Article

And She Be like 'Tenemos Frijoles en la Casa': Code-Switching and Identity Construction on YouTube

Michael Wentker [1],[*] and Carolin Schneider [2],[*]

1   Department of Anglophone Studies, University of Duisburg-Essen, 45141 Essen, Germany
2   Department of Languages, Literature and Communication, Utrecht University,
     3584 CS Utrecht, The Netherlands
*   Correspondence: michael.wentker@uni-due.de (M.W.); c.schneider@uu.nl (C.S.)

Abstract: This empirical case study explores the (co-)construction and negotiation of identities through code-switching (CS) as found on the video-sharing platform YouTube, disentangling the complexities of social practice anchored in a discursive online environment. Drawing on a YouTube comment corpus and paying special attention to the socio-technical affordances of the platform, the study examines users' positioning practices and metapragmatic replies in response to a culturally themed video priming discussion about LatinX family stereotypes. More specifically, it analyses how users discursively position themselves vis-à-vis the video and which linguistic strategies they exploit to (co-)construct and negotiate their cultural identity. Focusing on interrelated positioning devices such as code-choice, identity labels and quoting, this contribution proposes a multi-level model of analysis to account for the dynamic interplay between CS practices and identity construction in a heterogeneous online space. Following a social-constructivist approach to identity, CS is shown to reinforce in-group solidarity rooted in the shared experience and discussion of LatinX culture and provides evidence of a sense of togetherness in an emerging community of practice.

Keywords: code-switching (CS); computer-mediated communication (CMC); community of practice (CofP); digital discourse; identity construction; metapragmatics positioning; quoting; YouTube

1. Introduction

As observed by Bailey, "language is our primary semiotic tool for representing and negotiating social reality, including social identity categories" (Bailey 2000, p. 341). This paper focuses on the language employed on the social video-sharing platform YouTube (Johansson 2017) by a diverse set of international users. It investigates the discursive code-switching practices on YouTube by an emergent community of practice (CofP) that is jointly produced in discourse. According to Eckert, "[a] community of practice is a collection of people who engage on an ongoing basis in some common endeavor [. . . developing] ways of doing things, views, values, power relations, ways of talking" (Eckert 2009, p. 683). The two major pillars, or 'conditions' a CofP needs to meet are "shared experience over time and a commitment to shared understanding" (Eckert 2009, p. 683).

Focusing on the comments section of a BuzzFeed video titled 'Things our Latina moms say', this research is interested in the complex dynamics of code-switching (CS) and identity construction. In more detail, it investigates the various linguistic practices in which users position themselves in the digital discourse, illustrating both users' (implicit) assumptions of who they are and their (explicit) showing who they are. Considering various distinct, yet interconnected levels of identification, we propose a model of cultural and linguistic positioning which illustrates how users position themselves on various continua of identification, employing strategies of accommodation and inclusion to create a shared experience. While transgressing in- and outgroup boundaries, users are shown to ultimately work towards a shared LatinX[1] identity, i.e., establishing a mutually accessible community of practice.

The structure of this paper is as follows. The next section addresses the complex notions of identity and code-switching, and the affordances of YouTube, all of which lay the theoretical foundations for our analytic framework. Subsequently, Section 3 introduces the data and methodology used to conduct the analysis. The following Sections 4–6 present the findings of our analysis and discuss them in the light of community of practice before the main findings are summarised in Section 7, providing an outlook and impetus for future research.

2. Approaching Identity and Code-Switching on YouTube

As rightly pointed out by Joseph, "[. . .] our very sense of who we are, where we belong and why, and how we relate to those around us, all have language at their centre" (Joseph 2010, p. 9). This quote illustrates that we can use language as a tool to express who we are; and that language use is a powerful resource to explore our identities. In other words, CS, broadly understood here as the use of more than one language in interaction, and identity are two inseparably linked concepts. This is especially true for context-dependent communicative spaces such as the YouTube comments section, where visual cues usually present in face-to-face contexts are missing. There, identification depends solely on the linguistic output and background assumptions made about interlocutors. This section elaborates on each of the concepts consecutively.

2.1. Identity Construction

In this paper, we adopt a broad definition of social identity, conceptualizing it as "the social positioning of self and other" (Bucholtz and Hall 2010, p. 18). In line with social-constructivist research on the topic, we understand identity as something that is constructed rather than essential; something that is performed rather than possessed and dynamic rather than fixed. In addition, we understand identity to be emergent in discourse. When engaging in identity work with others, users commonly position themselves discursively in various ways.

As observed by Davies and Harré, "[a]n individual emerges through the processes of social interaction, not as a relatively fixed end product but as one who is constituted and reconstituted through the various discursive practices in which they participate." (Davies and Harré 1990, p. 46). In other words, discursive participation in the form of comments constitutes user positioning in an emergent polylogue, where users fluidly shift between stances and relationships, ultimately negotiating their own as well as other users' identities.

Users can perform or index facets of identities in various ways, for example by overtly mentioning identity categories and (self-)labels but also through their personal linguistic choices, such as code choice, among others, which may be indexically tied to identity categories. This means users can explicitly tell their interlocutors who they are or—rather implicitly—show them who they are.

Five Principles of Identity Construction

To understand identity from a socio-pragmatic angle, it is important to stress that identities are always constructed intersubjectively, i.e., in relation to one another, and that any identity account is in a way partial. What is more, identities can shift, i.e., they can be discursively negotiated and re-negotiated. In line with Bucholtz and Hall (2010), these conceptions about identity and identification are subsumed under the following five principles.

According to the emergence principle, identity is emergent in discourse, rather than pre-existing (cf. Bucholtz and Hall 2010, p. 19). Especially in a social media environment such as YouTube comment spaces where phenotypic cues are missing and interlocutors are taken to be unknown to and unfamiliar with each other, self- and other-identification is heavily dependent on users' linguistic output and subject to the unfolding discourse.

The positionality principle is rooted in what Bucholtz and Hall call "(a) macrolevel demographic categories; (b) local, ethnographically specific cultural positions; and (c) tem-

porary and interactionally specific stances and participant roles" (Bucholtz and Hall 2010, p. 21). While categories (a) and (b) comprise overarching demographic and cultural attributes, e.g., age or LatinX identity, (c) constitutes more locally produced or expressed stances by linguistic means.

Furthermore, identity is firmly rooted in the indexicality principle. According to this principle, identities are produced linguistically by means of various indexical processes, e.g., the use of labels or identity categories, evaluative stances towards the unfolding discourse, as well as "the use of linguistic structures and systems that are ideologically associated with specific personas and groups" (Bucholtz and Hall 2010, p. 21). The latter category goes beyond the mere act of choosing a code for communication and especially includes any code-mixing behaviour. The indexicality principle serves as the foundation of our analytic model (cf. Section 4).

The relationality principle proposes that identification is an intersubjective process that is based on dichotomies such as similarity and difference, which play a major role in the current analysis. Users make sense of self and others by intuitively or strategically comparing themselves to other discourse participants and thereby reinforcing this relation (cf. Bucholtz and Hall 2010, p. 23).

The final principle (Bucholtz and Hall 2010, p. 25), the partialness principle, asserts that the emerging identities are never complete and merely offer partial snapshots of enacted, situated identities which are influenced by a multitude of internal motivations and external factors. This is especially true also for the findings of this analysis and stresses, once more, the fluid nature of the identity product(s) resulting from the process of identification, which is fuelled by versatile linguistic strategies employed by an international discourse community triggered by various communicative needs.

2.2. Code-Switching

As mentioned above, indexicality is a major principle in identity construction, particularly when it comes to style and the individual code choices made by the different users. It is commonly established that code-switching (CS) is a common practice in oral communication between multilingual interlocutors, hence, it is hardly surprising that the phenomenon can also be found in the digital sphere. Written CS is particularly evident in online social media platforms that combine spoken and written linguistic features in complex multimodal settings. However, we agree with the line of argumentation presented by a number of scholars (e.g., Barasa 2016; Androutsopoulos 2013; Spitzmüller 2006) who caution against assuming the functions of spoken and written CS to be the same. While digital communication can be as spontaneous as spoken language, some discursive acts require conscious planning and hence can be identified as distinct to the medium to which they are tied (cf. Barasa 2016, p. 67). Relating back to Koch and Oesterreicher (1985, 2012), we assume that written CS, therefore, can be more intentional and planned regarding positioning and identity work, given it is not as immediate as spoken language (cf. Koch and Oesterreicher 1985, 2012).

In the classic definition of CS, it is understood as "[. . .] includ[ing] elements from two (or more) language varieties in the same clause [. . .]" (Myers-Scotton 2006, p. 241). However, clauses as unit boundaries are problematic and pose a methodological challenge in approaching our dataset. In this study, we therefore adopt a definition of CS as the combination of language material from more than one language (variety) within one comment. Due to the strong focus on the link between CS and identity construction in this study, structural differentiation (e.g., Muysken 2007) and respective debates within the field are neglected. For an overview of these, see Boztepe (2003).

The link between code-switching and identity construction in social media has gained scholarly interest and generated a multitude of studies, e.g., Hernández (2016), Wentker (2018), and Pérez-Sabater (2022), to name a few. To date, however, no functional approach has been modelled specifically for novel dynamic CMC contexts (Barasa 2016). Therefore, this study focuses on the three established functions of CS as identified by Gumperz (1977,

1982) and summarised concisely by Bailey (2000, p. 348): situational switching (provoked by a change of context), discourse contextualisation switching (structuring discourse by cues through CS), and metaphorical switching (invoking more than one cultural tie).

> Elements of setting, participants, activities, or perspectives that are conventionally associated with a code can be invoked by a switch into that code when such elements are not otherwise present or active in the conversation. Changes in language can thus constitute alternative cultural frameworks for interpreting experience and constructing social reality. (Bailey 2000, p. 349)

In other words, a switch to a different code widens the cultural scope of the interaction at hand, as certain cultural associations can be carried by, e.g., the Spanish code, which might otherwise not be invoked by the sole use of the English code.

As we will demonstrate, discourse contextualisation and metaphorical switching are particularly prevalent in our dataset as a wealth of quoting strategies interlink with the construction of a shared cultural experience. Culture in this context is to be understood as both, "[. . .] a manifestation of a group, or a community, and of an individual's experience within it, or apart from it" (Levy 2007, p. 105). Hence, community members engage in a 'shared social space' (Levy 2007, p. 105) which is constructed and negotiated. Linguistic strategies expressed in more than one code, for example positioning, narrating, and quoting define this fluid shared social space. Aspects of belonging are established through shared experience, which is achieved "[. . .] not through reference but by giving a representative direct quotation of their speech" (Bailey 2000, p. 342).

2.3. YouTube

Before introducing the data and methodology, a brief consideration of YouTube's sociotechnical affordances is necessary to better understand the context in which the analysed data was produced. Due to various intertwined technical and social factors, YouTube as a social video-sharing platform allows for the exchange of views across national boundaries, providing us with contributions of a diverse set of social actors. Burgess and Green (2018) understand YouTube as a site of participatory culture, which is an important but very complex domain for research. In their seminal studies, they conceive of YouTube as an example of popular culture, vernacular creativity, and, ultimately, a space for emerging communities. Other scholars, such as Benson (2017), highlight the notion of YouTube as a multimodal text that is expressed in various discursive spaces. This study adopts a socio-pragmatic approach to YouTube (cf. Johansson 2017) while focusing mainly on the discursive practices on the site.

In digital communication research, it has been proven useful to consider the technical and social factors of the discourse environment under investigation (see Herring 2007; Page et al. 2014). Among the social or *situation* factors, Herring (2007) lists *participation structure*, *participant characteristics* and *code*, all of which prove essential to the analysis in our study. We classify YouTube as a site of public (as opposed to private) communication (cf. Bös and Kleinke 2017) with a global, heterogeneous viewership, with users belonging to a multitude of cultures, with assumed recourse to a broad range of languages.

Elaborating on the technical or *medium* factors, communication in the YouTube comment spaces is mostly asynchronous, despite often being realised in a quasi-synchronous manner, and offers a persistent transcript, i.e., users have permanent recourse to fellow users' messages (if not deleted post-creation) and YouTube offers an interactive address-and-reply infrastructure, enabling users to refer to each other's contributions and thus more conveniently navigate the discursive space. Furthermore, YouTube comment sections are multimodal in the sense that they can contain text and emojis. In addition to that, they are multimodal in another sense; even though the comments are materially written, they are ever-increasingly conceptually oral (cf. Koch and Oesterreicher 1985, 2012). In other words, users tend to approximate styles of speaking in their writing. For an overview of linguistic features creating linguistic immediacy, see Bös and Schneider (2021, p. 94f).

While social and technical factors can be considered separately, they need to be regarded as being ultimately intertwined, i.e., complementing each other. In the domain of cultural and linguistic identity, for example, the degrees of (social) self-presentation and self-disclosure are dependent on the technical affordances of YouTube, e.g., the degree to which the platform allows for users to present themselves in terms of media richness (cf. Kaplan and Haenlein 2010).

What is more, YouTube is highly interactive on various levels, producing a complex terminology to refer to its users at the production and reception end of communication (see Dynel 2017). This research draws on the neutral terms *user* and *participant* as umbrella terms to include all people using the platform and/or participating in the discourse. In accordance with Dynel (2017), we understand communication on YouTube as mediated interaction'. More specifically, this comprises a communicative setting "that simulate[s] the impression of direct contact between participants who are removed in space and time" (Dynel 2017, p. 63), enabling intercultural encounters between participants in an online polylogue who would otherwise not have met outside the digital sphere.

It is important to note the following:

> YouTube commenting can create networks of audiences and different kinds of communities, depending on the topics covered by the respective videos as well as on what users as members of various communities actually do with their comments. (Johansson 2017, p. 174)

Such networks or communities on YouTube are marked by (pseudo-)anonymity or (pseudo-)familiarity. On the one hand, users do not necessarily know each other by merely having registered on the site. On the other hand, however, it can be assumed that by engaging in discourse (regularly), people develop a sense of who they are or portray themselves to be, and thus become increasingly familiar with each other in any given comment space below a particular video or set of videos. Given that channels on YouTube, like Buzzfeed, rely on subscriptions, the same pseudo-familiarity can be assumed for the viewership, including the participants in the comment section. This pseudo-anonymity enabled by the platform inevitably leads to positioning of the self within the comment section, where social actors, by using language, reveal layers of their identity, e.g., being or not being LatinX, relating or not relating to the contents presented in the video or comments made by others. Meanwhile, users negotiate a particular type of group identity, hence forming a virtual community of practice (cf. Section 6.4).

3. Data and Methodology

Our dataset consists of 6474 (4033 top-level, 2441 replies) comments which are reactions prompted by the Buzzfeed video "Things Our Latina Moms Say" (https://www.youtube.com/watch?v=FHVxWPdZxJk; accessed on 2 May 2022) published on 14 August 2015, which by the time of extraction had generated almost 4 million views.

In this study, we adopt a mixed-method approach using the software MaxQDA. Pursuing an iterative, i.e., inductive, as well as deductive approach (Dörnyei 2007, p. 243), the analysis draws on various research methodologies, e.g., content analysis and lexical keyword-in-context (KWIC) searches, to qualitatively investigate emergent trends.

First, code choices in the data were manually tagged and categorised by language (pairs). The quantitative findings are used to investigate how the two major codes, i.e., English and Spanish, are represented overall in the data and to quantify how code choice is used as a linguistic tool to elaborate and strengthen the pillars of a shared identity.

Zooming in on the emergent patterns obtained by a close reading of comments, the data were analysed from a discourse-analytic angle with a qualitative, metapragmatic focus. Sections 4–6 present selected examples from the dataset to show how a sense of togetherness is constructed and how group identity is negotiated among the members of that emerging community.

Following a quantitative methodology, we measured the users' uptake of cultural concepts and lexical patterns as prompted by the video and previous comments (cf. Section 4.1)

by means of keyword searches. Automated tagging was complemented by manual tagging to further refine and disambiguate the various categories.

It is noteworthy that the title of the video already primes the discussion in several ways. Given the personal pronoun *our*, the title already indicates the common aspect of LatinX heritage. Since users on YouTube navigate the site by clicking on subscribed content or using the search function, the title already attracts a certain viewership (cf. Section 2.3) and hence also predefines the imagined boundaries between the dichotomy of who is "we" and who are "they", which is an important prerequisite to evaluate the findings of our analysis. Furthermore, it is essential to point out that the dominant language of the video is English, and when speakers code-switch to Spanish in the video, subtitles in English are provided.

In our analysis, all examples are given in their original orthographic format, except that Spanish content was marked in italics by the authors. For convenience, English translations were added in square brackets and inverted commas following the Spanish element(s).

4. Analysis I: Cultural and Linguistic Positioning

For the analysis of the data, we propose a model of cultural and linguistic positioning (Figure 1) which assumes that each comment posted in response to the video is an act of user positioning on various axes or continua. We differentiate between two major levels of positioning. There is a linguistic level and a cultural level, both of which are briefly introduced now and discussed with examples in more detail below.
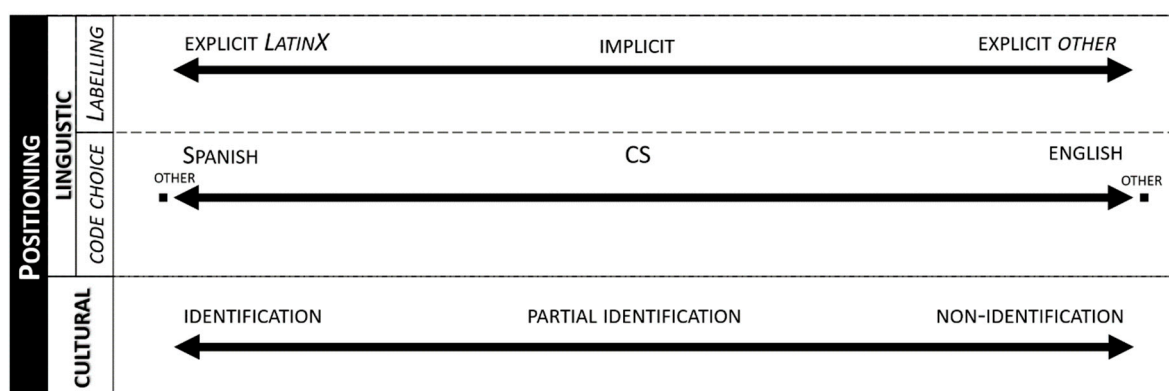

Figure 1. Continuum Model of Cultural and Linguistic Positioning.

On the linguistic level, bilinguals can draw on forms from two languages as well as hybrid forms. The linguistic level can again be divided into two sub-levels of expression, i.e., labelling and code choice. We thus understand labelling and code-switching as resources that are employed by (multilingual) users to position themselves. For example, users employ explicit labels of belonging (to 'LatinX') via implicit labels of (non-)belonging and explicit labels of non-belonging (cf. Section 4.2.1). By having recourse to various linguistic codes, users opt for a certain code or combination of codes to indexically position themselves on a continuum of code choice ranging from monolingual Spanish to monolingual English. In between there are comments that contain both languages which are classified as CS (cf. Section 4.2.2).

On the (socio-)cultural level, many users seem to be familiar with relatively diverse cultural frameworks for interpreting and evaluating the world and placing themselves and others within it. Again, this can be conceived of as a continuum which ranges from cultural identification via partial identification to non-identification with the LatinX culture (cf. Section 4.1).

4.1. Cultural Positioning

The video addresses various recurring phrases and situational family contexts, prompting users to identify themselves through this mutually shared cultural experience. It is essential to point out, however, that the content of the video—given its purpose of entertainment—provides a very limited window into what is labelled here as LatinX culture.

Simple expressions of agreement in the form of one-word comments such as *Yessss!*, *Yup*, *Yeeees* and brief evaluations, e.g., *100%*, *accurate*, *true*, *SOOO true* as well as the use of emoticons as a reaction to the video as a whole indicate the general tendency of identification with its content.

The specific stories mentioned in the video trigger the exchange of lived individual experiences, which are vividly discussed by the users. In the comment section, users prominently take up topics such as disciplinary actions (e.g., *chancla*, *cinturón*), food (e.g., *frijoles*), idiomatic expressions (e.g., *sana, sana culita de rana*) and linguistic idiosyncrasies of 'Latina moms' (e.g., *taki taki*), as illustrated in Table 1. In a total of 6474 comments, lexical searches provide us with the following numbers of uptakes which demonstrate the relative importance of these concepts and their potential for discursive identification or rejection.

Table 1. Instances of Cultural Concepts Mentioned in the Comments.

Cultural Concept	Lexical Item (as Featured in Video Prompt)	Uptakes in No. of Comments
Disciplinary actions	*chancla* ['sandal']	439
	sandal | *flip flop*	36 | 27
	cinturón | *cinto* ['belt']	95 | 49
	belt	105
Food	*frijoles* ['beans']	86
	beans	95
Idiomatic expressions	*sana, sana culita de rana* ['heal, heal, tail of a frog']	310

A wealth of comments demonstrates the nuanced ways of identification, reaching from complete identification via partial-identification to non-identification as depicted in the continuum illustrated in Figure 1.

On the one hand, comments illustrate that the user identifies with the prompts featured in the video. For example, (1) emphasises the perceived level of accuracy by means of reduplication (*sooooooooooooooo*) (cf. Danet 2010; Crystal 2006). On the other hand, a rejection of the mentioned parental disciplinary actions can be found in comments such as (2).

(1)　this is sooooooooooooooooo accurate "*chancla*" ['sandal'] and my mom still says the "*sana sana colita de rana, si no sana hoy, sanara mañana!*" ['heal heal tail of a frog, if it doesn't heal today, it will heal tomorrow!']

(2)　I never got hit. At all.

(3)　I am Mexican and American I don't know what's a *chancla* ['sandal'] can someone tell me what that means

Further, instances of non-identification can be found, for example when the cultural concept is unknown to the user (3). Most interestingly, we find cases of partial identification, where users identify with the cultural group, without being familiar with the cultural theme: *I am Mexican and American . . .* (3), or vice versa, *I'm not even Hispanic* (4). Comments show partial identification, where overlappings between cultures are pointed out, e.g., in example (4) the user expresses not to be Hispanic but to be familiar with both, the infamous saying 'we have beans at home' and the concept of 'la chancla'. The user in example (5) explicitly identifies as a member of the Arabic community and shows familiarity with *la chancla*.

(4)   *"tenemos frijoles en la casa"* ['we have beans at home'] 😂 😂 so TRUE I ALWAYS GET THAT IM NOT EVEN HISPANIC AND MY MOTHER USES A *CHANCLA* ['sandal']

(5)   In Arabic we know *la Chakla* ['the sandal'] almost

In addition to identifying with the concepts mentioned, users find additional examples to relate to the video. While the user in comment (6) shows familiarity with the concept, they also add that their mom merely needed 'the stare' instead of 'the sandal'. Implicit familiarity with the addition of 'the belt' is expressed in example (7).

(6)   Forget about "la chancla" ['the sandal'], my mom only needs "*la mirada*" ['the stare'] . . . . (¬_¬) . . . .

(7)   For me it was the *cinturon* ['belt'].

(8)   Did anyone's mom ever say *vas a ver en la noche el Diablo te va jalar las patas* ['you'll see in the night the Devil will yank your feet']

(9)   They missed something Latina moms say: *a mi por menos de ahí me hubiesen roto la boca* ['I would have had my mouth torn for less than that']. Which translate to for less than that my mom would've kill me [. . .]

We argue that this shared cultural experience and the discussion thereof is a crucial prerequisite for a virtual CofP. This becomes particularly evident in comments which invite other users to participate in the discourse, e.g., when asking for similar experiences, e.g., *Did anyone's mom ever say* (8), clarifications (3) and missing communalities, e.g., *They missed something Latina moms say* (9).

4.2. Linguistic Positioning

It is important to point out that the users in our dataset are not necessarily multilingual. However, even users who merely have recourse to one code position themselves by its use when commenting. Even though this might not be a conscious decision, every instance of language use is indexical or metaphorical of the speaker's identity. This section elaborates on two manifestations of linguistic positioning consecutively: *labelling* and *code choice*.

4.2.1. Labelling

The category of labelling concerns itself with the use of labels as a means of identification. When it comes to positioning via self-labelling, some identity claims are conscious, intentional, and referentially explicit. As is evident in examples (10)–(16), users explicitly position themselves on a continuum of being LatinX or belonging to another cultural group. In between those two poles, we find identity claims that are rather implicit (12)–(14).

(10)   This is actually my mom *porque yo es Latina* ['because I am Latina']

(11)   My mom is Latina and I'm Latina too

In (10–11), explicit labels such as *Latina* are used which indicate the speakers' belonging to the LatinX culture. In both examples, the cultural reference works via the shared family home and an other-identification, i.e., *This is actually my mom* (10) and *My mom is Latina* (11), which are then essentially complemented by a clear, more explicit self-identification *porque yo es Latina* (10) and *I'm Latina too* (11). Here, a direct link between family identity and personal identity is established.

(12)   Growing up Hispanic . . .

(13)   #I'm latino but I don't speak spanish!

(14)   My mom says the clean your room thing and in Spanish [. . .]

Examples (12)–(14) are more vague and rather implicit manifestations of labelling. In (12), the label *Hispanic* indicates a potential identification without explicitly invoking a clear LatinX identity. Example (13) displays partial identification with the LatinX culture. On the one hand, the label *latino* is used to identify with said group; on the other hand, however,

the user self-proclaims that there is no knowledge of the associated linguistic varieties. In example (14), the identification is indirect and solely works via the user's mother. It can be compared to (13) but features no subsequent self-identification. The user thus does not use a self-label and the own identity status remains relatively implicit.

(15)	I am Filipino
(16)	This makes me want to have a Latina mom!

In contrast to that, examples (15) and (16) illustrate explicit and implicit other-identification. In (15), the user employs a label which indicates their identification with a non-LatinX culture, while the other-identification works more implicitly in (16), where the user indirectly states that they do not have LatinX parents while expressing a desire or affection to LatinX culture.

4.2.2. Code Choice

As established in Section 2.1, linguistic choices are indexical of a user's identity; by drawing on their communicative repertoire (Casas 2016, p. 44), users identify themselves through monolingual and bilingual code choices. Table 2 below outlines the various code choices found in our dataset.

Table 2. Code choice distribution across all comments.

Code Choice	No. of Comments	Percentages
English only	3905	60.32
Spanish only	495	7.65
English-Spanish	1775	27.42
English-other language	15	0.23
Spanish-other language	3	0.05
English-Spanish-other language	1	0.02
Other language only	6	0.09
Miscellaneous (e.g., non-verbal)	276	4.26
Total	**6474**	

The dataset is composed of a majority of English only comments (60.32%), Spanish only comments (7.65%), and comments which include language material from both English and Spanish (27.42%). A minority of comments can be found combining language material from English (0.23%) or Spanish (0.05%) with another language as well as a small number of monolingual comments in a language other than English or Spanish (0.09%). We converged comments which only include emoticons, at-mentions, or unintelligible comments into the miscellaneous category (4.26%). Given the languages used in the video prompt (English as dominant language, with instances of Spanish), the lion's share of all comments contains English language material (87.98%), whereas Spanish, notably, is found in approximately one third of all comments (35.13%).

The level of code choice (cf. Figure 1), again, stipulates a continuum with two proposed poles, one being Spanish only, e.g., *A mi me lanzan la chancla* (18), the other being English only, e.g., *Nothing will ever scare me more than that look* (21). However, the two poles themselves are to be placed on an extended continuum, given that we find Spanish-Portuguese (17) and English-Dutch (22) combinations. Nevertheless, the focus here lies on the different realisations of CS between English and Spanish, with special attention to how CS is naturally used, presupposing different degrees of familiarity not only by the users themselves but also by other members of the YouTube community.

(17)　A […] *Aquí en Brasil llamamos de "chinelo" o "chinela". Si tu haces mierda y tu mamá agarra eso … … . .* ['Here in Brazil we call it "chinelo" or "chinela". If you goof up and your mum catches that']

(18)　*A mí me lanzan la chancla voladora o si eso falla el fajo nunca falla* ['I get the flying sandal or if that fails, the belt never fails']

(19)　Every Time I cried, my mom told me: "*Ya vas a ver, en la casa vamos a cortar el queso*" ['You'll see, at home we're going to cut the cheese'] […]

(20)　the *chancla* ['sandal'] is the scariest thing on earth

(21)　Nothing will ever scare me more than that look my mom gives once you annoyed her

(22)　My mother always used to say: "*Koppie stoot, luisie dood*".Wichtranslates from Dutch to: "Hit your head, lice are dead".

Across the continuum, we find various formal realisations of CS. Users make use of one-word insertions such as *the chancla is the scariest thing on earth* (20) or they switch intersententially: *Every Time I cried, my mom told me: "Ya vas a ver en la casa vamos a cortar el queso"*(19). By including more Spanish elements within one post, the user positions other users and assumes they are members of the same language community. In doing so, one could argue that users position themselves as members of an in-group, thereby inadvertently positioning non-Spanish speakers in the (linguistic) out-group.

Given the written context of YouTube comment spaces, participants use a range of punctuation markers to formally mark CS. On the one hand, CS appears formally unmarked as in (20): *the chancla is the scariest thing on earth*. On the other hand, comments such as (17), *"chinelo"*, indicate a CS by means of quotation marks. Further, we find quoting strategies indicated by colons and quotation marks as in (3): *my mom told me: "Ya vas a ver, en la casa vamos a cortar el queso"*. Our data suggest that in this dataset, quoting is a dominant function of CS, used for sharing personal stories which can be woven into the broader cultural frame of the community.

5. Analysis II: Quoting as Metacommunicative Acts

The wealth of quotations found in the comment section is hardly surprising, given the frame by the video which prompted these reactions. The title *"Things our Latina moms say"* can be understood as an implicit call for adding to the corpus of things said by the viewers' mothers. Quoting includes "an act of taking up text and, in doing so, performing a shift of context, focus and perspective" (Bublitz 2015, p. 2). This process of revaluation subsumes the metacommunicative acts of re-contextualising, re-focusing, and reflecting. We understand quoting as the written representations of oral contexts. As a general note, we are not interested in whether a quote is accurate or not; we are instead interested in its specific function for the current communicative exchange. As we demonstrate below, quoting is also the dominant function of CS established by the users; therefore, we analyse the different surfacing quoting practices in code-switched comments in further detail. Furthermore, we focus on the functional perspective of a CS in this context while their formal realisation is merely of secondary importance.

Table 3 provides an overview of the various strategies of quoting in our dataset, indicating that quotation marks are the most frequent strategy applied. In the total of 1775 comments that include CS (cf. Table 2), 966 sets of quotation marks, 1002 verba dicendi, and 586 echo quotes were found. However, it is important to note that the total of 2554 instances contain multiple tagging, i.e., more than one quoting strategy may occur per comment. Table 4, therefore, details the various overlaps, i.e., the multiple strategies employed per comment. Note that more than one instance of quotation marks or multiple verba dicendi may be found in one comment. It becomes evident that the majority (n = 1412; 79.55%) of code-switched comments contain at least one quoting strategy.

Table 3. Types and Instances of Quoting in Code-Switched Comments.

Quoting Strategy		Number of Instances
quotation marks	" … "	966
verbum dicendi	*say*	773
	answer	14
	ask	77
	yell	45
	quotative *like (e.g., go/be like)*	93
echo quote	A: …	586
	B: …	
Total		**2554**

Table 4. Distribution of Quoting Strategies across Code-Switched Comments.

(Multiple) Quoting Strategies per Comment	Number of Comments	% of CS Comments
quotation marks (only)	178	10.03
verbum dicendi (only)	329	18.54
echo quotes (only)	234	13.18
quotation marks + verbum dicendi	362	20.39
quotation marks + echo quotes	149	8.39
verbum dicendi + echo quotes	103	5.8
quotation marks + verbum dicendi + echo quotes	57	3.21
no apparent quoting strategy	363	20.45
Total comments including CS	**1775**	**100.00**
Total CS comments including quoting strategies	**1412**	**79.55**

Motivations for quoting in our data are, arguably, creating narrative immediacy, simulating dialogicity, but most dominantly—what Bublitz and Hoffmann (2011, p. 443) call "indicating affiliation/'devaluation' of the quoted text". Generally, the quotes or 'quote-switches' serve as an anchor of shared LatinX experience, reiterating dialogic exchanges between users and their mothers, often including themselves in the form of echo quotes.

(23) Or when you spend 4 hours after she said that you were leaving in 5 min and after you ask her to leave she be like "*hay como jodes hijode tu pinche madre*" ['how you fucked up you son of a bitch'] hahaha I love my crazy mom

(24) when you're in public
me: behaves bad in public
my mom: *Voy a decirle a ese policía que te lleve* ['I'm going to tell that policeman to take you'] (then says not directly to the police but acting like she is) *policía toma a este niñodesobediente* ['police take this disobedient child']
me: *no no no por favor mama te amo* ['no no no please mom I love you']
my mom: *que te calmas entonces* ['you calm down then']

(25) Every time I wanted to do something and my mum thought I'd get hurt, she'd be like "*Ándale pues, pero si te caes te voy a agarrar a chingazos*" ['go on then, but if you fall, I'm going to beat the shit out of you'] and I be like "bruh why? if I fall and I'm already hurt, why would you give me *chingazos* ['punshes']?"

Apart from using quotation marks, e.g., in (25): *Ándale pues, pero si te caes te voy a agarrar a chingazos*, acts of quoting are realised by various strategies, for example by using reported speech including verba dicendi, i.e., verbs of speaking, in *Or when you spend 4 hours after she said that you were leaving in 5 min* (23). One of the most common verbs of speaking found in the dataset is the quotative like, e.g., in (25): *she'd be like … I be like*. The use of echo

quotes is also quite frequently found in the comments section. Echo quotes appear in the form of a written dialogue in which speakers, e.g., *me* and *my mom* (24), and turn-taking are indicated. On a formal level, these 'quote-switches' are discourse contextualisation devices and anchors of reader guidance that indicate a CS by marking it formally. On a functional level, these quotes work as additional back-up to users' identity claims in establishing users' credibility (Bednarek and Bublitz 2006, p. 551). Taken together, it becomes clear that these quoting strategies create coherence in building a joint cultural narrative and can be regarded as being inclusive efforts.

6. Discussion: Negotiation of LatinX Identity

While each of the above-mentioned levels can be analysed separately and provide examples of positioning, these levels are interconnected and cannot be interpreted separately from each other. They must be understood holistically as an instance of omnipresent positioning. Each comment is regarded as a micro-act of identification, i.e., a smaller piece of a larger puzzle, which feeds into a macro-level of identification. Furthermore, positioning is realised consciously and unconsciously as well as in an explicit or implicit fashion, and their respective nuances in-between.

In addition to the levels postulated above, there is another, overarching level, that of metapragmatic negotiation strategies (cf. Figure 2). In our study, we understand *metapragmatic* to comprise comments which show users' awareness of relational, i.e., socio-cultural, aspects, including the codes used, and comments which negotiate intercultural meanings and different levels of identification (cf. Tanskanen 2007; Bublitz and Hübler 2007).
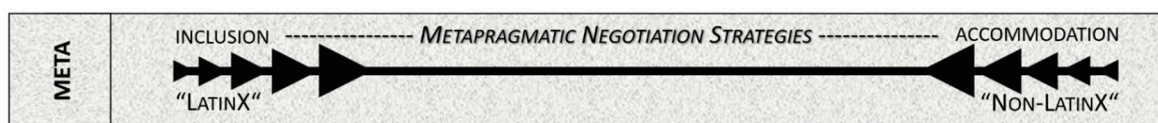

Figure 2. Metapragmatic Negotiation Strategies: Inclusion and Accommodation.

When posting a comment, users position themselves not only on a linguistics and/or cultural level but also on a continuum of metapragmatic negotiation, which work all at the same time. For example, by making observations about language use and code-choice on a meta-discursive level, the topic of CS is made explicit and becomes the focus of attention. For example, users may provide comments on the ongoing exchange and users' linguistic behaviour itself, thus expressing their metalinguistic awareness, and/or use situational code-switching (e.g., one-word insertions) to bring their intended meaning across. In other words, metapragmatic utterances either make code choice the topic being talked about and/or feature CS that is used to exemplify a point. CS and metapragmatic utterances thus have a propositional and/or a functional link.

More narrowly, we postulate two converging metapragmatic strategies, i.e., *inclusion* and *accommodation*. They are an outcome of positioning that takes place on the other levels. In our dataset, there is a trend towards a shared identity between "LatinX" and "Non-LatinX" users, which is achieved by complementing, mutually approaching negotiation strategies. The strategy of inclusion originates at the "LatinX" pole of the continuum and features metapragmatic comments aiming to include users from the other end of the continuum, i.e., users not identifying with LatinX culture. Conversely, strategies of accommodation originate at the other end and can best be understood as verbal utterances that move towards the "LatinX" pole, in which users display their open-mindedness about the LatinX culture and aspire towards it. Both strategies will be exemplified in the following.

6.1. Inclusion Strategies

On the left side of the continuum, located at the "LatinX" pole, we see various inclusion strategies at play. In (26), we see efforts made by LatinX users to provide translations of Spanish terms in brackets to create a shared understanding of the predominant cultural concepts. Example (27) highlights the user's metalinguistic awareness of not adequately representing one's own cultural group linguistically. The apologetic speech act shows the inclusion of one's own group and shares a potential vulnerability. Claiming that one's *writing/typing int spanish is really bad* can be seen as a partial disidentification, which, however, is not the decisive factor here.

(26)   Omg I remember *la chancleta* (the sandle) and *la coreya* (the belt) [. . .]

(27)   *Sana sana colita de rana si no sana oy sanara mañana* ['heal heal tail of a frog, if it doesn't heal today, it will heal tomorrow'] 🖐 if i missd accents or spelled it wrong i apologize my writing/typing int spanish is really bad

(28)   my mom whispering is more worse then yelling can anyone relate?

Those inclusive strategies are complemented by users who explicitly address other users, as depicted in example (28). Here, the user seeks agreement from the community: *can anyone relate?*

6.2. Accommodation Strategies

At the other end of the continuum, we see strategies of accommodation. In (29), the user also asks for similar experiences but from another point of view. Here, the user seeks affirmation from other users and invites contributions reporting about similar identification practices. A nomination (*any other white people . . .*) and a predication (*trying to translate the comments using the 2 years of spanish from middle school*) are combined to self-identify and their own interaction with the Spanish content. The final question (*just me?*) invites other contributors to accede to the user's lived experience.

(29)   any other white people here trying to translate the comments using the 2 years of spanish from middle school? just me?

Example (30) is metapragmatic in the sense that the user makes use of an echo quote of their own reaction (*Me: . . . . uhhhhhhhhhhhh . . . . . .*) to the Spanish code among the comments, expressing the efforts necessary to understand them. The second part of the comment (*AND THIS IS YET ANOTHER REASON WHY I'M TRYING TO LEARN SPANISH!!!!!!!!*) is an example of accommodation since the user expresses their willingness to learn or retrieve Spanish to be able to understand and actively participate in the discourse.

(30)   All the comments are in Spanish. Me: . . . . uhhhhhhhhhhhh . . . . . . AND THIS IS YET ANOTHER REASON WHY I'M TRYING TO LEARN SPANISH!!!!!!!!

An interesting pattern that emerges in the data concerns itself with cultural approximation. Comments such as (31)–(33) serve as examples of this type of accommodation. For example, non-LatinX users experience cultures such as Indian (31), African (32) and Albanian (33) as being similar to the LatinX one. By claiming that mothers across various cultures are the same, users highlight similarities and reduce differences between cultures, thus creating a sense of a more homogeneous discourse community with a shared identity across cultural boundaries. Such comments stress the common ground across members of the community.

(31)   Indian moms are the same

(32)   "We have beans in the house" is the African mother equivalent of "we have rice at home" 😂😂😂😂 my mums reply to every hunger request

(33)   My mom is exactly like this, 0.0001% different and my mom is Albanian

6.3. Metalinguistic Awareness

Comments such as (34) emphasise general metalinguistic statements that assess the linguistic practices of the whole community. In this example, the user expresses their emotional state of being, i.e., feeling *sad how some people don't speak Spanish anymore*. It points to the user's metalinguistic awareness. It is interesting to note that the user provides the English and Spanish version of the message, with an extended qualification in the Spanish part (*pues unos*), thus reserving this extended meaning component for Spanish-speaking users.

(34)   It's sad how some people don't know how to talk Spanish anymore: (es triste que estos muchachos ya no saben cómo hablar en español, pues unos ['it's sad that these guys don't know how to speak Spanish anymore, well, a few']

Finally, another striking example of the interconnectedness between both strategies of inclusion and accommodation and the dialogic nature of the interaction is provided in examples (35) and (36).

(35)   Is it Latina or Latino?
(36)   LatinA is for woman and LatinO is for men

In (35), a non-LatinX user asks for clarification with regard to terminology, which is politely answered by another user (36) in an attempt to work towards a sense of togetherness (*LatinA is for woman and LatinO is for men*). These and other metapragmatic exchanges illustrate that the community aims to reduce misunderstandings in a collaborative effort to create inter-cultural awareness and bridge cultural differences. Furthermore, it is striking that in our dataset, there are hardly any impolite comments or exclusive strategies which would lead to "othering", a phenomenon which is quite commonly found in various Internet spaces (cf. Aslan and Vásquez 2018).

6.4. Towards a Digital Community of Practice

Combining the findings from the various analytical levels, we postulate a holistic model of positioning (cf. Figure 3). To recapitulate, while cultural identification is an expression of cultural positioning on the linguistic and cultural levels (as expressed by labels and code-switching), we have identified a corresponding metapragmatic level featuring strategies of inclusion and accommodation, complementing the other ones. As we have exemplified, users are shown to widen the scope of shared (LatinX) identity from both sides. This includes stating why they feel a sense of (non-)belonging as well as demonstrating ways of loosening the narrowness of the term LatinX identity. The negotiations do not pertain to the terminology per se but focus on the metaphorical representation of what the concept LatinX denotes, e.g., a culturally-flavoured family context and upbringing.

All in all, the inclusion and accommodation strategies illustrate how LatinX identity is metapragmatically negotiated as a shared endeavour. The analysis further suggests that users do not explicitly make a distinction between an ingroup or outgroup, i.e., textually marking a difference between belonging ('us') and not-belonging ('them').

With regard to the movement along the axes, it can be stated that this is not measured for the linguistic and cultural levels as we have not measured users' positioning diachronically. However, on the third level, inclusion and accommodation can be conceived of as indicators of movement towards a shared identity. In the figure, this is indicated

by an emerging community of practice centred around the pole of LatinX culture, which is continuously extended on both sides to include the discourse community of users. A high degree of relatability and cooperation among users creates an inclusive and shared experience, working towards a digital CofP.
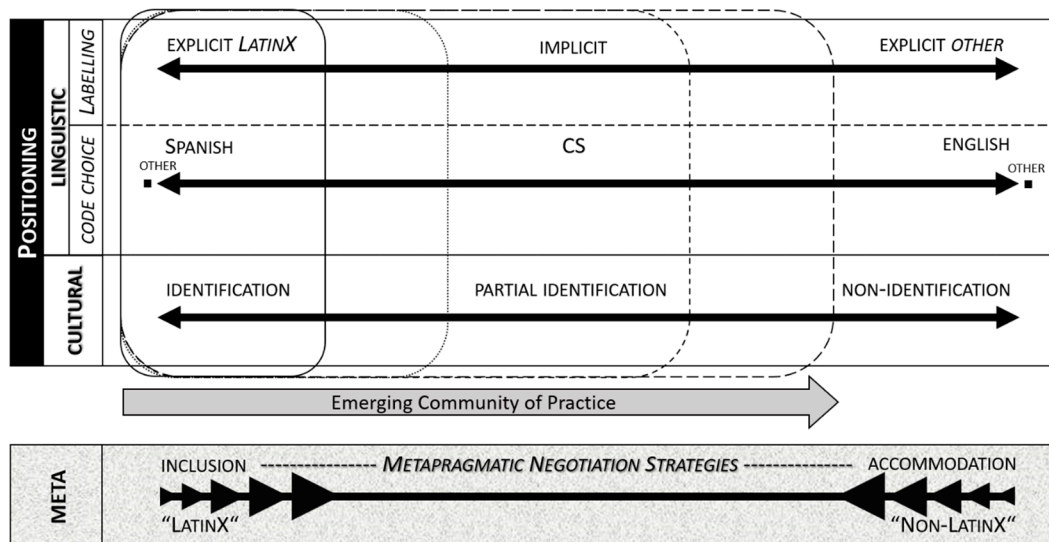

Figure 3. A Holistic Model of Positioning: An Emerging Community of Practice.

7. Conclusions

This study set out to investigate the link between English-Spanish code-switching and identity construction in the YouTube comments section of the culturally-framed video titled "Things our Latina Moms say". To serve as the foundation of the analysis, this study stipulates a model of cultural and linguistic positioning to illustrate and understand users' diverse positioning strategies and to capture the role of CS in the complex identity work at play in a heterogeneous audience on YouTube.

The findings indicate that on the cultural level, users identify with or reject the cultural concepts mentioned in the video, such as food, family upbringing or idiomatic expressions. On the linguistic level, users have been shown to position themselves either implicitly or explicitly in the form of (self-)labels as well as through their code choice. It is maintained that these categories are not mutually exclusive, but work simultaneously, while being highly interdependent and interwoven. Cultural positioning is shown to be fuelled by users' CS behaviour. In a collaborative endeavour, CS is used as a linguistic tool that goes beyond merely expressing identities. It is illustrated that CS also has the potential to negotiate identities, bridge intercultural gaps and create a discourse that is marked by inclusivity, rather than exclusivity, as well as inclusion and accommodation, rather than rejection of cultural positions and belonging. More specifically, CS is shown to reduce intercultural differences, equate various cultures, as well as reflect on users' own language behaviour metapragmatically. Focusing on the employed functions of CS, the empirical data illustrate that CS dominates the polylogue. CS is shown to be mostly discourse-contextualising, metaphorical, and, to a lesser extent, also situational. Since the various functions are at play simultaneously, it proves difficult to expose the dominant function. However, some general tendencies become apparent. As discussed, 'quote-switching' constitutes an intersection of discourse structure, thus creating reader guidance. Metaphorical switching is shown to enrich the discourse context and is inherently tied to users' (group) identities.

Overarchingly, switches into Spanish are metaphorical since these switches constitute the speaker's identity, thereby indicating cultural membership, association and/or affiliation. Rather than merely agreeing or disagreeing with other participants' propositions, the practice of using and especially switching into Spanish provides culturally-tinged examples that seem to activate cultural concepts in (fellow) discourse participants. This emphasises

once more that CS is not used to create an exclusive core of what it means to be LatinX but instead employed as a tool for the inclusion of others.

In addition to that, the analysis reveals that users engage in a number of quoting strategies, utilising quotation marks, verba dicendi, and echo quotes (cf. Table 4), thus adding to a collective narrative of a shared LatinX identity. The relative dominance of CS in conjunction with quoting strategies highlights the discourse-contextualising function of CS. These 'quote-switches' mark a distinction between what is said in the moment of production and what was said before. In this respect, it is not crucial whether the quote is an accurate reproduction of what was actually said. It is more important to mark the embedding of 'previous content' that is being recontextualised to adapt it to the current purpose of the exchange. The main function of this type is to formally mark this change in perspective, where CS and quoting cues (e.g., echo quote, colons, . . .) work in synergy to structure written discourse.

In addition to that, the study suggests an overarching metapragmatic level, exemplifying strategies of cultural inclusion and accommodation. The discourse community evidently displays a high degree of metalinguistic and metapragmatic awareness as emphasised in the preceding sections.

As main findings, the following points can be made:

1. Positioning is a common means of expressing cultural identity by using various (micro-)linguistic strategies, e.g., CS, among others. As the analysis demonstrates, those strategies are indexical of cultural identities on the personal level. The relational work (cf. Locher and Watts 2005) being invested by the discourse participants thus allows for establishing and reproducing (or contesting) aspects of social identities. By using, exploring or discussing the code associated with the LatinX community (i.e., Spanish) and engaging in CS behaviour between English and Spanish, participants orient towards the LatinX culture, actively participating in it and thereby strengthening its core.

2. In turn, the sum of communicative practices is indicative of a shared cultural identity on a group level, for which we propose the term of an (emergent) community of practice. It is emergent in the sense that it develops gradually, the more inclusive and the more accommodating the community acts by means of metapragmatic commenting. This community is not merely rooted in understanding the jointly constructed dialogue and sharing a cultural identity but is heavily characterised by (inter-)cultural awareness.

Finally, a mention of important limitations is in order. First, it needs to be noted that neither of the authors is a member of the LatinX culture or the particular YouTube community and therefore conducted this study from an outsider's perspective. Second, due to practical constraints, this paper cannot provide a comprehensive account of intralinguistic variation (e.g., lexical variations such as *fajo* for 'cinturón').

While the proposed model of cultural and linguistic positioning made an initial effort to uncover the complex interplay of the different levels at hand, there is room for further aspects which could be considered in future research, such as emoticons, terms of address, (im-)politeness, to name a few. Future scholars are invited to determine to what extent this model is applicable to other, maybe more hostile, online environments on YouTube and beyond, to which end the model can be extended, adapted, and further developed.

Data Availability Statement: Publicly available datasets were analysed in this study. The data were extracted on 18 February 2019.

Conflicts of Interest: The authors declare no conflict of interest.

Note

1   The term LatinX was chosen by the authors as an umbrella term to be as inclusive as possible, making no distinction between users' countries of origin and hence encompassing other existing labels (e.g., Hispanic, Latina/o, Latin@).

References

Androutsopoulos, Jannis. 2013. Code-switching in computer-mediated communication. In *Pragmatics of Computer-Mediated Communication*. Edited by Susan Herring, Dieter Stein and Tuija Virtanen. Berlin and Boston: De Gruyter Mouton, pp. 667–94. [CrossRef]

Aslan, Erhan, and Camilla Vásquez. 2018. Cash me ousside: A citizen sociolinguistic analysis of online metalinguistic commentary. *Journal of Sociolinguistics* 22: 406–31. [CrossRef]

Bailey, Benjamin. 2000. Social/Interactional Functions of Code-Switching among Dominican Americans. *Pragmatics* 10: 165–93. [CrossRef]

Barasa, Sandra. 2016. Spoken Code-Switching in Written Form? Manifestation of Code-Switching in Computer Mediated Communication. *Journal of Language Contact* 9: 49–70. [CrossRef]

Bednarek, Monika, and Wolfram Bublitz. 2006. Reported speech: Pragmatic aspetcts. In *Encyclopedia of Language and Linguistics*, 2nd ed. Oxford: Elsevier, pp. 550–53.

Benson, Phil. 2017. *The Discourse of YouTube. Multimodal Text in a Global Context*. London and New York: Routledge.

Bös, Birte, and Carolin Schneider. 2021. "We are all in this together"—Balancing virtual proximity and distance in online care partner discussions. *Anglistik* 32: 85–109. [CrossRef]

Bös, Birte, and Sonja Kleinke. 2017. Publicness and privateness. In *Pragmatics of Social Media*. Edited by Christian Hoffman and Wolfram Bublitz. Berlin and Boston: De Gruyter Mouton, pp. 83–121.

Boztepe, Erman. 2003. Issues in code-switching: Competing theories and models. *Studies in Applied Linguistics and TESOL* 3: 1–27.

Bublitz, Wolfram. 2015. Introducing Quoting as a Ubiquitous Meta-communicative Act. In *The Pragmatics of Quoting Now and Then*. Edited by Jenny Arendholz, Wolfram Bublitz and Monika Kirner-Ludwig. Berlin and Boston: De Gruyter Mouton, pp. 1–26.

Bublitz, Wolfram, and Axel Hübler. 2007. *Metapragmatics in Use*. Amsterdam and Philadelphia: John Benjamins.

Bublitz, Wolfram, and Christian Hoffmann. 2011. "Three men using our toilet all day without flushing—This may be one of the worst sentences I've ever read": Quoting in CMC. In *Proceedings: Anglistentag 2010 Saarbrücken*. Edited by Joachim Frenk and Lena Steveker. Trier: WVT, pp. 433–47.

Bucholtz, Mary, and Kira Hall. 2010. Locating Identity in Language. In *Language and Identities*. Edited by Carmen Llamas and Dominic Watt. Edinburgh: Edinburgh University Press, pp. 18–28.

Burgess, Jean, and Joshua Green. 2018. *Youtube: Online Video and Participatory Culture*, 2nd ed. Hoboken: John Wiley & Sons.

Casas, Marisol Pérez. 2016. Codeswitching and identity among Island Puerto Rican bilinguals. In *Spanish-English Codeswitching in the Caribbean and the US*. Edited by Rosa E. Guzzardo Tamargo, M. Carmen Parafita Couto and Catherine M. Mazak. Amsterdam: John Benjamins Publishing Company, pp. 37–60. [CrossRef]

Crystal, David. 2006. *Language and the Internet*, 2nd ed. Cambridge: Cambridge University Press.

Danet, Brenda. 2010. Computer-mediated English. In *The Routledge Companion to English Language Studies*. Edited by Janet Maybin and Joan Swann. London and New York: Routledge, pp. 146–56.

Davies, Bronwyn, and Rom Harré. 1990. Positioning: The Discursive Production of Selves. *Journal for the Theory of Social Behaviour* 20: 43–63. [CrossRef]

Dörnyei, Zoltán. 2007. *Research Methods in Applied Linguistics. Quantitative, Qualitative, and Mixed Methodologies*. Oxford: Oxford University Press.

Dynel, Marta. 2017. Participation as audience design. In *Pragmatics of Social Media*. Edited by Wolfram Bublitz and Christian Hoffman. Berlin and Boston: De Gruyter Mouton, pp. 61–82.

Eckert, Penelope. 2009. Communities of practice. In *Concise Encyclopaedia of Pragmatics*, 2nd ed. Edited by Jacob L. Mey. Oxford: Elsevier, pp. 109–12.

Gumperz, John J. 1977. The sociolinguistic significance of conversational code-switching. *RELC Jourrnal* 8: 1–34. [CrossRef]

Gumperz, John J. 1982. *Discourse Strategies*. Cambridge: Cambridge University Press.

Hernández, Nuria. 2016. Bilingual encounters online: Common linguistic code in German English WhatsApp. In *Anglistentag 2015 Paderborn*. Edited by Christoph Ehland, Ilka Mindt and Merle Tönnies. Trier: Wissenschaftlicher Verlag, pp. 253–66.

Herring, Susan C. 2007. A faceted classification scheme for computer-mediated discourse. *Language@ Internet* 4: 1–37.

Johansson, Marjut. 2017. Youtube. In *Pragmatics of Social Media*. Edited by Christian R. Hoffmann and Wolfram Bublitz. Berlin and Boston: De Gruyter Mouton, pp. 173–200. [CrossRef]

Joseph, John E. 2010. Identity. In *Language and Identities*. Edited by Carmen Llamas and Dominic Watt. Edinburgh: Edinburgh University Press, pp. 9–17.

Kaplan, Andreas M., and Michael Haenlein. 2010. Users of the world, unite! The challenges and opportunities of Social Media. *Business Horizons* 53: 59–68. [CrossRef]

Koch, Peter, and Wulf Oesterreicher. 1985. Sprache der Nähe—Sprache der Distanz. Mündlichkeit und Schriftlichkeit im Spannungsfeld von Sprachtheorie und Sprachgeschichte. *Romanistisches Jahrbuch* 36: 15–43. [CrossRef]

Koch, Peter, and Wulf Oesterreicher. 2012. Language of Immediacy—Language of Distance: Orality and Literacy from the Perspective of Language Theory and Linguistic History. In *Communicative Spaces. Variation, Contact, and Change: Papers in Honour of Ursula Schaefer*. Edited by Claudia Lange, Beatrix Weber and Göran Wolf. Frankfurt am Main: Peter Lang, pp. 441–73.

Levy, Mike. 2007. Culture, culture learning and new technologies: Towards a pedagogical framework. *Language Learning & Technology* 11: 104–27.

Locher, Miriam, and Richard J. Watts. 2005. Politeness Theory and Relational Work. *Journal of Politeness Research* 1: 9–33. [CrossRef]

Muysken, Pieter. 2007. Mixed Codes. In *Handbook of Multilingualism and Multilingual Communication*. Edited by Peter Auer and Li Wei. Berlin: Mouton de Gruyter, pp. 315–38.

Myers-Scotton, Carol. 2006. *Multiple Voices. An Introduction to Bilingualism*. Malden: Blackwell.

Page, Ruth, David Barton, Johann W. Unger, and Michele Zappavigna. 2014. *Researching Language and Social Media: A Student Guide*. London and New York: Routledge.

Pérez-Sabater, Carmen. 2022. *Mixing Catalan, English and Spanish on WhatsApp: A Case Study on Language Choice and Code-Switching*. Amsterdam: John Benjamins.

Spitzmüller, Jürgen. 2006. Spricht da jemand? Repräsentation und Konzeption in virtuellen Räumen. In *Aktualnije Problemi Germanistiki i Romanistiki*. Edited by Galina Kramorenko. Smolensk: SGPU, pp. 33–56.

Tanskanen, Sanna-Kaisa. 2007. Metapragmatic utterances in computer-mediated interaction. In *Metapragmatics in Use*. Edited by Wolfram Bublitz and Axel Hübler. Amsterdam and Philadelphia: John Benjamins, pp. 87–106.

Wentker, Michael. 2018. Code-Switching and Identity Construction in WhatsApp. Evidence from a (Digital) Community of Practice. In *The Discursive Construction of Identities in Online and Offline Contexts: Personal-Group-Collective*. Edited by Birte Bös, Sonja Kleinke, Nuria Hernández and Sandra Mollin. Amsterdam: John Benjamins, pp. 109–31. [CrossRef]