# Facebook Sentiment: Reactions and Emojis

**Ye Tian[1], Thiago Galery[2,3]**
[1]Laboratoire Linguistique Formelle,
Université Paris-Diderot, Paris, France
[2]DBpedia Association
& [3]Idio, London, UK
`tiany.03@gmail.com`
`tgalery@gmail.com`

**Giulio Dulcinati[4],
Emilia Molimpakis[4] & Chao Sun[4]**
[4]Division of Psychology
and Language Sciences,
University College London, London, UK
`g.dulcinati@gmail.com`
`emilia.molimpakis@ucl.ac.uk`
`chao.sun.13@ucl.ac.uk`

## Abstract

Emojis are used frequently in social media. A widely assumed view is that emojis express the emotional state of the user, which has led to research focusing on the expressiveness of emojis independent from the linguistic context. We argue that emojis and the linguistic texts can modify the meaning of each other. The overall communicated meaning is not a simple sum of the two channels. In order to study the meaning interplay, we need data indicating the overall sentiment of the entire message as well as the sentiment of the emojis stand-alone. We propose that Facebook Reactions are a good data source for such a purpose. FB reactions (e.g. "Love" and "Angry") indicate the readers' overall sentiment, against which we can investigate the types of emojis used the comments under different reaction profiles. We present a data set of 21,000 FB posts (57 million reactions and 8 million comments) from public media pages across four countries.

## 1 Introduction

We use social media not only to share information, but also to express emotions. This paper presents a data set of multi-cultural Facebook (FB) posts from public media pages, the readers' reactions and the emojis contained in the comments. We argue that this data set - one that can be up-scaled in size, in genres, and in languages/cultures - is a useful and cheap resource for investigating the types of emojis used in different emotional contexts.

## 2 Emojis and Sentiment - Some background

Emoticons, such as ";)", are representations of facial expressions using punctuation symbols. They were first used by the computer scientist Scott Fahlman in 1982 as a "joke marker" (Fahlman, 2012). Recently, emoticons have been gradually replaced by emojis, which are graphic symbols representing facial expressions (e.g. smiling), gestures (e.g. thumbs up), objects (e.g. vehicles) and even actions (e.g. dancing). They have gained popularity rapidly in smartphone texts, emails and social media. On certain platforms (e.g. Instagram), in some countries (e.g. Finland and France), over half of all online messages contain emojis, and this trend is going up worldwide (Dimson, 2015).

Emojis have attracted an increasing amount of research interest in sociology and in computer science. Sociological research is interested in how people with different demographic profiles (age, gender and culture) use emoticons and emojis, how it affects people's relationships and and how it fits the cultural contexts (Huffaker and Calvert, 2005; Sugiyama, 2015; Wolf, 2000; Kelly and Watts, 2015). Research in computer science has primarily focused on using emoticons and emojis as a cue for automatically analysing the sentiment of short messages, commonly tweets (Hu et al., 2013; Novak et al., 2015; Thelwall et al., 2010; Boia et al., 2013; Zhao et al., 2012; Hogenboom et al., 2013). It was found that positive emoticons and emojis are used more frequently than negative ones (Novak et al., 2015). The polarity of emoticons and emojis is relatively well correlated with the perceived emotional polarity of the entire text, but is poorly correlated with the perceived emotional polarity of the accompanying linguistic text alone (Boia et al., 2013). Using emoticons and

emojis as a cue for sentiment analysis of tweets results in better accuracy compared to using the linguistic text alone (Hogenboom et al., 2013; Hu et al., 2013; Zhao et al., 2012), to a level between 60% to 75%. Emojis tend to be a better indicator for an overall negative tweet than a positive one.

Although the polarity of emojis frequently mismatch the polarity of the accompanying linguistic text or even the entire message, little has been done to analyze the nature of these mismatches. The default assumption is that emojis express the user's emotional state, therefore they can be seen as an independent channel of communication from that of the linguistic text. For example, (Novak et al., 2015) offered sentiment scores for 751 most frequently used emojis, calculated using the sentiment rating of 1.6 million tweets in 13 European languages containing these emojis. This significant piece of work has provided a lot of information on the cross-linguistic usage of emojis in tweets. However, treating the average sentiment of tweets containing emojis as the sentiment score of the emojis themselves relies on the assumption that the meaning of emojis are consistent across linguistic contexts.

We argue that emojis and the linguistic text can modify the meaning of each other. The overall communicated meaning is not a simple sum of the two channels. A similar view has been voiced by some linguists. (Baron, 2009) points out that just like linguistic words, the meaning of emoticons and emojis are often under-specified. (Dresner and Herring, 2010) argues against the idea that emoticons are signs of emotion. Drawing on speech act theory, they argue that emoticons are indicators of the illocutionary force of the textual utterance that they accompany. They "neither contribute to the propositional content (the locution) of the language used, nor are they just an extralinguistic communication channel indicating emotion" (Dresner and Herring, 2010) [pp. 255].

We propose that an emoji can interact with the linguistic text in six ways. An emoji can

1. replace a word/phrase.
   e.g. *I want have a* 🍺.

2. repeat a word/phrase (accenting, adding focus)
   e.g. *Take note* 📝*Sam, this is how you season food, you are almost done there babe. Like you did the chicken* 🍗*the other nights.*

3. express the speaker's emotion or attitude independently.
   e.g. *(Facebook update from survivor of the Florida gay club shooting 2016-06-12): I am safely home and hoping everyone gets home safely as well.* 😭

4. enhance/ emphasize an emotion expressed in the text.
   e.g. *This would probably be really good* 😊.

5. modify the meaning of linguistic text (e.g. marking non-literal or non-serious use); implying propositional content
   e.g. *I bet you are enjoying your revision* 😉.

   *-A: Would you like to come to my party? -B:* 😬

6. be used for politeness.
   e.g. *Can you please cook us something that I tag you in instead of your 4am pastas? Thanks.* 😊

We hypothesize that compared to negative emojis, positive emojis are more often used not as direct reflection of emojis, but are used (1) ironically in a negative context, or (2) for politeness reasons (e.g. in a request or disagreement). These uses are also seen in smiles and laughter in natural dialogue (Mazzocconi et al., 2016). In face-to-face conversations, we may produce an ironic laughter to communicate that an attempted joke was not funny, or smile when we ask for a favour.

In order to study the meaning interplay between linguistic texts and emojis, we need to model contexts where the sentiment of the emojis are consistent with the overall sentiment of the texts, as well as contexts where they are inconsistent. Obtaining such data would normally require a large amount of manual labeling. Instead, we propose that we can cheaply obtain data set for this purposes by looking at Facebook Reactions and emojis in comments.

## 3 Facebook reactions

Facebook reactions, released in February 2016, are an extension of the old "Like" button. Its six options (Like, Love, Haha, Wow, Sad and Angry) are represented by slightly edited versions of several long-established Unicode Emojis, and they allow for a more nuanced expression of how

users feel towards a post. The emotions underlying these six reactions are supposed to be frequent and universal. If we assume that Facebook reactions reflect the readers' overall sentiment towards a post, we can investigate the distributions of emojis in readers' comments, under different emotional attitudes. Thus, if there is a mismatch in the emotional polarity between the overall profile of reactions (e.g. dominantly "Angry" - negative) and the sentiment of the emojis in the comments (e.g. "thumbs up" - positive), these emojis are likely used not to directly reflect emotions.

## 4 Current data set and analysis

We collected data from 21,000 posts in August 2016 on public media Facebook pages from four countries: UK (bbc, the Daily Mail, Daily Mirror, the Telegraph), US (CNN, Fox News, New York Times, Wall Street Journal, MSNBC), France (Le Figaro, Le Monde, Liberation) and Germany (Die WELT, Frankfurter Allgemeine Zeitung, Suddeutsche Zeitung). These posts were shared 15,273,365 times. There were 57,444,404 reactions and 8,463,602 comments to these posts. We tried to balance the political leanings of the selected media. The purpose of analyzing data from different countries was to see whether the way we use Facebook Reactions and emojis can be generalized cross-culturally/linguistically. We analyzed the reactions, sharing behaviours, and the emojis that appeared in the comments. The data set contains country, name of media, the title of the posts, the link to the full article (if any), the time of posting, the total times shared, the total number of reactions, a breakdown of reactions (Likes, Loves, etc), the total number of comments, and the texts of the comments[1].

## 5 Results - Facebook Reactions

In terms of reactions (figure1), we found that "Like", being the default reaction, is unsurprisingly the most frequent (overall 78.9%). The frequency of the other fie reactions ranks as "Love" (5.5%), "Angry" (5.4%) "Sad" (4.0%), "Haha" (3.7%), "Wow" (2.5%).

There are small but statistically different differences cross countries (p< 2.2e-16). "Angry' is the most frequent in France at 9% and the least frequent in the UK at 3%. "Love" is the highest in the

---

[1]As of 23/02/2017, the authors are still discussing with Facebook regarding the details of data release options.

Figure 1: Facebook Reactions distribution (the small pie shows the zoomed-in distribution of reactions other than "Like")
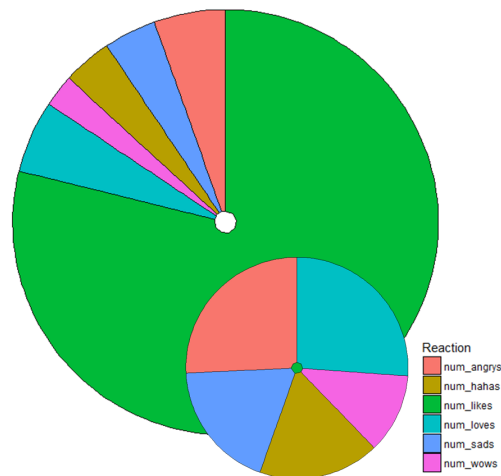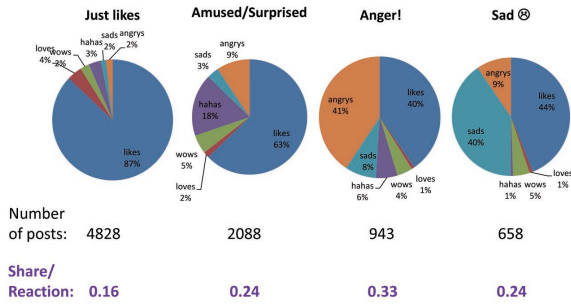


US at 6% and lowest in Germany at 2%. "Haha" is the highest in Germany at 6% and lowest in the UK at 3%. The overall comments to reaction ratio is 0.15, and sharing to reaction ratio is 0.27.

We calculated the proportions of each of the six Reactions for all posts (for example, a post could have 80% "Like", 10% "Wow", 5% "Haha", 5% "Love", and 0% of "Angry" or "Sad"), and applied K-means clustering algorithm to these Reaction proportions. We found four major profiles of Reactions (figure 2), which we label "Just likes", "Amused/Surprised", "Angry" and "Sad". The first cluster is characterized by almost all reactions being "Like". The second cluster has a significant percentage of "Haha"s but also some "Angry"s. In both the "Angry" and "Sad" clusters, less than half of the reactions are "Like"s. In the "Angry" cluster, 41% of reactions are "Angry" and 8% are "Sad". In "Sad" cluster, 40% of reactions are "Sad" while 9% are "Angry". We calculated the Share to Reaction ratios (number of Shares/ number of Reactions), and found them to be different in different Reaction profiles: 0.16 for "Just likes", 0.24 for "Amused/Surprised", 0.33 for "Angry" and 0.24 for "Sad": people are more likely to share a post when the reaction is something other than "Like", suggesting that stronger emotional attitude leads to more post sharing.

### 5.1 Results - Emojis

We randomly sampled 100,000 emoji-containing comments, and processed the emojis by matching them against a dictionary of emoji Unicode.

Figure 2: Four profiles of FB Reactions



| | | | |
|---|---|---|---|
| Number of posts: | 4828 | 2088 | 943 | 658 |
| Share/ Reaction: | 0.16 | 0.24 | 0.33 | 0.24 |

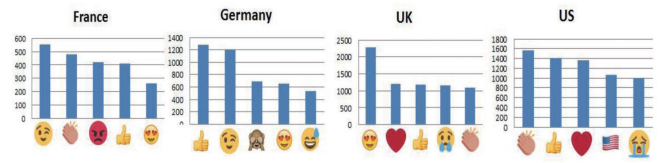Figure 4: FB Emoji distribution by country



These emojis can then be more easily counted and be matched with the emoji sentiment score from (Novak et al., 2015). We found that in our data, the most frequent emojis are "thumbs up", "face with heart shaped eyes", "claps", "angry face" and "winking face" (Figure 3). This is different from that of Twitter emoji distribution based on http://www.emojitracker.com/, where the most frequent is "face with tears of joy". We cannot generalize this to a difference between Facebook and Twitter emoji use. Our data, being being comments to news articles on public pages, are likely less personal and more evaluative, and hence the frequent uses of "thumbs up". Unlike words in natural language, emoji frequencies do not seem to follow a Zipfian distribution, leading to a hypothesis that the senses of emojis overlap more than that of linguistic words. We found that positive emojis are more frequent than negative ones, which is consistent with previous findings (Novak et al., 2015).

Figure 3: Emoji frequencies in FB comments on media pages
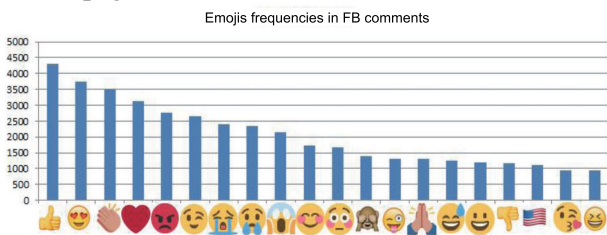


We found that emoji distributions vary across countries (figure 4). All countries use positive emojis more frequently than negative ones. France uses "angry face" frequently, consistent with the fact that the "Angry" reaction is the most frequent in France among the four countries. Both the UK and the US use "crying face"/ "face with tears" frequently. The US also frequently uses the "American flag" emoji, though this may be due

to the data been collected during the US election campaign (August 2016).

Emoji distributions are different in different Reaction profiles (figure 5). The most frequent emojis under "Just likes" are all positive. Under "Amused/Surprised" there are a mixture of positive emojis (e.g. "thumbs up") and negative/neutral emojis (e.g. "astonished face") indicating surprise. Interestingly, while the most frequent emojis under "Sad" all have negative sentiment, a significant amount of emojis under Angry are positive (e.g. thumbs up and clapping hands). This suggests that users often use emojis ironically when their overall attitude is Angry.

Using the sentiment scores compiled for emojis by (Novak et al., 2015), we calculated the average emoji-based sentiment score for each comment containing emojis. We assumed that repetitions of emojis likely express a stronger sentiment than a single use, but the marginal increase of each repetition should gradually diminish, i.e., the use of three "heart shapes" in a row express a stronger sentiment than one "heart shape", but not three times as much. Taking this into account, we calculated the emoji-based sentiment of comments using:

$$\frac{1}{n} \sum_{i=1}^{n} log((number\ of\ emoji_i) + 1) \times sentiment\ of\ emoji_i$$

$$n = the\ total\ number\ of\ distinct\ emojis$$

For example, if a comment contains three "faces with tears of joy" and one "winking face", the emoji-based sentiment of the comment would be calculated as ((log(3)+1)*0.22 + (log(1)+1)*0.46)/2 = 0.46. Then, we calculated the emoji-based sentiment score of each post by averaging the emoji-based sentiment scores of all comments for this post. The scores for the four Reaction profiles of posts are: 0.41 for "Just likes", 0.34 for "Amused/Surprised", 0.24 for "Angry" and 0.24 for "Sad". Though the scores for the first two profiles are higher (and therefore more positive) than the last two, the difference is small,

and the sentiment scores for "Angry" and "Sad" are still positive when negative values should be expected. We think two factors may have dampened the difference between scores of the positive profiles and the negative profiles. First, like we mentioned before, the emoji sentiment scores from (Novak et al., 2015) were calculated from the sentiment of entire tweets. If positive emojis are sometimes used in negative contexts (ironically or for politeness), this method would render lower scores of *positive* emojis than what what would have been the scores when they are used "literally" (directly reflect emotions). Second, we saw that positive emojis are often used in overall negative profiles, e.g. "clapping hands" are frequently used in the "Angry" Reaction profile. These are contexts where many of the emojis are used ironically, or are used for marking that the accompanying text is ironic. Therefore, using the sentiment scores of these positive emojis to calculate the sentiment of the entire comments will lead to misleading results. This further demonstrates the fact that emojis and linguistic texts can modify the meaning of each other, and it is important to study how the meaning interplay works.

Figure 5: FB Emoji distribution by Reaction profile
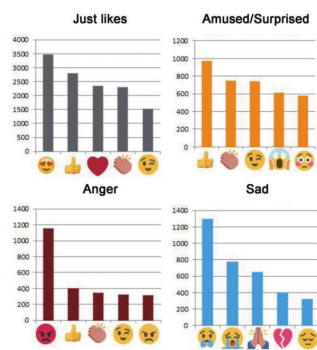


## 6 Future studies

Our next step is to model contexts of emojis, distinguishing those where emojis directly reflect emotions, and those where the meaning of emojis are modified. In addition, to know whether any of the the cross-country differences we found are indeed due to cultural factors rather than due to the major events happening in each country, we need to obtain data over a longer period of time from a wider varieties of FB pages.

## 7 Conclusions

Our results show that there is a reliable correlation between Facebook reactions and emoji usages, suggesting that emojis *can* be used to detect users sentiment, if we take into account of contexts where their meanings are modified (used ironically or for politeness). This study also demonstrates that Facebook reactions and comments are a good data source for investigating indicators of user emotional attitudes.

## References

Naomi Baron. 2009. The myth of impoverished signal: Dispelling the spoken-language fallacy for emoticons in online communication. In Jane Vincent and Leopoldina Fortunati, editors, *Electronic emotions: The mediation of emotion via information and comunication technologies*, pages 107–135. Peter Lang, Oxford.

Marina Boia, Boi Faltings, Claudiu-Cristian Musat, and Pearl Pu. 2013. A:) is worth a thousand words: How people attach sentiment to emoticons and words in tweets. In *Social Computing (SocialCom), 2013 International Conference on*, pages 345–350. IEEE.

Thomas Dimson. 2015. Emojineering: Machine learning for emoji trends by instagram. http://instagram-engineering.tumblr.com/post/117889701472/emojineering-part-1-machine-learning-for-emoji. Accessed: 2016-06-12.

Eli Dresner and Susan C. Herring. 2010. Functions of the nonverbal in cmc: Emoticons and illocutionary force. *Communication theory*, 20(3):249–268.

Scott E. Fahlman. 2012. Smiley lore:-. *Scott E. Fahlmans Homepage*.

Alexander Hogenboom, Daniella Bal, Flavius Frasincar, Malissa Bal, Franciska de Jong, and Uzay Kaymak. 2013. Exploiting emoticons in sentiment analysis. In *Proceedings of the 28th Annual ACM Symposium on Applied Computing*, pages 703–710. ACM.

Xia Hu, Jiliang Tang, Huiji Gao, and Huan Liu. 2013. Unsupervised sentiment analysis with emotional signals. In *Proceedings of the 22nd international conference on World Wide Web*, pages 607–618. International World Wide Web Conferences Steering Committee.

David A. Huffaker and Sandra L. Calvert. 2005. Gender, identity, and language use in teenage blogs. *Journal of Computer-Mediated Communication*, 10(2):00–00.

Ryan Kelly and Leon Watts. 2015. Characterising the inventive appropriation of emoji as relationally meaningful in mediated close personal relationships. *Experiences of Technology Appropriation: Unanticipated Users, Usage, Circumstances, and Design*.

Chiara Mazzocconi, Ye Tian, and Jonathan Ginzburg. 2016. Towards a multi-layered analysis of laughter. In *Proceedings of JerSem, the 19th Workshop on the Semantics and Pragmatics of Dialogue*.

Petra Kralj Novak, Jasmina Smailović, Borut Sluban, and Igor Mozetič. 2015. Sentiment of emojis. *PloS one*, 10(12):e0144296.

Satomi Sugiyama. 2015. Kawaii meiru and maroyaka neko: Mobile emoji for relationship maintenance and aesthetic expressions among japanese teens. *First Monday*, 20(10).

Mike Thelwall, Kevan Buckley, Georgios Paltoglou, Di Cai, and Arvid Kappas. 2010. Sentiment strength detection in short informal text. *Journal of the American Society for Information Science and Technology*, 61(12):2544–2558.

Alecia Wolf. 2000. Emotional expression online: Gender differences in emoticon use. *CyberPsychology & Behavior*, 3(5):827–833.

Jichang Zhao, Li Dong, Junjie Wu, and Ke Xu. 2012. Moodlens: an emoticon-based sentiment analysis system for chinese tweets. In *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 1528–1531. ACM.