire into how the novel signifiers are represented and used in the context of traditional media, and how social media represent and use visual elements specific to traditional media. Finally, concerning cross-platform studies, and in consonance with our approach to the dominance of English in accounts of textual communication, we foregrounded the use of novel signifiers in different social media.

Language is central to the constitution, in a relational fashion, of both personal and collective experiences: as Ferdinand de Saussure ([1916] 1983) argued in his seminars taught over a century ago, the value of specific words emerges from interactions among signs that

are contextually dependent. Applied to the study of social media this means that signifiers acquire different meanings in relation to other signifiers in the eyes of users situated in various parts of the world, encountering them on multiple media, and on diverse platforms. Because we signify comparing and we compare signifying, the study of language is an ideal way to build the epistemic perspective we advocate in this book.

This is a section of [doi:10.7551/mitpress/13999.001.0001](https://doi.org/10.7551/mitpress/13999.001.0001)

To Know Is to Compare

Studying Social Media across Nations, Media, and Platforms

By: Mora Matassi, Pablo J. Boczkowski

Citation:

To Know Is to Compare: Studying Social Media across Nations, Media, and Platforms

By: Mora Matassi, Pablo J. Boczkowski

DOI: [10.7551/mitpress/13999.001.0001](https://doi.org/10.7551/mitpress/13999.001.0001)

ISBN (electronic): 9780262374972

Publisher: The MIT Press

Published: 2023

The open access edition of this book was made possible by generous funding and support from MIT Press Direct to Open



The MIT Press

© 2023 Massachusetts Institute of Technology

This work is subject to a Creative Commons CC BY-NC-ND license.
Subject to such license, all rights are reserved.



The MIT Press would like to thank the anonymous peer reviewers who provided comments on drafts of this book. The generous work of academic experts is essential for establishing the authority and quality of our publications. We acknowledge with gratitude the contributions of these otherwise uncredited readers.

This book was set in Stone Serif and Stone Sans by Westchester Publishing Services.

Library of Congress Cataloging-in-Publication Data

Names: Matassi, Mora, author. | Boczkowski, Pablo J., author.

Title: To know is to compare : studying social media across nations, media,
and platforms / Mora Matassi and Pablo J. Boczkowski.

Description: Cambridge, Massachusetts : The MIT Press, [2023]. |

Includes bibliographical references and index.

Identifiers: LCCN 2022030583 (print) | LCCN 2022030584 (ebook) |

ISBN 9780262545938 (paperback) | ISBN 9780262374989 (epub) |

ISBN 9780262374972 (pdf)

Subjects: LCSH: Online social networks—Study and teaching. |

Social media—Study and teaching.

Classification: LCC HM742 .M365 2023 (print) | LCC HM742 (ebook) |

DDC 302.231—dc23/eng/20220707

LC record available at <https://lcn.loc.gov/2022030583>

LC ebook record available at <https://lcn.loc.gov/2022030584>