Exploring the Distinctiveness of Emoji Use for Digital Authorship Analysis

Karoline Marko

University of Graz, Austria

https://doi.org/10.21747/21833745/lanlaw/7_1_2a9

Abstract. *This exploratory study sets out to investigate the potential distinctiveness of emoji use on the social media platform Instagram. The use of emoji has become popular in digital media, as they provide additional information about a given message; they are said to serve similar purposes to non-verbal cues in face-to-face interactions (e.g Gawne and McCulloch 2019). Several studies have investigated emoji use related to geographical origin, personality traits, age and gender, but their distinctiveness and use for authorship analysis has remained relatively unexplored. Based on a sample of 60 individuals, this study researches not only group-specific characteristics of emoji use on Instagram, but also explores whether or not the use of emoji is distinctive enough to identify individuals simply based on their use of emoji. For this purpose, this study mainly draws on, but also expands, Evan's (2017) emoji function classification framework. The results suggest that individuals do indeed exhibit emoji usage patterns that can be valuable for authorship analysis.*

Keywords: *Emoji, uniqueness, authorship analysis, sociolinguistic profiling.*

Resumo. *Este estudo exploratório tem como objetivo investigar a possível singularidade da utilização de emojis na rede social Instagram. A utilização de emojis popularizou-se nas plataformas digitais por proporcionarem informações adicionais acerca de uma determinada mensagem; eles servirão fins semelhantes às indicações não-verbais em interações face-a-face (e.g Gawne e McCulloch 2019). Diversos estudos investigaram a utilização de emojis relativamente à origem geográfica, a traços de personalidade, idade e género, mas a sua singularidade e utilização para efeitos de análise de autoria continua por explorar. Com base numa amostra composta por 60 pessoas, este trabalho investiga, não só características da utilização de emojis específicas de determinados grupos no Instagram, como também se a utilização de emojis é suficientemente distintiva para identificar indivíduos simplesmente com base na sua utilização de emojis. Para o efeito, este estudo, sobretudo, baseia-se – mas também aprofunda – a grelha de classificação*

das funções dos emojis, de Evans (2017). Os resultados indicam que os indivíduos apresentam realmente padrões de utilização potencialmente preciosos para análise de autoria.

Palavras-chave: *Emoji, singularidade, análise de autoria, perfis sociolinguísticos.*

Introduction

Situated in the field of authorship analysis, this study addresses prevailing new challenges posed by digital media. Authorship analysis, which aims at comparing two or more different texts with the goal of investigating whether the authors of the texts are the same or different, or of creating sociolinguistic profiles based on the language used in order to classify the authors into some category (e.g. gender, age, education, etc.), has been facing many challenges. For instance, it has long been an issue that sample texts used in authorship attribution are too short to yield valid results through the use of statistical methods (e.g. Coulthard 2006; see also Eder 2015, Brennan and Greenstadt 2009). This particular problem has been intensified by technological developments, as newer forms of communication such as text messages are often even shorter than 100 words, and yet have frequently been implicated in crimes (e.g. Coulthard *et al.* 2017, Grant 2010). Sousa Silva (2018) and, even more recently, Heydon (2019) have pointed out that technology impacts on authorship analysis also due to developments like the inclusion of multimodal resources such as gifs, videos, and pictures (including emoji). The focus of this article will be on the use of emoji.

Reflective of its place of invention, emoji is Japanese for "picture character" (Goldman, 2018: 1231). Emoji are particularly common on social media platforms, which are sites "that promote social interaction between participants" (Page *et al.*, 2014: 5). Since emojis are often used to substitute for non-verbal cues and gestures (e.g. McCulloch 2019; Miller *et al.* 2016), it is hardly surprising that they appear so commonly in digital media. Evans (2017: 22) estimates that more than six billion emoji are exchanged on social media every day, and Goldman (2018) reports that "2-3 trillion mobile messages incorporate emojis in a single year" (p. 1229). Further, and particularly relevant for the context of this article, Dimson (2015) argues that approximately 50% of the texts on Instagram contain at least one emoji. These figures are indeed intriguing and give rise to important questions relating to authorship analysis, such as whether it is possible that the use of emoji differs so much among different people that it is possible to identify the author of a post simply based on their use of emoji. This is the question I will investigate in this paper. First, however, it is necessary to establish what previous research has found out about the use of emoji in relation to different groups of people and to provide an overview of research in the field of authorship analysis in relation to digital communication.

Emoji, which can represent emotions, objects, ideas, and even actions (Donato and Paggio, 2017), were introduced in 2011 and due to the sharp rise in use in 2015 (Evans, 2017: 10), many researchers have taken on the task of investigating how emoji use relates to individual users. Theoretically, as emoji (as opposed to emoticons, see e.g. Ai *et al.* 2017), are unified by the Unicode Consortium, all emoji should look the same on each platform. Practically, however, each platform adapts the original Unicode Code so that in reality, emoji appear differently on each platform or device. This has led to many misunderstandings, as Goldman (2018) points out. A much-discussed example in

the context of misunderstandings is the "Astonished face" (U+1F632[1]). As Figure 1 illustrates, the "Astonished face" is rendered quite differently across platforms. The black and white emoji on the left-hand corner is the original emoji provided by Unicode. As Goldman (2018) discusses, the renderings of this particular emoji on some platforms can be interpreted as death threats, while others are more clearly associated with feelings of anger, astonishment, shock, or annoyance. Similarly, the "Grinning Face with Smiling Eyes" (U+1F604) has been associated with both happiness and anger, depending on the platform on which they are depicted (see Miller *et al.* 2016).

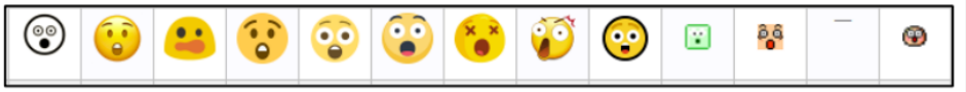

Figure 1. The "Astonished Face" across different platforms (Goldman, 2018: 1258).

Further, Chen *et al.* (2018), for instance, have researched emoji use in relation to gender. They have not only found that women use significantly more emoji than men, but they have also found that women and men prefer different kinds of emoji and use different emoji for expressing similar sentiments (p. 1). Even though many emoji are co-used by males and females, it appears that these emoji are used for different purposes and to convey different emotions or sentiments (p. 2).

Other researchers such as Li *et al.* (2018) have focused on the connection between emoji use and personality. They have correlated emoji use patterns with the Big Five Personality Traits[2] and reached the following conclusions: the trait of openness does not seem to be related to emoji use, while people with higher scores on conscientiousness tend to use fewer emoji. Further, people who scored low on extraversion used the most emoji; people scoring high on agreeableness also tend to use more emoji than those scoring low on this trait. Additionally, people who scored high on neuroticism are the ones who prefer the use of exaggerated expressions (pp. 649-650).

Another interesting strand of research in this respect has investigated links between emoji use and the living conditions of the respective users. Concentrating on a Twitter corpus, Ljubesic and Fiser (2016), for example, have focused on first, second, third and fourth world clusters[3] and obtained the following results: while people in the first world cluster used almost no face emoji, people in the second world cluster used many emoji conveying positive emotions. People in the third world cluster exhibited a high use of unhappy face emoji and the praying hands emoji, while people in the fourth world cluster used many hand gesture emoji. Thus, they conclude, it is possible to track the user's living conditions in different parts of the world simply by taking a close look at which emoji people use most commonly (pp. 87-89).

As outlined above, particular attention has been paid to how emoji use and patterns of emoji use are related to different groups of individuals. Other researchers have taken different approaches and have investigated emoji semantics (Barbieri and Camacho-Collados, 2018), redundancy (Donato and Paggio, 2017), the role of emoji in the law (Goldman, 2018), emoji as gestures (Gawne and McCulloch, 2019), and emoji ambiguity (Miller *et al.*, 2016, 2017). Even though these studies have investigated how people differ in their use and interpretation of emoji, none of them has looked at patterns of emoji use from an authorship perspective, even though Na'aman *et al.* (2017) have hinted at, but

do not elaborate on, idiolectal differences in the use of emoji, by stating that "there can be little question that individuals use emoji differently" (p. 141). The present study is not simply interested in the use of individual emoji, however. Rather, it is the aim to look at how people use emoji, as it is possible that individuals develop certain preferences for emoji use, comparable to "preferred co-selections" (Coulthard, 2008: 146) in language.

It is the aim of this paper to answer the following research questions: (1) Can the use of emoji on Instagram be related to demographic categories such as gender, age, and social group so that findings can be useful for sociolinguistic profiling? (2) Is the use of emoji distinctive enough to identify the author of posts only based on emoji usage patterns in a simple authorship comparison task?

Research in Digital Authorship Analysis

Research into short texts produced online has, for example, garnered the interest of researchers focusing on computational approaches to authorship analysis: Orebaugh and Allnutt (2009) have focused on the classification of instant messaging communications, and Layton *et al.* (2010) have concentrated on authorship attribution in a Twitter corpus. Other examples of automated authorship analysis are Sousa-Silva *et al.* (2011); Rocha *et al.* (2016), and Ishihara (2017). Interestingly, Sousa-Silva *et al.* (2011) have found that in their study, the use of emoticons "outperforms all other feature groups tested" (p. 167), demonstrating the usefulness of non-language features in authorship analysis.

MacLeod and Grant (2012), and Johnson and Wright (2014) have also investigated short online communications (Twitter and e-mails, respectively). Since the analysis of short texts usually defies the use of statistical methods and because "traditional [authorship analysis] methods do not easily translate into computer-mediated communication" (MacLeod and Grant, 2012: 210), these researchers have adapted the use of Jaccard's coefficient and Delta-s calculations to authorship analysis problems. This approach will also be taken in the present article.

Sociolinguistic profiling, in contrast to authorship attribution, is non-comparative (Ehrhardt, 2018). Rather, it categorizes individuals in terms of pre-defined social and demographic categories. As Coulthard *et al.* (2011: 538) state, "profiling involves taking a single example and, by matching it to a well-founded generalization, drawing a conclusion about that instance." Even though sociolinguistic profiling has been improved in recent years (see Nini 2018b), the potential of emoji has not yet been addressed in this connection.

Emoji Classification Systems

A variety of emoji classifications and categorizations exist. Many of these classifications have separated emoji into categories, such as 'traveling/commuting', 'events', 'places', 'feelings', 'people', 'eating and drinking', etc. (Donato and Paggio, 2017). Other such classifications are provided by Emojipedia (online), Lin *et al.* (2014), Barbieri *et al.* (2016), and Vidal *et al.* (2016). Although these classifications are interesting and important (as shown in experiment 2), they only classify individual emoji into different groups but disregard how these emoji function in context.

One of the most important classifications of emoji which takes their function into account comes from Evans (2017: 130-135). He differentiates between six different functions, which are similar to the functions non-verbal cues serve in face-to-face conversation, such as adding emphasis, repeating what is said, or referring to objects and lo-

cations, among others (see also Gawne and McCulloch 2019). The first of Evans' (2017) categories is substitution, which refers to the actual replacement of a word with an emoji. Secondly, emoji can serve the function of reinforcement. This, for instance, means that the emoji conveys the same meaning as the words do, which simultaneously emphasizes the meaning of the words. Further, emoji can be used in a contradictory way, which usually happens in cases when the writer intends to be ironic. Emoji can also serve a complementary function, which refers to something similar as a meta-comment to the words. This can also be regarded as a politeness strategy which has the potential of mitigating possible face-threatening acts. The fifth function is an emphasizing function, meaning that emoji are used to highlight an idea. Lastly, the discourse management function focuses on emoji in initial and final positions. For instance, utterance initial emoji are often used to respond to a previous message, while utterance final emoji can be seen as similar to transition relevance points in conversation, which signal that an idea is complete (see, e.g. Clift 2016).

From a semiotic and communication studies perspective, Danesi (2016) has additionally found that emoji are often used to replace punctuation marks at the end of sentences or salutation formulae at the beginning of messages – an idea that is compatible with Evans' (2017) discourse management function. Also similar to Evans (2017), Danesi (2016) differentiates between an adjunctive and a substitutive use of emoji, but he additionally discusses phatic and emotive functions. The former refers to ways of using emoji for small talk, such as utterance openers, utterance enders, and silence avoiders – thus, they simultaneously serve a pragmatic function; the latter refers to emoji being used to substitute for facial expressions or for emphasizing an idea visually, which is comparable to what Evans (2017) calls substitution or reinforcement.

Data & Methodology

The data for the present study was collected manually from the social media platform Instagram, which is a photo sharing "platform best known for selfies and self-representation" (Leaver *et al.*, 2020). Instagram was chosen for several reasons: first of all, little linguistic research has focused on Instagram. It allows the addition of yet another layer of multimodality to the study of emoji in that emoji can also be used in relation to the image, not just to the text (see below). The second advantage is that Instagram does not have a function that allows users to direct a post to someone in particular. Public posts are visible to potentially anyone who has access to the internet and it is rarely the case that someone is directly addressed through the use of tagging, which is a common feature on Twitter (Zappavigna, 2013). The lack of this affordance on Instagram diminishes the effect of the addressee on the post (Bell, 1984), as all individuals have potentially the same audience – any person with access to the internet.

The data sample in this study is composed of 60 individuals: 30 males and 30 females. The age range is 14 to 69, with a mean age of 26.4 years, a median of 24 years and a mode of 15 years and is thereby reminiscent of the general age distribution of Instagram (see Statista 2020).

8873 posts were included in this analysis. Posts were excluded for the following reasons: if the emoji were only used in hashtags, and if they were used in reposts with the original post present and thus in direct response to someone else's post. This was done for the following reasons: first of all, emoji in hashtags seemingly portray a different

function (see, e.g. Zappavigna 2013) which was not focused on in this analysis. Secondly, if posts are re-posts and include the original post, the language and emoji use of the individual is likely to be influenced by the audience (Bell, 1984). In order to prevent the skewing of findings, such posts were entirely excluded from the analysis.

Ethical Considerations

Great care needs to be taken to avoid harm to the people whose data is included in an analysis based on social media data (Townsend and Wallace, 2016). For this study, only public profiles that can be accessed and viewed by anyone with or without an Instagram account have been included. According to the Instagram policy (Instagram, 2018), data that is publicly available can be accessed and used by third parties. Importantly, since the minimum age for Instagram use is 13, no individuals younger than 13 are included in this study. The individuals' names have been anonymized and no inferences about their identities can be made. Since the posts are publicly available, the wordings of the posts used in this study have been altered to prevent any detection of the individuals through google searches by replacing content words, as previously practiced by Gawne and McCulloch (2019), who refer to Ayers *et al.* (2018).

Emoji Classifications

Emoji Functions

For this analysis, Evans' 2017 classification of emoji functions outlined above was adapted and linked in parts to Danesi's (2016) functions for practical reasons. For instance, the discourse management functions described by Evans (2017) could not be identified in the Instagram data on which this study is based, as there are no ongoing conversations to be analyzed. Thus, although emoji can appear in initial and final positions, they have rather different functions than responding to previous messages or serving as transition-relevance points. Further, based on the data, it was impossible to distinguish clearly enough between emoji that serve as reinforcements and those that serve as emphasizers. The remaining adapted categories are shown in Table 1.

Emoji classifications have concentrated on how emoji can be classified according to the function they serve in relation to the language they accompany. However, on Instagram, emoji use might not only be related to the language in the post; it might as well reflect the picture itself, complement, emphasize, or contradict it. This dimension has to be accounted for by the classification system. Thus, the classification system outlined in Table 1 was further adapted to suit this particular need (see column 4). In the present context this additional dimension is crucial, as individuals may have (un)intentional preferences for their emoji use either reflecting the picture and/or the text, which can have further important implications for authorship analysis.

Prior to the experiments, the applied categories of emoji functions were tested for inter-rater reliability in several steps. Following from these initial tests in which each trial was rated by two researchers and which revealed an inter-rater reliability of 68%, the definitions of the categories of emoji functions were slightly adapted to provide clearer boundaries between the categories. The adapted definitions can be seen in Table 2 below.

Function	*Definition*	*Example*	*Functions related to pictures*
Substitution	Replacing a word with an emoji	I❤️this dress!	n/a
Reinforcement, emphasis	Repeating meaning of words; emphasizing the idea the words convey	I love this dress❤️	Emoji can either reinforce or emphasize the picture they accompany
Contradiction	Emoji expresses the opposite of what the words convey; often used for irony	Thank you very much 🙄	Emoji can contradict the picture
Complementation	Meta-comment to the accompanying words; politeness strategy, frequently encountered in conversation openings (therefore, non-reactive)	Hi, how are you? 🙂	Emoji can complement the words or the image, either in semantic terms, or, for instance in terms of color
Discourse management; placement[4]	Initial position: response to previous turn; Final position: replacing punctuation marks, thereby	🙂 I agree Fantastic 🙂	n/a

Table 1. Emoji Functions (adapted from Danesi 2016 & Evans 2017)[4]

Function	*Definition*	*Example*
Substitution	Replacing a word with an emoji in a sentence; not applicable to emoji outside sentence/clause boundaries	I❤️this dress!
Reinforcement, emphasis	Repeating meaning of words actually present in the sentence; emphasizing the idea the words convey; illustrative purpose	I love this dress❤️
Contradiction	Emoji expresses the opposite of what the words convey; often used for irony	Thank you very much 🙄
Complementation	Meta-comment conveying an additional sentiment/idea; politeness strategy to save face; meaning of the emoji is not present as a word within the sentence it complements	Hope you're ok 🙂
Discourse management; placement	*Excluded as a function*	

Table 2. Adapted Emoji Functions.

Table 2 shows the changes made to the original definitions (Table 1): The category of Substitution is only used when a word inside a sentence is replaced by an emoji. In order to avoid confusion between Reinforcement/emphasis and Complementation, the former will only be used in cases where an emoji conveys the same meaning as the words within the sentence. That is, for an emoji to count as reinforcement or emphasis, a word or phrase needs to be present that has the same meaning as the emoji itself. If an emoji conveys an additional thought or concept that is not directly encoded in the

words used, the emoji will be coded as serving a complementary function. The Discourse Management function was excluded from the present analysis for the reason that any emoji serving any of the other functions can also simultaneously serve a discourse management function. The placement of emoji was thus looked at separately from the emoji functions, since it is still expected to be of analytical value (see Table 4).

After these adaptations were made, another researcher was asked to classify the emoji into their respective categories. This time, the researcher received more prior input and more detailed instructions in addition to the adapted definitions. With these changes in definitions and preparations, inter-rater reliability increased to 85%.

Emoji Taxonomy

The second experiment conducted for this study is not based on the functions the emoji serve but rather on the use of emoji types. Therefore, the following taxonomy of emoji based on Apple's iOS version, as shown in Table 3, is used. The original version of the taxonomy was adapted to the specific needs of this study. Thus, some categories were split into further descriptive categories in order to make follow-up calculations more accurate.

Level 5	Level 4	Level 3	Level 2	Level 1
Emoji	Smileys & People	Smileys	Positive	
			Neutral	
			Negative	
		People	Body parts	
			Real people	Male
				Female
			Mythical people	
		Families		
		Hand gestures		
		Clothing & accessories	Male	
			Female	
		Other	Positive	
			Neutral	
			Negative	
	Animals & Nature	Animals		
		Nature	Astronomy	
			Weather	
			Plants	
			Other	
	Food & Drink	Fruit		
		Vegetables		
		Meals		
		Beverages		
		Utensils		
	Activity	Sports		
		Music		
		The Arts		
		Hobbies		
		Other		
	Travel & Places	Scenes		
		Locations		
		Buildings		
		Modes of transport		
	Objects	Household items		
		Celebrations		
		Stationary		
		Miscellaneous objects		
	Symbols	Hearts		
		Clocks		
		Arrows		
		Signs		
		Shapes		
	Flags	Country		
		Other		

Table 3. Emoji Taxonomy and Corresponding Levels.

Jaccard's Coefficient & Delta-S

As mentioned above, many computational and statistical methods tend to be unreliable with short texts. The use of Jaccard's Coefficient, however, provides a solution to this problem. Jaccard's Coefficient, as outlined in detail in MacLeod and Grant (2012), Grant (2013) and also in Johnson and Wright (2014) and Nini (2018a), was used in the present study to evaluate distances between texts based on emoji. Jaccard's coefficient "establish[es] degrees of similarity between cases" (MacLeod and Grant, 2012: 2013) and Grant (2013: 482) further outlines that Jaccard's coefficient "is a correlation for binary values"; features identified in texts are assigned either 1 (presence) or 0 (absence). Results close to 0 indicate that the investigated texts are completely different; results close to 1 indicate that the texts are the same. An important advantage of Jaccard's Coefficient outlined by

Grant (2013) is that "it does not inflate similarity on the basis of two absences [and thus] does not risk overstating the explanatory power of a single text" (p. 482).

In order to make Jaccard calculations possible, the following features were identified as variables in the analysis: emoji functions (see Table 2), specifics of use, and the placement or position of the emoji (see Table 4). These variables have emerged as important features of emoji use patterns in previous research and in the present corpus and were thus chosen for the analysis.

Variable	*Definition*	*Examples*
Emoji functions	See Table 1	See Table 1
Specifics of use • Strings • Compositions	Strings: consecutive use of the same emoji Composition: use of several different emoji	String: 🌴🌴🌴🌴🌴🌴🌴🌴🌴🌴 Composition: ✨🏀💫
Position	Beginning or end of the post; around a word or phrase for emphasis; "naked" (Provine et al., 2007) or stand-alone	Initial position: 🤩 This! Final position: Happy to be home 💗 Emphasis: 🤩 Fantastic🤩 Stand-alone: 💫

Table 4. Variables used in the Jaccard calculation.

Delta-S (Δs), an extension of Jaccard, "allows the recognition of similar but not identical stylistic choices" (MacLeod and Grant, 2012: 213). Originally used in marine biology and adapted to forensic psychology (Woodhams *et al.*, 2007), MacLeod and Grant (2012) have successfully applied these measures to attribute authorship of Twitter messages. In contrast to the first experiment which uses Jaccard measures, Δs will be used in order to attribute messages to authors based on the emoji types used in the posts rather than on the emoji functions.

In order to simulate a simple authorship attribution analysis, twelve individuals were randomly chosen as authors. These individuals were then randomly grouped into pairs of two, resulting in six pairs. From each individual in each pair, ten posts from the respective data collection period were randomly chosen ('known' writings) and treated as a collective in the analysis; further five posts from one individual in each pair from outside the data collection period were chosen to represent the 'unknown' writings. The reason for choosing only a minimum of posts from each individual is that it is rarely the case that plenty of material for comparison is available in real world authorship analysis cases, and thus a method has to deliver useful results under extreme conditions (e.g. Johnson and Wright 2014). Future studies will take into account other scenarios, such as authorship attribution with more than two candidate authors, but for this exploratory study, it was decided to focus on a simple authorship attribution problem.

Analysis & Results

In total, 10573 emoji (690 different ones) were used by the individuals in the sample. Only the 25 most common ones occur more than 100 times, and 199 occur only once. Among the 20 most common emoji are seven different types of hearts, with the red heart (U+2764) as the most frequent one. The second most common emoji are the camera (U+1F4F7) and the camera with flash (U+1F4F8), which is likely explained by the fact that Instagram is a photo sharing platform and the camera emoji is used to substitute for the phrase "picture credits".

Emoji Use for Sociolinguistic Profiling

As a first step, the 60 individuals' emoji use related to gender was investigated in more detail. This analysis has revealed the following patterns: females use 56.3% of all emoji in the data set, while males use only 43%. Interestingly, a female dominance in emoji usage also emerges when the variety of emoji use was investigated: the females in the sample use a mean of 63.9 different emoji in their posts, while the males use a mean of 46.3 different emoji.

A further aspect that can be important for sociolinguistic profiling is the estimation of age of an individual. Figure 2 shows that the age group of 14-19 exhibits the highest emoji use with a combined total of 43% of all emoji in the data. It is clearly visible that emoji use steadily declines with increasing age.

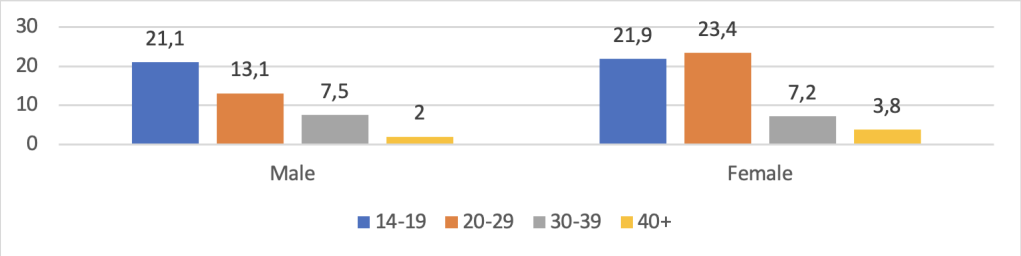

Figure 2. Emoji use according to age and gender (percentages).

Figure 3 further shows that the females in the age group of 20-29 use slightly more different emoji than the younger ones, which is also visible in Figure 2 with regard to overall emoji use. The high variety of emoji use in the category male/40+ is due to one outlier. If this outlier is disregarded, the figure is at 18.5. Thus, the findings support the results of previous studies and, for the first time, demonstrate that emoji use patterns are similar on Instagram with regard to sociodemographic features as it is on other social media platforms or instant messaging, thereby suggesting potential generalizability of results.

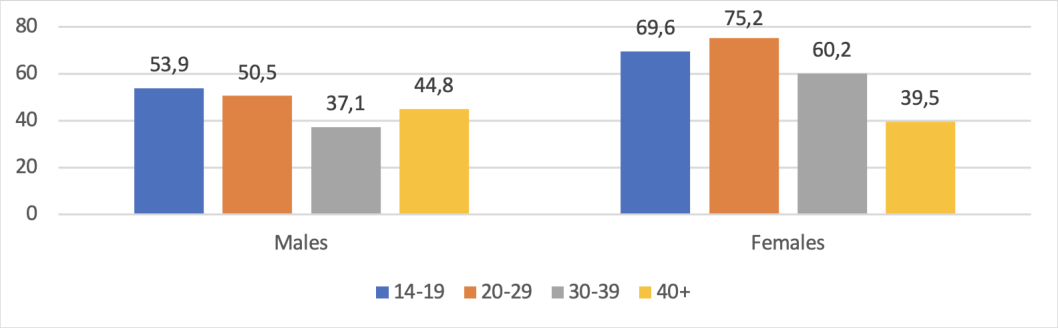

Figure 3. Comparison of mean variety of emoji according to age and gender.

Further highly promising results were obtained when investigating actual emoji use by the different age groups. Firstly, based on the 20 most common emoji in each group, the youngest age group uses hearts most frequently, followed by faces and gestures. In the age group of 30-39-year-olds, gestures are the most common, followed by objects and faces, while hearts are the least common. The age group of 20-29-year-olds employs faces and hearts the most, while gestures are used more sparsely. The oldest age group makes common use of hearts and almost no use of gestures (see Figure 4). Additional research is required in order to allow generalizations of these findings on any level, but these initial results from this exploratory study seem promising for authorship profiling.

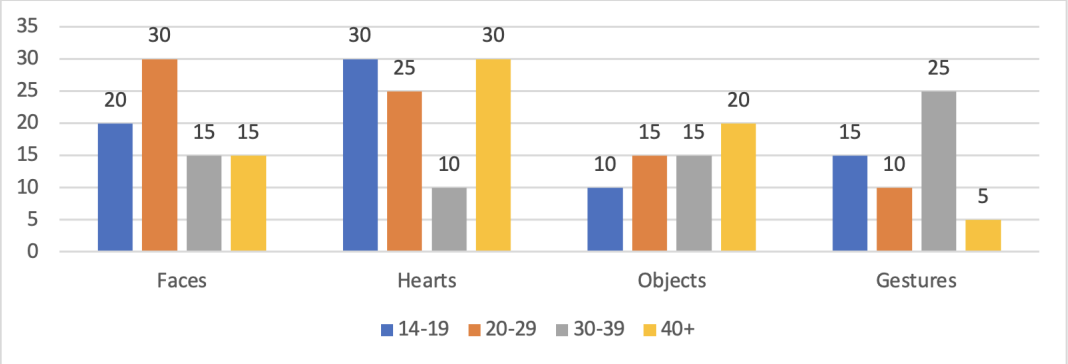

Figure 4. Emoji use according to age group (percentages).

Differentiating Individuals Based on their Use of Emoji

Subsequently, the findings of the Jaccard and Δs calculations will be outlined. The first experiment employs Jaccard's coefficient to attribute 'unknown' posts to one individual in a pair based on the functions of the emoji used (see Table 2). For the second experiment employing Δs, it was not the emoji functions but emoji types that were focused on, since each emoji can potentially serve any of the functions outlined in Table 2.

Experiment 1: Jaccard's Coefficient

As briefly outlined above, 12 individuals were randomly chosen from the data set and then grouped in pairs, resulting in six pairs (i.e. six trials). Each trial consists of 25 posts in total. The analysis proceeded as outlined in section 'Jaccard's Coefficient & Delta-S' above.

Results

In order to illustrate the most important findings from this experiment one trial will serve as an example, but results will be provided for all six trials.

The following examples are taken from Trial 5. Examples (1) to (3) are selected from the 'anonymous' posts of one of the authors in this trial; examples (4) to (6) are taken from the ten randomly selected posts by M.S. and examples (7) to (9) are randomly selected from the ten posts by S.A. The respective categories are indicated in square brackets.

Examples (1) – (3): "Anonymous" posts

(1) [name] is getting married, sorry guys 💍 [complementary] 🤦🏼‍♀️ [reinforcement]

(2) We are now officially open – [name] and I have been working hard, we are really excited! Check it out!!! 💗 [complementary]

(3) My sweet [name], you are the best boyfriend on this planet and I love you! Happy birthday! 💕[reinforcement]

Examples (4) – (6): M.S.'s posts

(4) Americans 🇺🇸 [reinforcement]

(5) This is✈️ boring [substitution]

(6)⚠️ [reinforcement of the picture showing a similar sign]

S.A.'s posts

(7) Only a few more weeks with this guy 🚣🏻‍♂️ [reinforcement of the picture showing the ocean]

(8) First day in Venice ⭐ [complementary]

(9) smiling all day 🌼 [reinforcement of the picture showing yellow flowers]

Applying Jaccard's Coefficient to Trial 5 revealed the following results: S.A.'s posts are slightly closer to the anonymous posts in terms of emoji functions compared to M.S.'s posts. Both of them use emoji in a similar way, but M.S. uses emoji in more different positions than S.A. does. In contrast, S.A. uses emoji more directly in conjunction with the respective picture. Table 4 below shows the results of all six trials.

Trial	Jaccard (A + anonymous)	Jaccard (B + anonymous)	Results & Solution
1	0.4 (B.Y.)	0.5 (Co.C.)	Correct
2	0.6 (C.S.)	0.25 (C.P.)	Correct
3	0.6 (S.Cu.)	0.7 (Ca.S.)	Correct
4	0.3 (M.Y.)	0.4 (R.J.)	Incorrect
5	0.5 (M.S.)	0.6 (S.A.)	Correct
6	0.6 (K.C.)	0.6 (J.B.)	Undecided

Table 5. Overview of Results of Experiment 1 (Jaccard).

As indicated in Table 5, four trials can be considered successful, while one trial lead to a wrong result and one is undecided. It can be seen that the outcome of most trials is very close, with the exception of Trial 2, which shows a very clear result. This can be attributed to the test itself not being sensitive enough. That is, the test only differentiates whether specific emoji functions are present or absent but disregards the extent to which a specific function is employed.

A further interesting finding related to the meanings of emoji has emerged. Examples (10) and (11) below show that although both H.B. and J.P. use the rainbow emoji, they use them to denote different things: H.B.'s use of the rainbow emoji is related to her sport: nature, the ocean, and (actual) rainbows. In contrast, J.P.'s use of the rainbow in example (11) is connected to gay pride. Although these are only two examples, the same use of the rainbow by these individuals can be found throughout their posts. Chen et al.'s (2018) study has yielded similar results in relation to gender differences, and these differences might arise from the different contexts these individuals find themselves in. Thus, in this respect, even though the rainbows are potentially used with the same function (even though this is not the case in Examples 10 and 11), they are used to refer to different concepts and ideas, which indicates a qualitative difference the purely quantitative analysis cannot account for.

Example (10) H.B.

> Surfing 🌈 📸: @name

Example (11) J.P.

> Proud to have been a part of this… thank you @name and happy pride 🌈 💗

Further interesting individual differences emerge when looking at the use of emoji by individuals in relation to how emoji are officially defined. Two common emoji in the dataset will serve as examples, namely the

(U+1F919) and the

(U+26A1). The first one is illustrated in Example (12). K.S. uses the emoji with the meaning 'hang loose' and thereby signifies her identity as either Hawaiian and/or member of a particular local community. However, the official name of this emoji is "call me" (emojipedia, online) – a meaning which is not intended in this context.

Example (12) K.S.

Mahalo NY for all the good memories!! 🙏👍🏻

A further case in point is the following post by J.H. (Example 13), who uses the

⚡

emoji to represent movement. Officially, it is described as a symbol for high voltage, or a representation of lightning. Thus, J.H. clearly adapts the latter meaning and uses it in a metaphorical sense.

Example (13) J.H.

Another wave⚡

These examples illustrate the individual use of emoji in the following ways: individuals use emoji with a specific meaning in mind. This specific meaning might be shared by their community of practice (see Eckert 2006), which could lead to other meanings of the same emoji being lost. As Goldman (2018) has pointed out, this has already happened for a small number of emoji, resulting in what he calls 'emoji dialects'. Taking these individual meanings into account might be of high value in authorship analysis, even though further research is required.

Experiment 2: Δs

A detailed description of the procedure of calculating Δs can be found in Woodhams *et al.* (2007) and in Izsak and Price (2001). As mentioned above, Δs allows for the recognition of similarities much more so than Jaccard's coefficient does. Therefore, this measure was chosen to investigate whether or not it is possible to differentiate between authors in the trial data simply based on the emoji types the individuals use. The emoji taken from examples (14) to (22) are presented below for illustration of how they were classified according to Table 3 above.

Examples (14) – (16): "Anonymous" posts

(14) 💍 [Smileys & People/Clothing & Accessories/Female] 👰🏼 [Smileys & People/People/Real/Female]

(15) 💗 [Symbols/Hearts]

(16) 💕[Symbols/Hearts]

Examples (17) – (19): M.S.'s posts

(17) 🇺🇸 [Flags/Country]

(18)✈️ [Travel & Places/Modes of Transport]

(19)⚠️ [Symbol/Sign]

Examples (20) – (22): S.A.'s posts

(20) 🗻 [Animals & Nature/Nature/Other]

(21) ⭐ [Animals & Nature/Nature/Astronomy]

(22) 🌼 [Animals & Nature/Nature/Plants]

All emoji in the selected posts were classified according to this taxonomy. Table 6 shows the results of all trials.

Trial	**△s (A + anonymous)**	**△s (B + anonymous)**	**Results & Solution**
1	0.4 (B.Y.)	0.5 (Co.C.)	correct
2	0.33 (C.S.)	0.16 (C.P.)	correct
3	0.32 (S.Cu.)	0.33 (Ca.S.)	incorrect
4	0.05 (M.Y.)	0 (R.J.)	correct
5	0.08 (M.S.)	0.2 (S.A.)	correct
6	0.43 (T.W.)	0.45 (J.B.)	incorrect

Table 6. Overview of Results of Experiment 2 (Delta-s).

Four of the anonymous posts were attributed to their actual authors (trials 1, 2, 4, 5), while two trials resulted in incorrect attributions (trials 3, 6). Since it was not expected that posts can be attributed to individuals based only on the actual emoji used, these results are quite surprising. These results indicate that it is indeed worth looking at emoji types in addition to emoji functions, particularly when emoji meanings are taken into account as well. Moreover, the taxonomy of emoji has much potential for development, which will also very likely improve results.

Discussion & Conclusions

The overview of existing literature has revealed that emoji are used quite differently by different groups of people; and yet, emoji have been largely neglected in authorship analysis. For the purpose of filling this research gap, two analyses were carried out: first of all, the data was used to reveal whether or not patterns identified in previous studies, particularly those relating to gender and age, would also be found in posts on the social media platform Instagram. The findings support the results of previous studies and they indicate similar trends: females and younger people use both more and a larger variety

of emoji. Another promising avenue for future research is the investigation of emoji functions in relation to age and gender in order to investigate whether differences exist in this respect as well. Additionally, it was shown that even though people use the same emoji, they might use them in different contexts, thereby denoting different meanings. An important issue for authorship analysis that needs to be investigated further is how far individuals might be aware of different meanings of emoji and whether or not they consciously switch between these meanings, or whether they consistently use the same emoji with the same meaning. In the latter case, several of the more ambiguous emoji could prove useful in authorship analysis. If these findings are tested and developed further, they could make a valuable contribution to sociolinguistic profiling tasks in digital media.

Secondly, in the context of a simple mock authorship attribution task, an experiment was conducted in order to find out about individual's use of emoji with regard to the emoji functions. The calculation of Jaccard's coefficient demonstrates that an investigation of emoji functions can indeed be valuable for authorship analysis, even though the classification system itself still needs to be improved. The main limitation of this study remains the inter-rater reliability. Importantly, however, this pilot study reveals the potential of an analysis of an individual's emoji use in addition to a purely linguistic analysis. Regarding emoji functions, this paper demonstrates the importance of complementary qualitative analyses in conjunction with quantitative analyses. Neglecting a qualitative analysis in this context could result in a loss of valuable information and might even mislead the analysis. Additionally, using complementary qualitative analyses that can easily be explained to a judge or jury is likely to be viewed favorably in actual forensic cases (e.g. Grant and Baker 2001; Solan 2013; Grant and MacLeod 2020).

Thirdly, Δs was used to investigate an individual's use of emoji regardless of the emoji function. The relative success of the second experiment might be due to emoji being used as identity markers, as previous research has indicated (e.g. Robertson *et al.* 2018; Ge 2019). Further, this portrayal of identity might be particularly strong on Instagram (Leaver *et al.*, 2020), which could explain the relatively good results for this platform.

Overall, it was the aim to investigate the potential of emoji as authorship markers. Even though any other linguistic markers were neglected in this paper, the results are very promising. Nevertheless, the classification systems require further refinement to be used in forensic cases. As this paper demonstrates, even though emoji use should not be relied upon as an individual marker of authorship, it should not be neglected either and can serve as a valuable addition to authorship analysis methods.

Notes

[1]Codes in brackets refer to the original Unicode Codepoints.

[2]Agreeableness, openness, extroversion, conscientiousness, and neuroticism (see, e.g., Roccas *et al.* 2002).

[3]The first world cluster includes North America, Western Europe, the Russian Federation, and Australia; the second world cluster covers most of South America, India and China, Eastern Europe, Morocco, Algeria, and Tunisia; the third world cluster includes Angola, Nigeria, Sudan, Jordan, Saudi Arabia, Yemen, Pakistan, Nepal, and the Philippines; the remaining African states are subsumed under the fourth world cluster (Ljubesic and Fiser, 2016: 86)

[4]The description of the 'Discourse Management Placement' category is based on research carried out by Evans (2017) and Danesi (2016). Due to this category overlapping to a large degree with some of the other categories, the discourse management function was later abolished for the purposes of this study (see Table 2) and later reintroduced and adapted an additional feature of emoji use (see Table 4).

References

Ai, W., Lu, X., Liu, X., Wang, N., Huang, G. and Mei, Q. (2017). Untangling Emoji Popularity through Semantic Embeddings. In *Proceedings of the 11th Int'l AAAI Conference on Web and Social Media*, 2–11.

Ayers, J. W., Caputi, T. L., Nebeker, C. and Dredze, M. (2018). Don't quote me: reverse identification of research participants in social media studies. *Digital Medicine*, 1(1), 30.

Barbieri, F. and Camacho-Collados, J. (2018). How Gender and Skin Tone Modifiers Affect Emoji Semantics in Twitter. In *Proceedings of the 7th Joint Conference on Lexical and Computational Semantics*, 101–106.

Barbieri, F., Ronzano, F. and Saggion, H. (2016). What does this emoji mean? In *Language Resources and Evaluation Conference, LREC.*, Portoroz, Slovenia.

Bell, A. (1984). Language Style as Audience Design. *Language in Society*, 13(2), 145–204.

Brennan, M. and Greenstadt, R. (2009). Practical attacks against authorship recognition techniques. In *Proceedings of the 21st Innovative Applications of Artificial Intelligence Conference*, 60–66.

Chen, Z., Lu, X., Ai, W., Li, H., Mei, Q. and Liu, X. (2018). Through a Gender Lens. In *Proceedings of the 2018 World Wide Web Conference on World Wide Web - WWW '18*, 763–772, New York, New York, USA: ACM Press.

Clift, R. (2016). *Conversation Analysis*. Cambridge: Cambridge University Press.

Coulthard, M. (2006). …and then… Language Description and Author Attribution. Inaugural Lecture at Aston University.

Coulthard, M. (2008). By their words shall ye know them: On linguistic identity. In C. R. Caldas-Coulthard and R. Iedema, Eds., *Identity Trouble*. New York, NY: Palgrave Macmillan, 143–170.

Coulthard, M., Grant, T. and Kredens, K. (2011). Forensic linguistics. In R. Wodak, B. Johnstone and P. Kerswill, Eds., *The SAGE handbook of sociolinguistics*. London: SAGE, 531–544.

Coulthard, M., Johnson, A. and Wright, D. (2017). *Introduction to Forensic Linguistics*. London England & New York, NY: Routledge.

Danesi, M. (2016). *The Semiotics of Emoji*. Bloomsbury.

Dimson, T. (2015). Emojineering Part 1: Machine Learning for Emoji Trends. *Instagram Blog, Engineering*.

Donato, G. and Paggio, P. (2017). Investigating Redundancy in Emoji Use: Study on a Twitter Based Corpus. In *Proceedings of the 8th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*, 118–126.

Eckert, P. (2006). Communities of Practice. In K. Brown, Ed., *Encyclopedia of Language & Linguistics*. Amsterdam, The Netherlands: Elsevier Science, 683–685.

Eder, M. (2015). Does size matter? Authorship attribution, small samples, big problem. *Digital Scholarship in the Humanities*, 30(2), 167–182.

Ehrhardt, S. (2018). Authorship attribution analysis. In J. Visconti and M. Rathert, Eds., *Handbook of communication in the legal sphere, 14*. Berlin, Germany: de Gruyter, 169–200.

Evans, V. (2017). *The Emoji Code*. New York: Picador.

Gawne, L. and McCulloch, G. (2019). Emoji as Digital Gestures. *Language@Internet*, 17.

Ge, J. (2019). Emoji Sequence Use in Enacting Personal Identity. In *Companion Proceedings of The 2019 World Wide Web Conference*, 426–438, New York, NY, USA: ACM.

Goldman, E. (2018). Emojis and the law. *Washington Law Review*, 92, 1227–1291.

Grant, T. (2010). Txt 4n6: Idiolect free authorship analysis. In M. Coulthard and A. Johnson, Eds., *Routledge Handbook of Forensic Linguistics*. Abingdon and New York: Routledge.

Grant, T. (2013). Txt 4n6: Method, consistency, and distinctiveness in the analysis of SMS text messages. *Journal of Law and Policy*, 21(2), 467–494.

Grant, T. and Baker, K. (2001). Identifying reliable, valid markers of authorship: A response to Chaski. *Forensic Linguistics*, 8(1), 66–79.

Grant, T. and MacLeod, N. (2020). *Language and Online Identities*. Cambridge: Cambridge University Press.

Heydon, G. (2019). *Researching Forensic Linguistics*. London & New York: Routledge.

Instagram, (2018). Privacy and Security.

Ishihara, S. (2017). Strength of linguistic text evidence: A fused forensic text comparison system. *Forensic Science International*, 278, 184–197.

Izsak, C. and Price, A. R. G. (2001). Measuring β-diversity using a taxonomic similarity index and its relation to spatial scale. *Marine Ecology Progress Series*, 215, 69–77.

Johnson, A. and Wright, D. (2014). Identifying idiolect in forensic authorship attribution: an n-gram textbite approach. *Language and Law / Linguagem e Direito*, 1(1), 37–69.

Layton, R., Watters, P. and Dazeley, R. (2010). Authorship attribution for Twitter in 140 characters or less. *IEEEI*, 1–8.

Leaver, T., Highfield, T. and Abidin, C. (2020). *Instagram*. Cambridge: Polity.

Li, W., Chen, Y., Hu, T. and Luo, J. (2018). Mining the Relationship between Emoji Usage Patterns and Personality. In *Proceedings of the 12th International AAAI Conference on Web and Social Media*, 648–651.

Lin, T., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P. and Zitnick, C. L. (2014). Microsoft coco: Common objects in context. In *European Conference on Computer Vision*, 740–755.

Ljubesic, N. and Fiser, D. (2016). A Global Analysis of Emoji Usage. In *Proceedings of the 10th Web as Corpus Workshop and the Empirist Shared Task*, 82–89.

MacLeod, N. and Grant, T. (2012). Whose Tweet? Authorship analysis of micro-blogs and other short-form messages. In *Proceedings of the International Association of Forensic Linguists' 10th Biennial Conference*, 210–224, Birmingham, UK: Aston University.

McCulloch, G. (2019). *Because Internet*. New York: Riverhead Books.

Miller, H., Kluver, D., Thebault-Spieker, J., Terveen, L. and Hecht, B. (2017). Understanding Emoji Ambiguity in Context: The Role of Text in Emoji-Related Miscommunication. In *Proceedings of the 11th International AAAI Conerence on Web and Social Media*, 152–161.

Miller, H., Thebault-Spieker, J., Chang, S., Johnson, I., Terveen, L. and Hecht, B. (2016). "Blissfully Happy" or "Ready to Fight": Varying Interpretations of Emoji. In *Proceedings of the 10th AAAI Conference on Web and Social Media*, 259–268.

Na'aman, N., Provenza, H. and Montoya, O. (2017). MojiSem: Varying linguistic purposes of emoji in (Twitter) context. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics- Student Research Workshop*, 136–141.

Nini, A. (2018a). An authorship analysis of the Jack the Ripper letters. *Digital Scholarship in the Humanities*, 33(3), 621–636.

Nini, A. (2018b). Developing forensic authorship profiling. *Language and Law / Linguagem e Direito*, 5(2), 38–58.

Orebaugh, A. and Allnutt, J. (2009). Classification of instant messaging communications for forensic analysis. *The International Journal of Forensic Computer Science*, 1, 22–28.

Page, R., Barton, D., Unger, J. W. and Zappavigna, M. (2014). *Researching Language and Social Media*. London & New York: Routledge.

Robertson, A., Magdy, W. and Golwater, S. (2018). Self-representation on Twitter using emoji skin color modifiers. In *Proceedings of the 12th International AAAI Conference on Web and Social Media*, 680–683.

Roccas, S., Sagiv, L., Schwartz, S. H. and Knafo, A. (2002). The Big Five personality factors and personal values. *PSPB*, 28(6), 789–801.

Rocha, A., Scheirer, W., Forstall, C., Cavalcante, T., Theophilo, A., Shen, B., Carvalho, A. R. B. and Stamatatos, E. (2016). Authorship attribution for social media forensics. *IEEE Transactions on Information Forensics and Security*, 12(1), 1–30.

Solan, L. (2013). Intuition versus algorithm: The case of forensic authorship attribution. *Journal of Law and Policy*, 21, 551–576.

Sousa Silva, R. (2018). Computational Forensic Linguistics: An Overview. *Language and Law / Linguagem e Direito*, 5(1), 80–96.

Sousa-Silva, R., Laboreiro, G., Sarmento, L., Grant, T., Oliveira, E. and Maia, B. (2011). 'twazn me!!! ;(' Automatic Authorship Analysis of Micro-Blogging Messages. In R. Muñoz, A. Montoyo and E. Métais, Eds., *Lecture Notes in Computer Science 6716 Springer 2011*, volume Natural La, 161–168, Berlin and Heidelberg: Springer – Verlag.

Statista, (2020). Distribution of Instagram users worldwide as of April 2020, by age and gender.

Townsend, L. and Wallace, C. (2016). *Social media research: A guide to ethics*. University of Aberdeen & ESRC.

Vidal, L., Ares, G. and Jaeger, S. R. (2016). Use of emoticon and emoji in tweets for food-related emotional expression. *Food Quality and Preference*, 49, 119–128.

Woodhams, J., Grant, T. and Price, A. R. G. (2007). From marine ecology to crime analysis: improving the detection of serial sexual offenders using a taxonomic similarity measure. *The Journal of Investigative Psychology and Offender Profiling*, 4, 17–27.

Zappavigna, M. (2013). *Discourse of Twitter and Social Media. How We Use Language to Create Affiliation on the Web*. London: Bloomsbury.