

**“BRANDS: THEY’RE JUST LIKE US!” NON-STANDARD LANGUAGE USE
AND THE CONSTRUCTION OF BRAND IMAGE IN DUOLINGO’S SOCIAL
MEDIA MARKETING.**

Jay Barber

English Language, 2023

Abstract: This paper discusses the use of non-standard language and features of computer-mediated communication (CMC) in digital marketing, and how this relates to the construction of brand image. To do so, a small-scale investigation of the brand Duolingo’s language use on Twitter and Facebook is conducted. The study seeks to identify which non-standard linguistic features are used and highlight the factors that influence Duolingo’s decision to implement such features in their digital marketing. Qualitative and quantitative analyses are conducted to answer these questions. The non-standard linguistic features identified in Duolingo’s marketing are the use of minimalist typography, non-standard lexis, non-standard punctuation and the personification of Duolingo through the use of first-person singular pronouns. It is found that the implementation of non-standard language is affected by the choice of social media platform and post content, with posts on Twitter that relate less to Duolingo’s products having a higher rate of non-standard features. It is argued that the use of non-standard language constructs a “relatable” brand image through the personification of brand mascot Duo. Linguistic features are used to attract a youthful audience and imply a close consumer-brand relationship.

Keywords: sociolinguistics, CMC, social media, marketing, non-standard language, brand image, linguistic personification

Supervisor(s): Dr. Heike Pichler

Table of Contents

1. Introduction	2
2. Literature Review	3
2.1. Computer-Mediated Communication	3
2.2. Non-standard Language in Marketing	6
2.3. Language Use, Performance and Brand Image.....	9
2.4. Research Questions and Hypotheses	10
3. Methodology	12
3.1. Choice of Brand and Social Media Platforms.....	12
3.1.1. Duolingo	12
3.2. Twitter and Facebook.....	12
3.2.1. Rationale	14
3.3. Data Collection	14
3.4. Method for Analysis.....	15
3.5. Limitations	16
4. Results	16
4.1. Qualitative Results	17
4.2. Quantitative Results	18
5. Discussion	22
5.1. Use of Non-Standard Linguistic Features.....	23
5.2. Use of Language – Twitter vs. Facebook.....	27
5.3. Non-Standard Language and ‘Relatable’ Brand Image	30
6. Conclusion	33
References	35

“B r a nTh e y Jurs et Us i!” Non-standard language use and the construction of brand image in Duolingo’s social media marketing.

1. Introduction

Marketing is a near-inescapable force in the average person’s life, as consumers encounter around 5000 overt advertisements each day (Vasile, et al., 2021: 31). With 58.4% of the global population now active social media users (Digital 2023: Global Overview Report), it is no surprise that brands turn to social media platforms to increase their visibility amongst the masses. Despite the prevalence of marketing in our everyday lives, linguistic analysis of the subject has been relatively sparse, especially in relation to digital marketing.

With that in mind, this study focuses on the use and impact of non-standard linguistic features in brands’ social media marketing. Specifically, I analyse the brand Duolingo’s language use on Twitter and Facebook to provide a discussion of how non-standard linguistic features are implemented to construct a “relatable” brand image (Sligh and Abidin, 2020). At present, this subject has primarily been discussed in the field of digital studies, with relatively sparse linguistic analysis. Due to the relatively niche nature of the marketing style I discuss, I refer readers to Roth-Gordon et al., (2020), Smirnov (2020), and Sligh and Abidin (2022), all of which are discussed in my literature review, for further examples.

As brands increasingly occupy spaces on social media platforms designed for human interaction and mimic the modes through which individuals communicate with one another, I argue that linguistic attention must be paid to how brands use language to “transform traditional linear, unidirectional marketing models” (Chen, et al., 2015: 97). All instances of communication “necessarily involve meaningful social relationships as prerequisite, conduit and outcome” (Blommaert, 2018: 67) – whereas this has primarily been discussed in relation to communication between individuals, I argue that the shifting power balance and increased potential for relationships between brands and consumers necessitates this discussion. In doing so, I aim to highlight the means through which language can be used to further blur the lines between brands and consumers.

The present study conducts a small-scale linguistic analysis of the brand Duolingo’s posts on Twitter and Facebook. Using both qualitative and quantitative methods, I investigate which non-standard linguistic features are implemented in digital marketing, suggest factors that influence Duolingo’s decision to use non-standard language, and provide further discussion on how non-standard language is used to construct brand image. Primarily, my discussion

focuses on non-standard lexical items, orthographic features, and the personification of brands through the use of first-person singular pronouns.

This paper introduces relevant literature from the fields of linguistics, marketing, and digital studies in **section 2**, before presenting the research questions and hypotheses central to my study. In **section 3**, I outline how the study was conducted, addressing data collection, quantitative and qualitative methods of analysis, and expanding upon potential limitations to my research. **Section 4** then displays the qualitative and quantitative results. **Section 5** expands on the results in relation to the research questions presented and provides further discussion of how a ‘relatable’ brand image is constructed through the use of non-standard language.

2. Literature Review

As stated, this paper introduces a discussion regarding brands’ use of non-standard language in digital marketing and how this relates to the construction of brand image. In this section, I introduce relevant literature to facilitate this discussion. First, I introduce key linguistic features of computer-mediated communication (CMC), attitudes towards CMC in comparison to Standard Language, and suggest factors that influence the use of CMC. I then introduce prior studies that explore non-standard language use in marketing, discussing the features used, motivations identified for brands to advertise this way, and potential audience responses. Finally, I discuss concepts related to the linguistic construction of identity and brand image. This section then concludes with the introduction of my research questions and hypotheses.

2.1. Computer-Mediated Communication

Since much of the present study is focused on the ‘features of CMC’ and other instances of ‘non-standard language’, it is necessary to define these terms. Prior studies have amassed a large range of terminology to describe language used online – netspeak, chatspeak, cyber slang, and more (Paterno, et al., 2022: 1980). My decision to refer to the ‘features of CMC’ is intended to simplify this. By the ‘features of CMC’, I specifically refer to the linguistic characteristics of online discourse that is primarily casual in nature, such as communicating with friends through short-messaging services (SMS), as opposed to formal communications conducted online, such as work-related emails.

As features of CMC are instances of non-standard language, I use both terms throughout the paper. This therefore necessitates a discussion of what is meant by ‘standard’. Standard English (or any ‘Standard’ language, for that matter) is a register of language that is “promoted by institutions of such widespread hegemony that they are not ordinarily recognised as [a] distinct [register]” (Agha, 2005: 24). Standard variants dominate both written and verbal forms of communication, promoted by institutions such as the education system and official media (Blommaert, 2013: 9; Gates and Ilbury, 2019). These varieties are ideologically loaded and seen as in opposition to non-standard variants, which are often deemed lesser by comparison. When applied to language online, features of CMC are criticised for not ascribing to rules of ‘proper’ punctuation, spelling and grammar (Squires, 2010: 476).

With these definitions in mind, I will now introduce the features of CMC central to this paper. In written CMC, non-lexical information such as tone and volume must be constructed through linguistic innovations and an understanding of these conventions between users. This is referred to as digital orality, which “manifests interpersonal, informal spoken conversation in the written sphere” (Soffer, 2010: 396). Ruhl (2016) provides a case study of how orthographic features are used to encode pragmatic information on the blogging platform Tumblr. Ruhl (2016) states that a flat tone of voice and implied informality is constructed “through the non-standard convention of all lowercase [typing]” (Ruhl, 2016: 12). Referred to as “minimalist typography” (McCulloch, 2019: 147), this feature of CMC has gained meaning as the technology that CMC is conducted through has developed. Whereas typing in all-lowercase was previously described as “sav[ing] a keystroke” on a computer, therefore making capitalisation the marked form (Crystal, 2001: 87), the development of mobile phones with automatic capitalisation has altered this. The continued use of minimalist typography is indicative of the feature carrying social meaning that users deem worth the effort.

The non-standard use of punctuation in CMC can also construct alternative meanings than when used traditionally, such as the use of a question mark to simply pose a question. Punctuation can provide emphasis, deliver speech acts, or indicate physical actions (Yasa et al., 2023: 106 – 107; Yus, 2011: 173). Question marks and exclamation points can indicate confusion and excitement respectively, with meanings intensified through reduplication (Ruhl, 2016: 8). In this way, a message simply reading “???” can be meaningful, indicating confusion that would be evident visually in face-to-face conversation, or even implying a

confused noise. As well as constructing tone, punctuation can be used to demonstrate physical actions in written form, as asterisks can be used to indicate movement (i.e., *high five*). Full-stops can also be a marked feature in CMC, as the lack of punctuation observed in minimalist typography has led some to perceive sentence-final full-stops as blunt or insincere (McCulloch, 2019: 113). Finally, punctuation is sometimes used in a ‘decorative’ sense – for example, asterisks and tildes have been used to surround text ~*~* like this *~*~ as a stylistic choice. This feature can evoke multiple meanings, ranging from providing emphasis, constructing sarcasm, or evoking ‘cuteness’ (McCulloch, 2019: 138). Interestingly, the development of emoji within CMC has led to the use of punctuation in this way to be replaced with emoji in some cases, leading the older variant to be perceived as ‘archaic’ by some.

Other features associated with CMC are lexical substitutions (*b4* = ‘before’), onomatopoeic signs (*haha* = laughter) and abbreviations (*BRB* = be right back) (Soffer, 2012). Although thought of by many as integral features of CMC, the actual use of these lexical features is often minimal, occurring only 142 times out of a corpus of 10,000 words in a study by Squires (2010: 483). The importance of lexical items in constructing CMC is perhaps due to the stark difference they create between CMC and Standard Language. The use of these items is associated with younger generations, who are often found to be the primary users of CMC (Schwartz et al., 2013: 9). However, as time passes since the introduction of social media platforms and online messaging services, it has been found that millennials are increasingly using these non-standard lexical items (Paterno, et al., 2022).

As previously stated, CMC is often viewed in opposition to Standard Language, which leads to public metalinguistic discourse assessing its ‘appropriateness’ in certain contexts. Squires (2010: 460) refers to this as the “imperative of containment”, the idea that features of CMC must be confined to certain domains. This is often framed through the narrative that features of CMC must not encroach on ones’ ability to write formally. Although the imperative of containment was first discussed in relation to the use of CMC in ‘offline’ settings, I expand the use of this term to facilitate a discussion of where CMC is deemed as appropriate within the digital landscape.

As well as perceived appropriateness, a user’s familiarity with technology as a whole may influence the use of CMC. Prensky (2001) coined the terms “digital native” and “digital immigrant” to refer to those who grew up immersed in digital culture, and those who did not.

In what could now be considered an archaic description, Prensky states that the “accent” of a digital immigrant may involve things such as “printing out [their] e-mail” (Prensky, 2001: 3) to edit it. As over two decades have passed since the publication of Prensky’s paper, it can be argued that these categories are now reductive. However, for the sake of this paper, I utilise these terms to refer to those who are ‘fluent’ in the linguistic conventions of CMC, and those who are not. This may still correlate with age, as younger generations have higher rates of social media use (Politte-Corn, et al. 2023) and therefore greater exposure to the use of CMC.

2.2. Non-standard Language in Marketing

As stated in section 1, social media marketing alters the hierarchical structure inherent in earlier forms of marketing by taking place on platforms that allow brands and consumers to interact (Chen, et al., 2015), as opposed to the one-sided nature of print and television advertisements. Social media replies and reshares can be seen as the new word of mouth, leading some brands to engage with widely reshared trends or “memes”, defined as “trending online topic[s] initiated by individuals or organised groups” (Vasile, et al., 2021: 31).

The use of CMC in digital marketing was first popularised by American fast-food chain ‘Denny’s’, in 2015. The brand’s social media presence (primarily on Twitter and Tumblr) has been described as “akin to a hallucinogenic experience” (Smirnov, 2017: 66), utilising features of CMC such as minimalist typography to attract a “younger market that enjoys surrealist humour” (Smirnov, 2017: 66).



Figure 1. Smirnov, 2017.

The publicity that Denny’s gained through this marketing strategy led other companies to follow in their footsteps. Video streaming service Netflix was found to enact a “fannish persona” on Instagram accounts made to promote their original series (Sligh and Abidin, 2022). To do so, Netflix mimicked the digital style of a fan through using emotive language and first-person singular pronouns, “position[ing] its voice as an individual rather than a promotional team or an embodiment of the Netflix brand” (Sligh and Abidin, 2022: 12). Although this study was not linguistic in nature, it highlighted this aspect of language use as contributing to the construction of ‘relatability’, as brands seek to “foster feelings of intimacy and build rapport” with their audience (Sligh and Abidin, 2022: 12). By positioning a brand as an individual and peer to their audience, parasocial relationships that “create an illusion of intimacy at a distance” can be formed (Yuksel and Labrecque, 2016: 306).

Another example of non-standard language use by brands is Roth-Gordon et al.’s (2020) study of brands’ use of African American Vernacular English (AAVE) on Twitter. They argued that the “disembodied voicing” of social media works to construct “corporate cool”, where corporations use AAVE and references to Black culture while remaining racially ambiguous (Roth-Gordon, et al., 2020: 116). The use of AAVE for profit is an example of linguistic appropriation and was found by Roth-Gordon et al., to demonstrate a potential “lack of fluency” (Roth-Gordon et al., 2020: 117) where the lexis of AAVE was used incorrectly. A particularly intriguing intersection is present in the discussion of brands using AAVE and features of CMC, as many lexical items originating in AAVE are popularised online and later attributed to “Gen Z” or, more generally, “online slang” (Voice Magazine, 2022). Both non-Black social media users and brands therefore “reap the indexical qualities” (Ilbury, 2022: 259) of AAVE without recognising important cultural context.



Figure 2. Roth-Gordon et al., (2020)

The use of AAVE by brands in social media marketing relates to its status as an enregistered variant, meaning that it is a “socially recognised register of forms” (Agha, 2003: 231) that is ideologically linked to a certain type of person (see section 2.3.). Similarly to AAVE, enregistered regional dialects have been used in marketing. King and Wicks (2009) provide an analysis of the use of Newfoundland English in a TV commercial that featured an actor’s approximation of the dialect with subtitles ‘translating’ to Standard English underneath. Instead of using this variant to appear ‘cool’ or construct ‘relatability’, King and Wicks (2009) argue that the commercial used linguistic features to ‘other’ and ridicule speakers of Newfoundland English for the entertainment of speakers of more standard variants.

It is evident that brands’ use of enregistered variants that are associated with minority groups or those who face accent discrimination poses a greater risk of causing harm or offence than the use of CMC features. This is because CMC is not tied to any geographic location or specific social groups, other than its association with younger generations (Soffer, 2012: 1104). However, the use (or misuse) of CMC is not without risk of criticism. When a brand causes offence through language or is simply perceived as ‘inauthentic’, the ‘Silence, Brand!’ meme has been created as a humorous response from an unimpressed audience (Totempool, 2021).



Figure 3. *Totempool, 2021.*

Interestingly, both Smirnov (2017) and Roth-Gordon et al. (2020) reference Twitter as a social media platform on which brands use non-standard language in marketing. Though limited by the lack of linguistic research into this phenomenon, this may still suggest that the choice of social media platform influences a brands’ decision to implement non-standard linguistic features in their marketing. Referring back to Squires’ (2010) concept of the imperative of containment, this may suggest that Twitter is viewed as a more ‘casual’

platform where the use of non-standard language is acceptable. When compared to other social media platforms such as Facebook (to be discussed), Twitter is said to have a “context collapse” where multiple audiences convene (Marwick and boyd, 2011). Rather than viewing this only from the direction of users to audiences, it can be looked at from the perspective of users following a wider range of accounts. In this case, the use of non-standard language could be due to brands’ position amidst a wider range of accounts on Twitter when compared to platforms such as Facebook, which primarily connect users with acquaintances they know ‘in real life’. Equally, the use of non-standard language to construct relatability or humour may be motivated by the potential to go ‘viral’ through the resharing of tweets to a wide audience, especially as many Twitter users do not use privacy settings to limit their accounts’ viewers (Miller and Melton, 2015).

Additionally, the studies presented all provide examples in which brands use the first-person singular pronoun alongside linguistic features such as non-standard lexis and minimalist typography, though only Sligh and Abidin (2022) comment upon this. The use of first-person singular by brands is not typical within marketing, and prior studies have suggested that first-person singular pronouns are negatively received by consumers when not used in direct conversation (Labrecque, et al., 2014). Typically, brands use the first-person plural (“we”) alongside second person (“you”) to refer to the brand as a whole and directly communicate with a consumer (Sela et al., 2012). As stated by Sligh and Abidin (2022), the use of first-person singular seemingly intends to personify the brand as an individual to foster relatability. This therefore makes the use of first-person singular an interesting feature to analyse.

2.3. Language Use, Performance and Brand Image

The studies by Roth-Gordon et al., (2020), King and Wicks (2009) and Sligh and Abidin (2022) all examine how brands use language that is, to some extent, enregistered. These varieties are associated with a certain type of person or social characteristic, meaning that the use of such variants can index these characteristics. The intentional use of enregistered varieties is referred to as “stylistation”, through which individuals “construct personae and perform stance” (Ilbury, 2022: 485).

In a digital landscape, language can be used to index certain social groups or characteristics that a user does not, in reality, belong to or have. Ilbury (2020; 2022) provides in-depth linguistic analyses of how individuals stylistically perform identities, assessing how white, gay men on Twitter use AAVE to construct what he labels the ‘sassy queen’ persona (2020)

and how the persona of white, British, working-class women is indexed through the ‘Hun’ meme (2022). In these studies, Ilbury examines which linguistic features are used to construct these personae. Across both studies, orthographic variation is used to replicate phonological features, lexical items associated with the indexed group are used, and thematic content seeks to “evoke the heteroglossic connotations of the ‘voice’ of the imagined author” (Ilbury, 2022: 495),

For both brands and individuals who construct personae through language, it is essential to recognise this as an act of performance. I utilise Goffman’s dramaturgical model (1990) as a framework through which to discuss this within this paper. The dramaturgical model states that presentation of the self is constructed through intentional ‘performances’ that can be altered based upon factors such as the imagined audience (Marwick and boyd, 2011). Rather than discuss the construction of identity, I use the dramaturgical model to discuss how brand image is created through a performance using non-standard language. Brand image, defined as “how customers perceive a brand and its attributes” (Ekhlassi et al., 2018: 5), avoids the potentially problematic associations of ‘identity’, which has been criticised as contributing to essentialism through suggesting that those who share social characteristics are a homogenous group (Bucholtz and Hall, 2005: 373).

Additionally, as brand image relies on consumers’ perceptions, I focus on the use of non-standard language to construct ‘relatability’ and imply a close consumer-brand relationship, as identified in Sligh and Abidin’s study of Netflix (2022). Although an actual study of consumer perception is beyond the scope of this paper, linguistic theory supports my discussion of brand image. As authenticity, or “enoughness” is “assessed by audiences” rather than being viewed as an inherent quality possessed by some speakers (Androutsopoulos, 2015: 76), the language used by brands online reveals not only the persona constructed, but indicates who the imagined audience may be. Language use can be seen as both constructing and reflecting the intended audience, summarised best by Agha as “reformulating one’s persona [involves the use of] interactional tropes that seek strategically to align the performed image of speaker with that of target audiences” (2003: 268).

2.4. Research Questions and Hypotheses

In summary, this section has introduced key concepts related to brands’ use of non-standard language in digital marketing. I have illustrated the key features of CMC and suggested that its use is associated with younger “digital natives” (Prensky, 2001) and linked to acceptability

through the imperative of containment (Squires, 2010). By introducing prior studies of non-standard language use in digital marketing, I have shown how language constructs brand image and can create parasocial relationships with consumers. The construction of brand image has been discussed in relation to Goffman's dramaturgical model (1990).

I now relate this general discussion to my own investigation, which focuses on the brand Duolingo's language use on Twitter and Facebook (expanded on in section 3). Based upon the prior research, the following research questions and hypotheses are posed:

RQ1: Are certain non-standard linguistic features used in Duolingo's digital marketing more frequently than others?

H1: I hypothesise that there will be a lower rate of non-standard lexis (i.e., use of AAVE, acronyms, etc.) due to the potential harm this can cause (Roth-Gordon et al., 2020) and the fact certain forms aren't used frequently in CMC (Squires, 2010).

RQ2: Is non-standard language used more by Duolingo on Twitter than on Facebook?

H2: I hypothesise that non-standard language will be used more on Twitter, based on the platform's prevalence in prior studies, functionality, and the "context collapse" (Marwick and boyd, 2011) of audiences.

RQ3: Does the age demographic of a social media platform affect the rate at which Duolingo uses features of CMC?

H3: I hypothesise that a younger age demographic will lead to increased use of non-standard features associated with CMC, as it is suggested that younger users are "digital natives" who are more likely to be fluent in the implied meanings of CMC's features (Prensky, 2001).

RQ4: Does the content of a post influence the rate at which non-standard language is used?

H4: I hypothesise that posts that are intended to entertain, rather than overtly advertise Duolingo's products, will include a higher rate of non-standard features than posts due to the imperative of containment (Squires, 2010) and CMC's casual nature relating to digital orality (Soffer, 2010).

After answering these research questions, I will introduce a further qualitative discussion focusing on defining the brand image constructed by Duolingo's use of non-standard language.

3. Methodology

In this section, I outline the study I conducted to answer the research questions presented above. To do so successfully, background information on the brand Duolingo and social media platforms Twitter and Facebook is provided, alongside the reasoning for these choices. I then provide overview of the data collection and method for analysis.

3.1. Choice of Brand and Social Media Platforms

3.1.1. Duolingo

Founded in 2011, Duolingo is an American educational technology company, best known for their language learning app of the same name. The flagship app boasts over 500 million learners across over 100 language courses and the company has recently introducing two other apps, Duolingo ABC and Duolingo Maths. For clarity, 'Duolingo' will be used to refer to the company as a whole, and references to the app itself will be specified. Duolingo reports that, in the US, around 60% of their learners are under 30 years old, demonstrating a younger overall userbase (Duolingo Blog, 2022).

In Duolingo's published brand guidelines, the brand personality is described as the following: inspiring, inclusive, can-do, curious, and quirky. Most significant here is the use of "quirky" – as a brand, Duolingo aims to "put the fun in language fundamentals" (Duolingo Brand Guidelines, n.d.), demonstrating a clear standard they have for their brand image. A key component of this is the mascot, Duo (a green cartoon owl) who is prominent across all marketing.

3.2. Twitter and Facebook

Launched in 2006, Twitter is a micro-blogging social media platform through which users 'follow' and are 'followed by' other users. Posts on Twitter are statuses up to 280 characters long known as 'tweets'. If an account is set to 'public', these can be 'retweeted', meaning an original post is reshared by another user onto the 'timeline' of those following them, or 'quote tweeted', in which the original tweet is shared with additional commentary. Businesses on Twitter are distinguished from individuals' accounts by a gold verification checkmark, but

otherwise interact with Twitter users in the same way as an any individual account. It is necessary to clarify that this is the status of Twitter as of the time of writing (early 2023), but this may be subject to change as Elon Musk’s acquisition of Twitter has led to increasingly frequent changes in the platform’s functionality.

Twitter reports 436 million monthly active users as of early 2022 (Statista, 2023). In April 2021, it was reported that the majority of Twitter users (38.5%) were between 25 and 34 years old. Users between 13 and 24 years old made up 23.7% of the demographic, and those aged 35+ made up 37.7%.

Facebook was founded in 2004 and expanded its reach to users worldwide in 2006. Users of Facebook create a profile comprised of their personal information (such as name, date of birth, education, etc.) through which they can share text posts and multi-media content. Facebook users interact with one another through becoming ‘friends’ on the site, which allows users to see each other’s posts. It is also possible for individuals’ accounts to be public, making posts visible to everyone. A brands’ presence on Facebook is functionally different than on Twitter as, on Facebook, brands have pages that can be ‘liked’ by users, rather than profiles that can be friended.

Facebook reports 2.91 billion monthly active users in 2021, making it the most used social media worldwide. As of January 2022, the majority of Facebook users (29.9%) were between the age of 25 and 34 (Statistica, 2023).

Table 1. Age Demographics of Twitter and Facebook

Age Group	Twitter	Facebook
13 – 24	23.7%	26%
25 – 34	38.5%	30%
35+	37.8%	44%

Table 1’s age groupings have been reduced significantly to enable comparison between platforms as Twitter’s demographic data had fewer age categories than Facebook.

Additionally, the grouping of 35+ was decided based on the majority of Duolingo’s users being under 30. Facebook statistics have been rounded to nearest whole number as original data totalled 100.1%.

3.2.1. Rationale

As this study aims to assess the use of non-standard language in constructing brand image, it followed that it would be more informative to select a brand that is, to some extent, using non-standard language in this way, as opposed to seeking out the prevalence of this trend overall. This therefore led me to select Duolingo as the brand at the centre of this analysis. Duolingo has received praise for their marketing in recent years, being awarded ‘Social Marketer of the Year’ in 2021 by Ad Age (Pittsburgh Business Times, 2022). Their alternative marketing strategy originated on the video sharing platform TikTok, on which they began sharing humorous videos featuring mascot ‘Duo’, references to pop culture, and the features of CMC discussed prior in this paper. As a result, their following on TikTok grew from 50,000 followers in September 2021 to over 6 million in January 2023. Duolingo was therefore chosen for this analysis because their success is both relevant and recent, demonstrating a repeated use of nonstandard language to construct brand image.

Although Duolingo’s marketing success originated on TikTok, the present study turns its attention to Twitter and Facebook. This decision was made for multiple reasons. On a practical level, Facebook and Twitter are both primarily text-based platforms, allowing for more effective data collection when compared to the audio-visual emphasis of TikTok. Alongside the practical reasoning, Twitter and Facebook were selected due to their nature as more long-standing social medias. TikTok was founded in 2016 and has a much younger demographic, with 38.9% of users under the age of 25 (Statistica, 2023). It is therefore of interest to examine if the use of nonstandard language in marketing is extended to ‘older’ platforms, both in longevity and user base. Further, Twitter and Facebook allow a compelling cross-platform comparison due to their similar demographics and different functionality.

3.3. Data Collection

The data collected for this analysis was sourced from the official Duolingo accounts on Twitter and Facebook (@duolingo on both). Although Duolingo has multiple verified accounts on Twitter, such as @duolingoUK and @duolingoBrasil, data was collected only from the original account. This was decided because @duolingo is the original Twitter account and most likely to be managed alongside the Facebook page, whereas the country-specific accounts are much smaller in following and less centralised.

A total of 130 posts were collected from Twitter and Facebook (65 per social media platform) with the final 5 individual posts each month from January 2022 to January 2023 collected.

Selecting the final 5 posts ensured that the sample was not biased towards posts using non-standard language. Additionally, the timeframe was selected to align with Duolingo's success on TikTok, allowing time for the social media strategy implemented on TikTok to potentially be introduced across Twitter and Facebook.

The process of data-collection specified that 'individual' posts were selected for analysis. This was decided as Twitter allows for replies to other users' posts and I found that most of Duolingo's individual responses were related to customer service inquiries, rather than marketing. Any tweets in which Duolingo responded to its own account were excluded purely due to limitations in Twitter's search function. Finally, during data collection, a series of tweets related to Duolingo's 'Streak of the Week' competition were excluded due to their repetitive nature and the fact they would have dominated the sample of posts from Twitter.

3.4. Method for Analysis

To adequately answer the research questions central to this study, it was necessary to conduct a qualitative analysis that, due to the scope of this project, relied on the judgement calls of only one person. Due to this, I turn to Ilbury (2022: 490) as basis for a statement of positionality: as a 21-year-old who spends much time immersed in digital culture, I argue that I have the lived experience that allows me to recognise linguistic features used in marketing that aim to target those with a similar investment in digital culture. A combination of my lived experience as a digital native (Prensky, 2001) and previous literature as explored in section 2 therefore provides the basis for my analysis.

After initial data collection, posts were annotated with their linguistic features and any relevant context (e.g. images in posts) before being categorised by their content into 'themes' to facilitate further analysis (Ilbury, 2022: 489). Thematic analysis was conducted in stages – first, posts were annotated with multiple themes that I deemed relevant to the message, before these themes were collated into broader categories. The judgement of themes was decided independently, as stated prior, but conducted methodically, with conciseness and continuity in mind.

After the initial thematic analysis, posts were further sorted into one of three categories: conventional, unconventional, or mixed. These categories are what I label the posts' 'style', categorised mainly by the linguistic features central to the post. Unconventional style indicates the use of non-standard linguistic features and conventional style the use of Standard English. The mixed style was introduced to cover certain nuances in the data – for

example, instances where a post displayed a mix of features, or where thematic content had to be taken into consideration. Due to the scope of this paper, the mixed style has often had to be grouped with the unconventional style to simplify data analysis.

The labels ‘conventional’ and ‘unconventional’ were chosen to avoid the connotations of the term ‘traditional’. Conventional marketing, for Duolingo, still presents a cheerful and comedic tone. By comparison, the label ‘unconventional’ is used to highlight posts that diverge from Standard English or include content unrelated to Duolingo’s language-learning functions. Once data was codified by theme and style, I was able to conduct a small-scale quantitative analysis to demonstrate the distribution of themes and styles across social media platforms, as well as examine the frequency of non-standard features identified.

3.5. Limitations

Prior to my results, I first wish to address the limitations of this study. Most obvious is the sample size of my data, both in terms of posts collected and the fact only one brand is analysed. This limitation is due to the scope of this project and results should therefore be viewed as preliminary findings that are not intended to be generalised.

Further, I recognise that there is an element of overlap in my method for analysis. I have taken efforts to avoid this where possible and make apparent when it has taken place. Specifically, there some instances where posts’ style is influenced by thematic category, and where a non-standard linguistic feature is present in an otherwise conventional post. Despite this, I argue that the overall patterns within my results remain accurate and I use the forthcoming discussion to expand on these instances. In future studies, I would recommend a much larger data sample to allow for further categories within the analysis, as this would have resulted in an ineffective analysis of the present data.

4. Results

In this section, I outline the results of the study conducted. As my quantitative analysis relied on the results from my qualitative analysis, I first present the thematic categories identified across Duolingo’s social media posts before outlining the key linguistic features that contributed to the categorisation of a posts’ style. I then present the results of my quantitative analysis, which focuses on highlighting the distribution of post style, thematic content, and non-standard linguistic features.

4.1. Qualitative Results

Table 2 presents the key thematic categories that emerged from my qualitative analysis of Duolingo’s posts on both Twitter and Facebook. Each theme has been given a name, described briefly, and an example from Twitter has been provided. In total, six specific themes were identified, with the addition of a ‘miscellaneous’ category.

There is a clear divide in the themes presented throughout Duolingo’s social media presence. The categories of ‘app related’, ‘company image’ and ‘educational’ all align closely with Duolingo’s existence as an educational company. Through these themes, Duolingo promotes the app itself, the company as an equitable employer, and displays itself as knowledgeable through the sharing of informative articles or promoting causes related to social justice. Comparatively, the categories of ‘media references’ and ‘brand personification’ often made no mention of Duolingo’s products. Instead, posts centred on media and pop culture references (such as *The Owl House*, the upcoming *Barbie* movie, and Elon Musk’s acquisition of Twitter) or the brand’s mascot, Duo, who was implied to be the author of many of these posts.

Table 2. *Thematic categories prominent in Duolingo’s posts on Twitter and Facebook, with examples from Twitter.*







Theme	Description	Example:
App Related	Relating to Duolingo’s apps, e.g., new language courses.	"Say bonjour to our newest course: Haitian Creole! HT Learn the revolutionary language of the first Black independent nation that is spoken by over 12 million people worldwide. Vin jwenn ak nou (that means ""join us"") and start learning today!    https://bit.ly/34lfh9T "
Company Image	Promotion of Duolingo as a company, e.g., work ethic and communication with consumers through ‘fan art’.	"We  our interns! Duolingo's 2022 intern cohort is our largest yet, and this #NationalInternDay, we're celebrating all of the incredible work they've done this summer. Even Lily's proud!  Read more about our (award-winning!) internship experience: https://bit.ly/3baLc04 "
Educational	Educational material relating to language learning, cultural issues and social justice.	"Hi Mayor @FrancisSuarez! You asked, ""how can I help?"" Well... if you can keep a Haitian Creole learning streak for 30 days, we'll donate \$30,000 to @HaitiansInTech. Deal?"
Media Reference	Focusing on other media and culture references, e.g., TV shows, celebrities.	“three tickets to barbie please !”
Brand Personification	Posts that construct the identity of brand mascot ‘Duo’.	an owl has gotta pay the bills somehow
Miscellaneous	Posts that do not fit into other categories.	Happy Owlidays! Wishing everyone a restful holiday season (not too restful though, don't forget your lesson!) 

Table 3 illustrates the linguistic features associated with conventional and unconventional styles of posts. As stated, posts categorised as mixed style may display a combination of these linguistic features or have their stylistic category influenced by thematic content. For clarity, I note that linguistic features associated with unconventional style will be referred to as ‘non-standard linguistic features’. Although the use of first-person singular pronouns is not non-standard in relation to individual communication, I assign it this label due to the negative associations previously discussed regarding its use in marketing.

Table 3. *Key linguistic features associated with ‘conventional’ and ‘unconventional’ style.*

Feature	Conventional Style	Unconventional Style
Orthography	Standard English orthography, including punctuation and capitalisation.	Minimalist typography – all lowercase, lack of punctuation.
‘Decorative’ Features	Tone indicated through word choice and standard use of punctuation. Use of emojis relate to post contents.	Non-standard use of punctuation to construct tone.
Pronouns	Use of 1 st person plural (we, us, our) to refer to Duolingo as a corporation.	Use of 1 st person singular (I, me, mine) to personify ‘Duo’ as authoring posts.
Lexical Features	Standardised spellings of lexical items and use of ‘regular’ vocabulary, as opposed to dialect features.	Use of lexical features of CMC e.g., acronyms. Use of lexical items tied to certain vernaculars.

4.2. Quantitative Results

Table 4 presents the distribution of posts by their style across Twitter and Facebook. Of the 65 Facebook posts collected, 86.4% were classified as conventional, indicating primary use of Standard English, with the remaining posts categorised as unconventional (9.2%) or mixed (6.2%). This indicates a strong preference against the use of non-standard linguistic features on Facebook. Comparatively, 41.5% of posts on Twitter were classified as conventional, less than half of that on Facebook. On Twitter, 50.8% of posts were unconventional, indicating frequent use of non-standard features, and the remaining 7.7% as mixed. Twitter has 4 mixed posts with 0 non-standard features while Facebook has 2 mixed posts with no non-standard features.

Additionally, table 4 reveals that there is a greater diversity in the styles used on Twitter compared to Facebook. Though both social media platforms feature all styles of post, there is greater variation on Twitter than on Facebook. Although the use of Standard English remains the most prominent variant across both platforms, the use of non-standard features remains significant.

Table 4. *Distribution of posts by ‘style’ on Twitter and Facebook (N and %).*

Style	Twitter		Facebook	
	N	%	N	%
Conventional	27	41.5	55	84.6
Unconventional	33	50.8	6	9.2
Mixed	5	7.7	4	6.2

Table 5 illustrates the frequency of posts by theme across Twitter and Facebook. On Facebook, the majority of posts belong to the categories ‘app related’ (40%) and ‘educational’ (30.8%), amounting to 70.8% in total. As previously stated, the frequency of these themes indicates that Duolingo’s communication on Facebook primarily relates to the function of their apps. Comparatively, posts on Twitter are much more varied. The most frequent theme is ‘brand personification’ (24.6%), though this is followed closely by ‘app related’ (23.1%), ‘educational’ (20%) and ‘media references’ (20%).

Table 5. *Distribution of posts by ‘thematic category’ on Twitter and Facebook (N and %).*

Theme	Twitter		Facebook	
	N	%	N	%
App	15	23.1	26	40
Company Image	6	9.2	9	13.8
Educational	13	20	20	30.8
Media References	13	20	2	3.1
Brand Personification	16	24.6	2	3.1
Miscellaneous	2	3.1	6	9.2

Figure 6 displays the categorisation of posts categorised as ‘unconventional’ or ‘mixed’ on Twitter and Facebook over the course of the 13 months that data was collected from. The decision to group these styles together was made to enable a simpler comparison, as ‘mixed’ posts were judged as relating more closely to unconventional posts. For clarity, the first time period (January 2022 to May 2022) includes an extra month and therefore has a larger sample size in comparison to the latter two periods.

The data displayed here indicates a steady increase in the proportion of posts with non-standard features on Twitter, rising from 32% (January – May 2022) to 75% (June – September 2022) and finally 80% (October 2022 – January 2023). Comparatively, Facebook

had only 4% and 5% of posts in non-conventional styles in the first two time periods. Interestingly, between October 2022 and January 2023, this increased significantly to 40%.

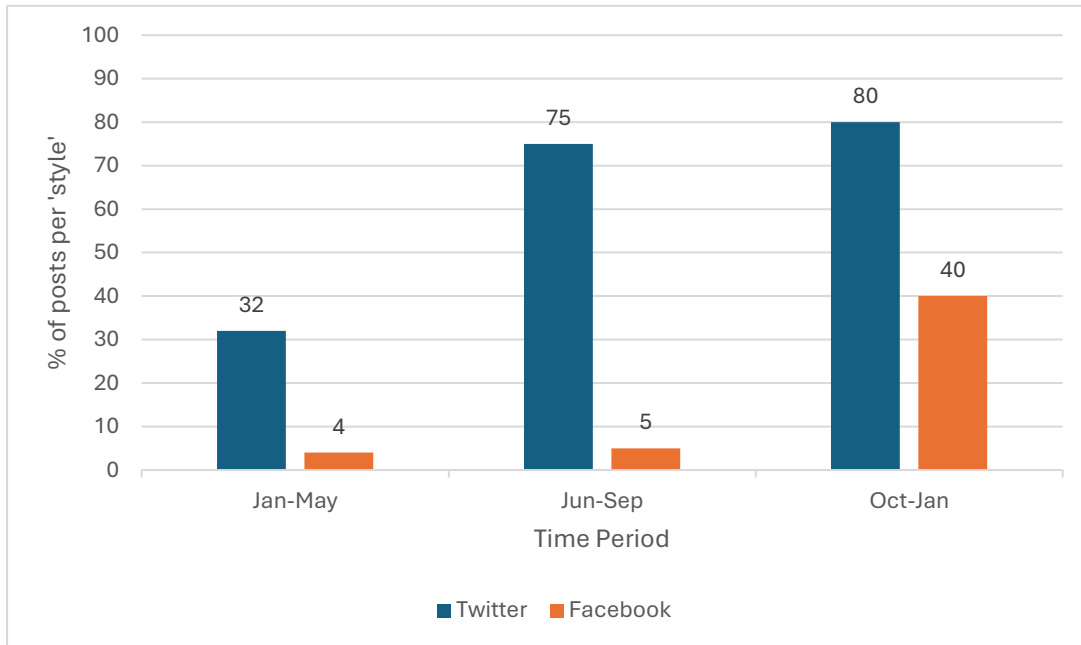


Figure 4. *Percentage of posts in unconventional or mixed style on Twitter and Facebook, across time periods.*

Figure 7 illustrates the link between post style and thematic content, showing which percentage of each theme was categorised as conventional or unconventional / mixed. For simplicity, figure 7 combines the data from Facebook and Twitter into one – table 5 should be kept in mind for the raw number of tokens.

This illustrates that the themes most associated with Duolingo’s function as an educational company are much more likely to be written in conventional style, while the posts relating to media references and brand personification are exclusively categorised as unconventional or mixed. As style was primarily decided based upon linguistic features, this is indicative of certain linguistic features being associated with certain topics discussed – put plainly, topics that are more ‘formal’ are presented in Standard English.

As stated throughout this paper, certain overlaps between style and theme were unavoidable. However, I argue that this data remains accurate in representing the larger pattern, despite some instances of overlap.

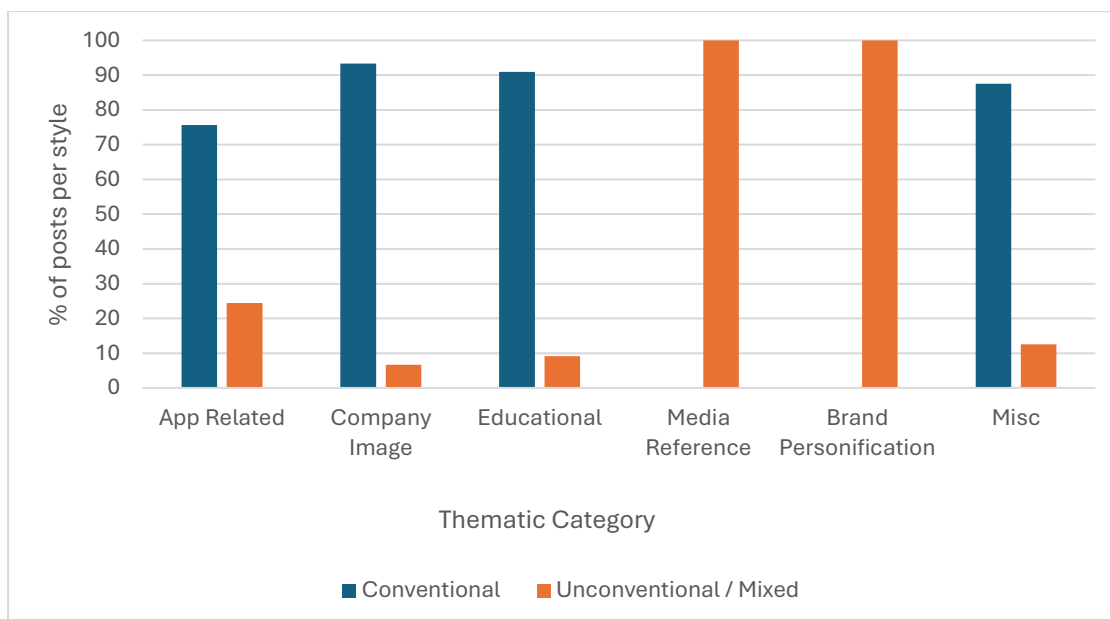


Figure 5. *Percentage of posts categorised as ‘conventional’ or ‘mixed / unconventional’ per thematic category.*

Finally, table 6 illustrates the frequency of nonstandard linguistic features across Duolingo’s posts on Twitter and Facebook. The four categories of features is primarily based on table 3 – however, the category of ‘pronouns’ has been expanded further to what I have labelled ‘implied authorship’ to encompass all posts that suggested they were authored by brand mascot ‘Duo’. This includes posts that explicitly feature first-person singular pronouns as well as posts that imply authorship through other means, such as through context in images or the use of statements that only make sense when read as from an individual author, as opposed to a company statement. This is expanded upon fully in section 5.

This table includes all instances of nonstandard linguistic features regardless of post style. Interestingly, three Facebook posts and two Twitter posts of conventional style included a non-standard feature – these posts were classed as conventional because these non-standard features were not judged to outweigh the rest of the post’s style or content.

This table demonstrates that on Twitter, the features of minimalist typography and implied authorship are much more common than non-standard lexis and decorative / tone-indicating features. Comparatively, on Facebook, non-standard features are used at a much more equal level, perhaps due to the much smaller rate of use in comparison to Twitter. From this, it appears as though Duolingo perceive certain features as more acceptable to use, or perhaps

easier to implement. It should equally be noted that on Twitter, it was common for unconventional posts to use multiple non-standard features at the same time.

As this table only aims to display the number of tokens per feature, a breakdown of the instances of these features is available in section 5.

Table 6. *Number of tokens of each non-standard linguistic feature.*

Feature	Twitter (N)	Facebook (N)
Minimalist Typography	28	3
Implied Authorship	27	5
Non-standard Lexis	13	3
Non-Standard Punctuation	8	3

5. Discussion

In this section, I discuss my results in relation to the research questions and hypotheses posed in section 2.4. I expand on my results to provide further explanations of the patterns observed and suggest potential reasons for these patterns. Once the research questions have been addressed, I introduce a qualitative discussion of the brand image Duolingo constructs through the use of non-standard language.

To briefly summarise section 4, my qualitative analysis identified the thematic categories present throughout Duolingo’s social media and defined the linguistic features associated with what I labelled ‘conventional’ and ‘unconventional’ post styles. The following quantitative analysis highlighted the links between social media platform, post style and thematic content. Broadly, Duolingo’s activity on Facebook primarily used Standard Language to discuss thematic content relating to Duolingo’s existence as an educational company. Comparatively, Twitter displayed greater variation, with a slight preference towards non-standard language use and discussion of thematic content unrelated to Duolingo’s functionality.

Prior to this discussion, I once again reiterate the goals and limitations of this paper. Broadly, I have aimed to provide a linguistic analysis of the non-standard language used in digital marketing, focusing on the use of features of CMC and construction of brand image. As well as the well-discussed limitation of sample size, significant elements of my results and the subsequent discussion have relied on my own judgement. Though confident in my

judgements, I encourage future research to continue this discussion beyond the scope of my paper.

5.1. Use of Non-Standard Linguistic Features

This research has frequently relied on post style as an indicator of non-standard language use. However, I aimed to provide a more specific analysis of the non-standard linguistic features used, leading to the introduction of the following research question:

***R Q 1** Are certain non-standard linguistic features used more frequently than others in Duolingo's social media posts?*

In this section, I will address the hypothesis posed from this question, before providing a more detailed discussion of all non-standard linguistic features used. Based on prior literature demonstrating the potential of negative reactions to the use of enregistered dialects in marketing (Roth-Gordon et al., 2020; King and Wicks, 2009), I hypothesised that there would be a lower rate of non-standard lexical items, especially those tied to specific dialects. Table 6 presented the rate at which non-standard linguistic features were used and confirmed my overall hypothesis, as there were only 16 instances of non-standard lexis used across the 130 posts collected. However, upon a closer analysis of the non-standard lexis used, I argue that my reasoning for this hypothesis may not be entirely accurate, especially in relation to the use of enregistered features.

Table 7 shows the non-standard lexical features found in my data. The rate of abbreviations and contractions (8 in total) aligns with Squires' finding that similar lexical shortenings were used infrequently in text conversations (Squires, 2010: 483). However, the use of four separate phrases originating in AAVE sheds doubt on the hypothesis that lexis from enregistered dialects would be avoided entirely due to past controversy surrounding linguistic appropriation. Although this rate of use is still low, it is higher than I expected – however, I pose a possible explanation for this.

As discussed previously, the appropriation of lexical items often leads to the loss of the original cultural context (see Laing, 2021 for a full discussion of linguistic appropriation). This has been tied to the misrepresentation of some features of AAVE as 'Gen Z Slang'. Equally, in Ilbury's study of AAVE used by white gay men on Twitter, it was found that lexical features were used more frequently than morphosyntactic features such as copula absence, invariant be, etc. (2020: 253 – 254). If the features of AAVE present in Duolingo's

social media posts have been indexed by non-Black social media users, their use by Duolingo could represent a lack of knowledge of the phrases' cultural origins. If Duolingo were aiming to construct "corporate cool" (Roth-Gordon et al., 2022) or the "sassy queen" persona (Ilbury, 2020), one would expect to find use of phonological or morphosyntactic features, which were not present in my data.

Table 7. *Non-standard lexical items in the data collected.*





Feature	Tokens	Post
Contractions	<gonna> <i>going to</i>	"who's gonna tell them...? http://duolingo.com/abc "
	<gotta> <i>got to</i>	an owl has gotta pay the bills somehow
	<y'all> <i>you all</i>	see y'all on tumblr
	<'em> <i>them</i>	"My Children Plushies are for sale! Your wish is our command. Check ' em out here: http://store.duolingo.com #Duocon
AAVE	<serving tea> <i>spreading gossip, originating in Black drag culture</i>	Vanity Feather is serving all the latest Duolingo tea.  Comment with  if you'd read this issue.
	<turnt> <i>intoxicated</i>	turnt tuesday
	<it's giving> <i>something is emitting a particular vibe, originating in Black Twitter</i>	It's giving me postmodern Duo
	<glow up> <i>improvement, often physical, originating from rapper Chief Keef</i>	"Duolingo ABC, our children's literacy app, just got a major glow-up! ✨ The app teaches kids (ages 3-8) how to read in English with bite-sized lessons and stories featuring Duo and his friends.   https://duolingo.com/abc "
Abbreviations	<s2g> <i>swear to God</i>	if i don't make Big Bird's twitter circle i s2g
	<rn> <i>right now</i>	my villain origin story really being put on blast rn
	<plz> <i>please</i>	"ah yes, we'll take a side of brains plz. #duolingo365"
	<sesh> <i>session</i>	"who wants to join me for my weekly cry sesh? will provide snacks. #duolingo365"
Internet Specific Phrases	<slide into dms> <i>send someone a direct message, often flirtatious</i>	when you're trying to slide into Khaleesi's DMs
	<tag yourself> <i>identify what you relate to, often in an image</i>	tag yourself , i'm "man in green bird costume"

Table 6 revealed that the non-standard use of punctuation was the most infrequent of the features identified within my data, with a total of 11 instances across both Twitter and Facebook. Of these instances, three involved the use of asterisks to provide emphasis in posts

otherwise categorised as conventional due to the primary use of Standard English. This may suggest that this specific use of punctuation is becoming more widely known as a means of transmitting tone. Comparatively, a post uploaded to both Twitter and Facebook used asterisks to indicate a physical action, separating it from posts constructing digital orality through written statements (Yasa et al., 2023: 106). This demonstrates Duolingo's fluency in the multiple metapragmatic meanings that can be construed through punctuation, used purposefully and stylistically based upon context. On Twitter, there were two instances of hyphens being used to 'strike trough' text. Though this feature was extremely rare and not identified by previous literature, it provides an interesting point of discussion for future research when digital orality is considered (Soffer, 2010). The strikethroughs seemingly represent a backtracking or self-correction of sorts, as one would restart an utterance in natural speech. In written text, this can be seen as humorous and as a blurring of the lines between written and 'spoken' communications.

My results found that that the most common non-standard linguistic features were the use of minimalist typography (31 instances) and implied authorship (32 instances). The preference of these linguistic features indicates their significance in the construction of Duolingo's brand image, which will be discussed fully in section 5.3. I suggest that the high rate of overlap that minimalist typography had with other non-standard features indicates that its use is a key indicator of the unconventional style, as 29 posts using minimalist typography featured another non-standard feature.

The frequent use of implied authorship is evidence of an increasing personification of Duolingo's brand image, particularly on Twitter. As stated in section 4, implied authorship was extended to cover posts that did not directly use the first-person singular. Of the 32 instances, 15 directly used "I" or "me", while the remaining posts were categorised as such based on relevant context, often provided in images attached to the post, reoccurring 'inside jokes' on Duolingo's social media pertaining to mascot Duo, or through specific phrases used that otherwise implied individual authorship. To illustrate, I provide the following examples.

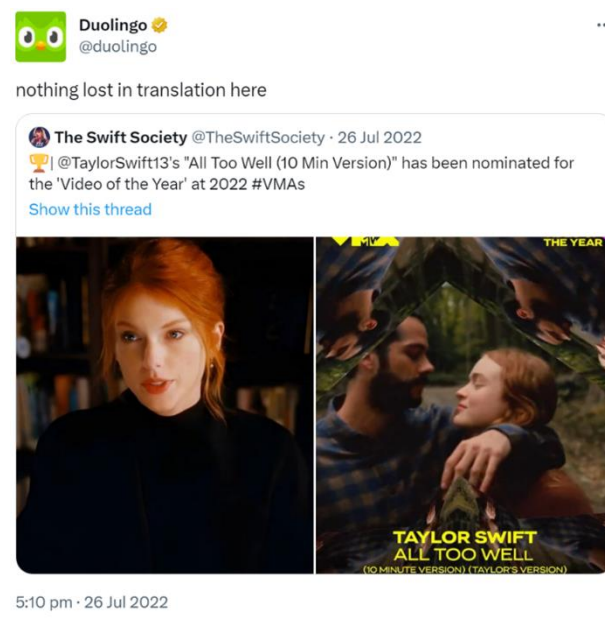
Example 1 demonstrates implied authorship through the joining of Duo and the audience, referring to the migration of users from Twitter to Tumblr after Elon Musk purchased the platform. The phrase "see y'all" is authored as though from an individual viewpoint and mimics much of the discourse that could be found on Twitter during that time.

Example 1. “see y’all on tumblr”



Comparatively, example 2 demonstrates Duolingo’s construction of a “fannish persona” (Sligh and Abidin, 2022), illustrated through the voice of brand mascot Duo. The reference both implies support of popular artist Taylor Swift through referencing the song’s lyrics, and ties it to the brand’s functionality as a language-learning app. Though the implication of a singular voice is not as evident here, I argue that this constructs an important part of Duolingo’s brand image (explored further in section 5.3).

Example 2. “nothing lost in translation here”



Finally, example 3 shows a quote tweet in which Duolingo replies to another personified brand (*Cult of the Lamb*, a video game). As the original post from @CultoftheLamb uses first-person singular and addresses Duo, it follows that this tweet should be viewed as authored from an individual.

Example 3. “*man it’s 3 PM on a Wednesday*”



5.2. Use of Language – Twitter vs. Facebook

Alongside assessing which non-standard features are used by Duolingo, I have investigated the factors that influence Duolingo’s decision to implement non-standard forms. The remaining three quantitative research questions address potential constraints, the first of which is:

R Q 2Is non-standard language used more by Duolingo on Twitter than on Facebook?

From RQ2, I hypothesised that Twitter would have a higher rate of non-standard language use than Facebook. This hypothesis was based partially on Twitter’s prevalence in prior studies investigating non-standard language use online (Smirnov 2017; Roth-Gordon et al., 2022; Ilbury, 2020). More significantly, however, this hypothesis was rooted in Marwick and boyd’s (2011) discussion of Twitter’s context collapse, in which users follow and are followed by a greater range of people than would be present on social media platforms such as Facebook. My hypothesis was confirmed, as Twitter had a much higher rate of non-standard linguistic features (table 7). When categorising posts by style, it was found that 50.5% of tweets were

classed as unconventional, compared to only 9.2% on Facebook. However, future studies may find the difference between Twitter and Facebook is less significant, as figure 6 illustrated that Facebook had a sharp increase in non-conventional posts in the final months from which data was collected.

Alongside addressing social media platform as a potential constraint on language use, I aimed to investigate the effect of age demographics:

R Q 3 Does the age demographic of a social media platform affect the rate at which Duolingo uses features of CMC?

From RQ3, I hypothesised that a higher rate of non-standard features would correlate with a younger age demographic. This was based on prior research associating CMC features with younger audiences (Presnky, 2001; Squires, 2010; Smirnov, 2017). However, I don't believe that my results and the available data on age demographics presented in table 1 allow for a confirmation of my hypothesis as the majority of users on both Twitter and Facebook are between 25 and 34 years old, yet there is a significantly higher rate of non-standard linguistic features present on Twitter. Although Facebook has a slightly higher percentage of users above the age of 35 than Twitter, at 44% and 37.8% respectively, this difference isn't significant enough to make any claims regarding actual audience age to language use, especially as most of Duolingo's app users are under 30.

Despite not finding a direct link between platform age demographic and non-standard language use, I present two important caveats to this conclusion. First, which I will expand upon in my discussion of brand image, is the important difference between actual audience and imagined audience (Marwick and boyd, 2011). As discussed, language can be used to foster 'relatability' and create parasocial relationships (Sligh and Abidin, 2022; Yuksel and Labrecque, 2016). This performance relies as much, if not more, on the audience a person or brand wishes to target, as it does the actual audience. Secondly, it is imperative to note that the available data on age demographic is only quantitative in nature and does not reveal information relating to individuals' qualitative experiences of social media use, ranging from time spent on specific apps to user enjoyment. Instead, it only reveals the ages associated with accounts on the platform. Future studies should therefore consider gathering improved qualitative information to gain a fuller picture of social media habits by age group and reassess if age demographic correlates with language use.

The final research question relating to constraints on language use was:

R Q 4 Does the content of a post influence the rate at which non-standard language is used?

I hypothesised that posts that primarily intend to entertain the audience, as opposed to those that more explicitly advertise Duolingo's products, would have a higher rate of non-standard features. This hypothesis was formed on the basis that CMC is associated with casual conversation, and so digital orality is more likely to be constructed when post content is intended to entertain (Soffer, 2010). Equally, as CMC is associated with the "imperative of containment" and seen as especially inappropriate in formal writing (Squires, 2010), I theorised that the use of non-standard language in formal posts would reflect negatively on Duolingo's brand image.

To answer this, I utilised the categorisation of posts by style and theme as indicators of a post's linguistic features and key content. Figure 7 illustrates the link between post theme and style. As stated previously, the thematic categories of 'app-related', 'company-related' and 'educational' posts were largely written in conventional style, whereas 'media references' and 'brand personification' posts were exclusively written in unconventional or mixed style. The clear differences between these two groupings of thematic categories can be further simplified – the former categories (app-related, company-related and educational) can be seen as 'informative', whereas the latter (media references and Duo's characterisation) are for the most part instances of 'phatic' communication (Blommaert, 2018: 61). My hypothesis is therefore confirmed that message content, and the intention of such messages, influences the rate at which non-standard language is used. Specifically, as predicted, informative posts that can be seen as more formal are affected by the "imperative of containment" (Squires, 2010) and so non-standard language is seen as inappropriate, whereas phatic communication is seen as acceptable for non-standard forms.

Importantly, however, the correlation between post content and linguistic features must not be viewed independently of other factors. As shown in table 5, the distribution of thematic content is varied across Twitter and Facebook, with the categories relating to phatic communication appearing more frequently on Twitter. The constraints of social media platform and post content are therefore not able to be viewed as orthogonal. Future research should seek to untangle the relationship between post content and social media platform.

5.3. Non-Standard Language and ‘Relatable’ Brand Image

From the results and discussion presented thus far, it has been shown that there is a clear division in Duolingo’s language use that is influenced by both social media platform and thematic content. It is now necessary to introduce a qualitative discussion that assesses not only the patterns of language use, but the motivations behind these choices. Utilising the answers to my research questions, I return to the concept of brand image to argue that the choice of thematic content and non-standard linguistic features are implemented to construct a “relatable” brand image, as found in Sligh and Abidi’s analysis of Netflix’s Instagram activity (2020).

Although this paper lacks the scope to conduct a full analysis and discussion of the brand image Duolingo constructs through their conventional posts, I return to the company’s statement discussed in section 3.1., in which it was stated that Duolingo aims to “put the fun in language fundamentals.” As this is achieved through the use of Standard Language, it follows that a separate – if similar – brand image is constructed through the use of non-standard language. From my analysis and discussion so far, I conclude that there are three key approaches used by Duolingo to present a brand image that is not just fun, but seeks to construct relatability:

1. Intentional use of Twitter as a platform for non-standard language to create potential for parasocial relationships to form.
2. Use of implied authorship to position ‘Duo’ as an individual representing Duolingo, the company, and as similar to the imagined audience.
3. Use of minimalist typography to construct informality and replicate conversational styles of CMC.

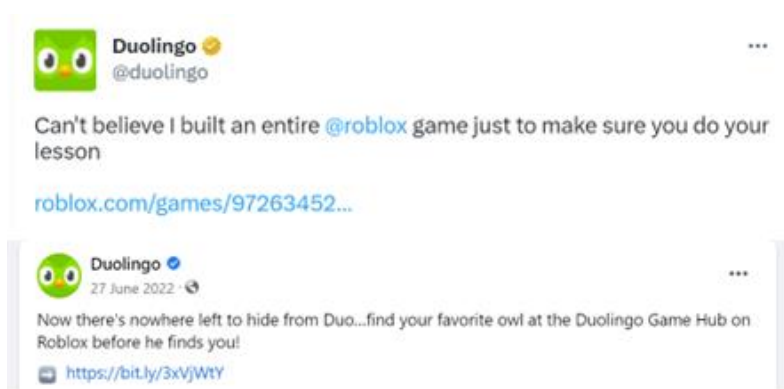
I address each of these in turn to build my argument. First, the use of Twitter as opposed to Facebook as the primary platform for non-standard linguistic features and thematic content relating to phatic communication is, in my view, intentional. Returning to the dramaturgical model (Goffman, 1990), I argue that Twitter and Facebook are separate ‘stages’, in which Duolingo produces separate performances. It has been argued that Twitter is a more impersonal platform than Facebook, as users interact with followers rather than friends (Miller and Melton, 2015: 680). This creates an interesting dichotomy – despite Facebook being perceived as more personal, Duolingo’s brand image on the platform is much less ‘personable’ than on Twitter. I argue that this is because of the positioning of brands differs

between the platforms, as explained in section 3.1. Users on Facebook ‘like’ Duolingo’s page while they ‘friend’ individual accounts (often those they know in real life), whereas users on Twitter ‘follow’ Duolingo as they would all accounts.

Amongst the context collapse of Twitter (Marwick and boyd, 2011), Duolingo can therefore capitalise on their position as one of many accounts a Twitter user follows, whereas their status as a brand on Facebook may be more apparent. This is furthered through the functionality of Twitter allowing Duolingo to interact with consumers more easily, and publicly (through retweets), than on Facebook. Even if Duolingo does not, in reality, interact with followers more frequently on Twitter, the implication that they could fosters the potential for parasocial relationships to be by their consumers.

The use of implied authorship on Twitter furthers this construction of relatability. Most obviously, the use of first-person singular pronouns explicitly constructs ‘Duo’ as an author of the posts. The most direct evidence of this in relation to Twitter and Facebook’s existence as separate ‘stages’ is displayed in example 3. Despite discussing the same content (the promotion of Duolingo’s collaboration with online game, ‘Roblox’) the tweet uses first-person singular whereas the Facebook post does not.

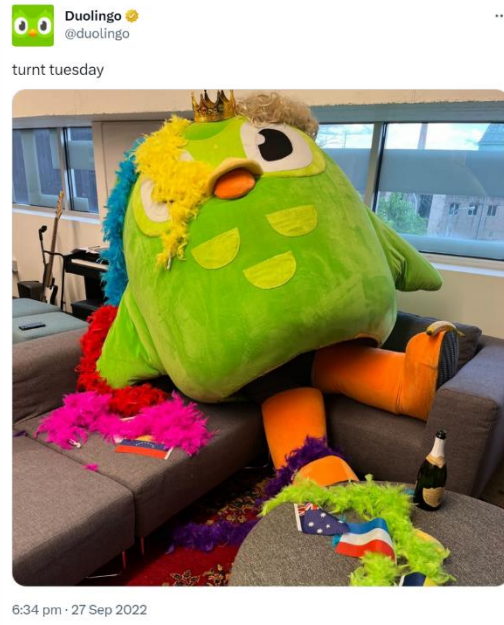
Example 4. *Twitter (above) and Facebook (below) posts relating to Duolingo’s collaboration with Roblox.¹*



By positioning Duo as an author, users therefore interact with a perceived individual – although consumers will undoubtedly understand that there is no individual author behind these posts, at least in the way that one is implied, the construction of such can still foster feelings of intimacy (Sligh and Abidi, 2022). Implied authorship is also used to index the constructed social life of Duo, as shown in examples 5, through presenting Duo as drunk

through the use of the non-standard lexical item, ‘turnt’. Equally, example 6 constructs Duo as ‘feeling’ sassy.

Example 5. “*turnt tuesday*”



Example 6. “*feeling spicy but what’s new*”



This is extremely similar to the practices identified in Sligh and Abidi’s study of Netflix, where it was found that the brand-run account “post[ed] about its (non-existent) mother and dating life” (Sligh and Abidi, 2022: 12). Equally, it relates to the thematic analysis conducted

in Ilbury's study of the 'Hun' meme, which indexed social practices and tropes that were "deeply ingrained in observations of British classed, gendered, and ethnic social dynamics" (Ilbury, 2022: 498). Despite my analysis not finding a link between Twitter's user age demographics and the increased rate of non-standard language use, I suggest that the combination of implied authorship and the indexed characteristics and interests presents Duo as relatable to a younger audience. Specifically, Duo is personified through the attribution of personal characteristics, ranging from his media interests, social life, and more. Again, this relates to Netflix's construction of a fannish persona – however, I believe Duolingo furthers this technique, as media references are unrelated to the company's products (instead discussing Taylor Swift, cartoons, movies, etc.), whereas Netflix enacts a fannish persona to promote their own shows (Sligh and Abidin, 2022).

Finally, these messages are mostly delivered in the stylised nature of CMC through the use of minimalist typography and, occasionally, other non-standard linguistic features identified in this paper. Though more subtle than the use of enregistered dialects, the implementation of CMC's stylistic features serves to further the construction of a relatable brand image. As a reminder, minimalist typography is associated with an informal register and flat tone of voice (Ruhl, 2016; McCulloch, 2019). Most importantly, minimalist typography is associated with casual conversations in CMC. If social media users associate minimalist typography with messages one would receive from a friend, the use of this by a brand reduces the formality associated with traditional forms of marketing and Duolingo purposefully implies a close relationship with their audience. Further, these messages may no longer be viewed as advertisements – when written to construct the voice of an imagined author (Ilbury, 2022: 495) through implied authorship and orthographic choices, the brand can gain promotion from its intended audience by purposefully mimicking their linguistic styles and associated interests (Agha, 2003). In relation to Goffman (1990), a performance is orchestrated to foster relatability through the character of Duo, while Duolingo's social media team is, in a sense, hidden from the audience 'back stage'.

6. Conclusion

Through an analysis of Duolingo's posts on Twitter and Facebook, this study has shown how non-standard language is used in the company's digital marketing. In summary, Duolingo constructs relatability through the considered use of non-standard linguistic features to imply a close consumer-brand relationship. This is furthered through references to elements of

popular culture associated with a younger audience, and through the personification of brand mascot 'Duo' as the voice through which these messages are delivered. Importantly, it has been revealed that the use of non-standard language is still constrained by post content and social media platform, indicating that Standard Language remains the 'acceptable' format through which formal topics should be communicated.

As stated, the majority of Duolingo's social media posts during the period of data collection were written in conventional style. However, throughout this time period, the proportion of non-conventional posts on both Twitter and Facebook increased, suggesting that the use of non-standard language is increasing in popularity within Duolingo's marketing strategy. At present, this paper contributes an in-depth case study of a phenomenon in marketing that relies primarily on linguistic features. In doing so, I have highlighted the importance of language in the construction of brand image.

Future studies should examine language use across a larger range of brands and social media platforms. Equally, to assess if the brand image companies aim to construct through their language use is successful, future studies should consider gaining qualitative data regarding consumer perceptions.

References

- Agha, A. (2003) 'The social life of cultural value', *Language & Communication*, 23(3–4), pp. 231–273. Available at: [https://doi.org/10.1016/S0271-5309\(03\)00012-0](https://doi.org/10.1016/S0271-5309(03)00012-0).
- Agha, A. (2005) 'Registers of Language', in *A Companion to Linguistic Anthropology*. John Wiley & Sons, Ltd, pp. 23–45. Available at: <https://doi.org/10.1002/9780470996522.ch2>.
- Androutsopoulos, J. (2015) 'Negotiating authenticities in mediatized times', *Discourse, Context & Media*, 8, pp. 74–77. Available at: <https://doi.org/10.1016/j.dcm.2015.06.003>.
- Blommaert, J. (2013) 'Writing as a sociolinguistic object', *Journal of Sociolinguistics*, 17(4), pp. 440–459. Available at: <https://doi.org/10.1111/josl.12042>.
- Blommaert, J. (2018). *Durkheim and the Internet: on sociolinguistics and the sociological imagination*. London: Bloomsbury Academic.
- Brand Narrative - Duolingo Brand Guidelines* (no date). Available at: <https://design.duolingo.com/writing/brand-narrative#brand-personality> (Accessed: 21 April 2023).
- Bucholtz, M. and Hall, K. (2005) 'Language and Identity', in *A Companion to Linguistic Anthropology*. John Wiley & Sons, Ltd, pp. 369–394. Available at: <https://doi.org/10.1002/9780470996522.ch16>.
- Chen, K.-J. *et al.* (2015) 'Would You Be My Friend? An Examination of Global Marketers' Brand Personification Strategies in Social Media', *Journal of Interactive Advertising*, 15(2), pp. 97–110. Available at: <https://doi.org/10.1080/15252019.2015.1079508>.
- Crystal, D. (2001) *Language and the Internet*. 2nd ed.. Cambridge: Cambridge University Press.
- Digital 2023: Global Overview Report — DataReportal – Global Digital Insights* (no date). Available at: <https://datareportal.com/reports/digital-2023-global-overview-report> (Accessed: 21 April 2023).
- Duolingo's viral TikTok account earns it 'Social Marketer of the Year' award from Ad Age* (2021) *Pittsburgh Business Times*. Available at: <https://www.bizjournals.com/pittsburgh/news/2022/04/26/duolingo-wins-ad-age-award-tiktok-account.html> (Accessed: 24 April 2023).
- Ekhlassi, A., Niknejhad Moghadam, M. and Adibi, A.M. (2018) *Building Brand Identity in the Age of Social Media: Emerging Research and Opportunities*. IGI Global (Advances in Marketing, Customer Relationship Management, and E-Services).
- Gates, S.M. and Ilbury, C. (2019) 'Standard Language Ideology and the Non-Standard Adolescent Speaker', in C. Wright, L. Harvey, and J. Simpson (eds) *Voices and Practices in Applied Linguistics*. White Rose University Press (Diversifying a Discipline), pp. 109–126. Available at: <https://www.jstor.org/stable/j.ctvs09r4p.9>
- Gen Z 'slang': A blatant appropriation of AAVE* (2022) *Voice Magazine*. Available at: <https://www.voicemag.uk/blog/10958/gen-z-slang-a-blatant-appropriation-of-aaave> (Accessed: 28 April 2023).
- Goffman, E. (1990) *The presentation of self in everyday life*. Harmondsworth, England: Penguin.
- How does language learning differ across generations?* (2022) *Duolingo Blog*. Available at: <https://blog.duolingo.com/dear-duolingo-how-does-language-learning-differ-between-generations/> (Accessed: 21 April 2023).
- Ilbury, C. (2020) "'Sassy Queens": Stylistic orthographic variation in Twitter and the enregisterment of AAVE', *Journal of Sociolinguistics*, 24(2), pp. 245–264. Available at: <https://doi.org/10.1111/josl.12366>.

- Ilbury, C. (2022) 'U Ok Hun?: The digital commodification of white woman style', *Journal of Sociolinguistics*, 26(4), pp. 483–504. Available at: <https://doi.org/10.1111/josl.12563>.
- King, R. and Wicks, J. (2009) "'Aren't We Proud of Our Language?": Authenticity, Commodification, and the Nissan Bonavista Television Commercial', *Journal of English Linguistics*, 37(3), pp. 262–283. Available at: <https://doi.org/10.1177/0075424209339924>.
- Laing, Rachel Elizabeth, "Who Said It First?: Linguistic Appropriation Of Slang Terms Within The Popular Lexicon" (2021). Theses and Dissertations. 1385. <https://ir.library.illinoisstate.edu/etd/1385>
- Marwick, A.E. and boyd, d. (2011) 'I tweet honestly, I tweet passionately: Twitter users, context collapse, and the imagined audience', *New Media & Society*, 13(1), pp. 114–133. Available at: <https://doi.org/10.1177/1461444810365313>.
- Mcculloch, G. (2019). *Because Internet: Understanding the New Rules of Language*. New York: Riverhead Books.
- Miller, R. and Melton, J. (2015) 'College students and risk-taking behaviour on Twitter versus Facebook', *Behaviour & Information Technology*, 34(7), pp. 678–684. Available at: <https://doi.org/10.1080/0144929X.2014.1003325>.
- Pateno, G.M., Cotejo, D.G. and Bongcales, M. (2022) 'The Growing Popularity Of Netspeak Among Millennials: An Analysis Of The Emerging Lingo On Facebook Messenger', *Webology*, 19(3), pp. 1978–1990.
- Politte-Corn, M., Nick, E.A. and Kujawa, A. (2023) 'Age-related differences in social media use, online social support, and depressive symptoms in adolescents and emerging adults', *Child and Adolescent Mental Health*, n/a(n/a). Available at: <https://doi.org/10.1111/camh.12640>.
- Prensky, M. (2001) 'Digital Natives, Digital Immigrants Part 1', *On the Horizon*, 9(5), pp. 1–6. Available at: <https://doi.org/10.1108/10748120110424816>.
- Roth-Gordon, J., Harris, J. and Zamora, S. (2020) 'Producing white comfort through "corporate cool": Linguistic appropriation, social media, and @BrandsSayingBae', *International Journal of the Sociology of Language*, 2020(265), pp. 107–128. Available at: <https://doi.org/10.1515/ijsl-2020-2105>.
- Ruhl, M. (2016) 'Welcome to my Twisted Thesis: An Analysis of Orthographic Conventions on Tumblr'. Available at: <https://doi.org/10.13140/RG.2.2.15033.77929>.
- Schwartz HA, Eichstaedt JC, Kern ML, Dziurzynski L, Ramones SM, et al. (2013) Personality, Gender, and Age in the Language of Social Media: The OpenVocabulary Approach. *PLoS ONE* 8(9): e73791. doi:10.1371/journal.pone.0073791
- Sela, A., Wheeler, S.C. and Sarial-Abi, G. (2012) 'We Are Not the Same as You and I: Causal Effects of Minor Language Variations on Consumers' Attitudes toward Brands', *Journal of Consumer Research*, 39(3), pp. 644–661. Available at: <https://doi.org/10.1086/664972>.
- Sligh, C. and Abidin, C. (2022) 'When Brands Become Stans: Netflix, Originals, and Enacting a Fannish Persona on Instagram', *Television & New Media*, p. 15274764221134778. Available at: <https://doi.org/10.1177/15274764221134778>.
- Smirnov, Kristen. "Marketing on Tumblr." *Strategic Uses of Social Media for Improved Customer Retention*, edited by Wafaa Al-Rabayah, et al., IGI Global, 2017, pp. 59-83. <https://doi.org/10.4018/978-1-5225-1686-6.ch004>
- Soffer, O. (2010) "'Silent Orality": Toward a Conceptualization of the Digital Oral Features in CMC and SMS Texts', *Communication Theory*, 20(4), pp. 387–404. Available at: <https://doi.org/10.1111/j.1468-2885.2010.01368.x>.
- Soffer, O. (2012) 'Liquid language? On the personalization of discourse in the digital era', *New Media & Society*, 14(7), pp. 1092–1110. Available at: <https://doi.org/10.1177/1461444812439550>.

Squires, L. (2010) 'Enregistering internet language', *Language in Society*, 39(4), pp. 457–492. Available at: <https://doi.org/10.1017/S0047404510000412>.

Statista (2023). Topic: TikTok. [online] Statista. Available at: <https://www.statista.com/topics/6077/tiktok/#topicOverview> (Accessed: March 4 2023)

Statista (2023). Topic: Twitter – Statistics and Facts. [online]. Available at: <https://www.statista.com/topics/737/twitter> (Accessed: 21 April 2023)

Statista (2023). Topic: Facebook. [online] Available at: <https://www.statista.com/topics/751/facebook> (Accessed: 1 May 2023).

Totempool, B. (2021) 'Silence Brand! What Can This Meme Teach You in Marketing?', 31 May. Available at: <https://totempool.com/blog/silence-brand/> (Accessed: 21 November 2022).

Twitter - Statistics & Facts | Statista (2023) Available at: <https://www.statista.com/topics/737/twitter/#topicOverview> (Accessed: 21 April 2023).

Vasile, S., Mototo, L. and Chuchu, T. (2021) 'USING "MEMES" AS A MARKETING COMMUNICATION TOOL IN CONNECTING WITH CONSUMERS WITHIN THE AGE OF DIGITAL CONNECTIVITY', *International Review of Management and Marketing*, 11(3), pp. 30–35. Available at: <https://doi.org/10.32479/irmm.11313>.

Yasa, I.G.S., Ayomi, P.N. and Maharani, P.D. (2023) 'Stylistic Features of Netspeak Language on 9GAG's Instagram Account', *Journal of Language and Literature*, 23(1), pp. 101–114. Available at: <https://doi.org/10.24071/joll.v23i1.4398>.

Yuksel, M. and Labrecque, L.I. (2016) "'Digital buddies": parasocial interactions in social media', *Journal of Research in Interactive Marketing*, 10(4), pp. 305–320. Available at: <https://doi.org/10.1108/JRIM-03-2016-0023>.

Yus, F. (2011) *Cyberpragmatics: internet-mediated communication in context*. Amsterdam ; Philadelphia: John Benjamins Pub. Co (Pragmatics & beyond new series; v. 213).