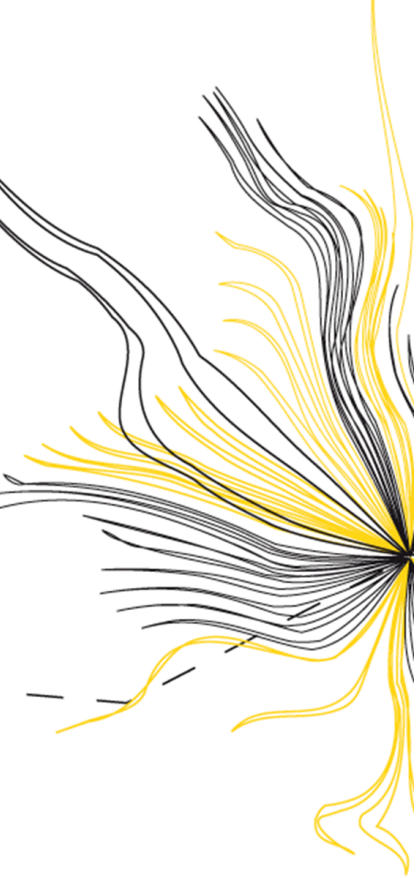


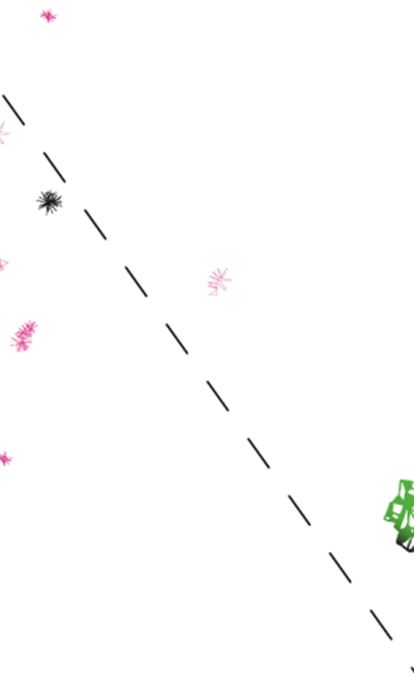
UNIVERSITY OF TWENTE.

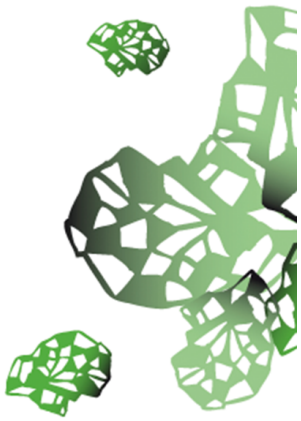
Faculty of Electrical Engineering, Mathematics & Computer Science

Emoji Recommendation for Text Messaging 💬🧐🔍💡

Lilian Sung

M.Sc. Thesis Interaction Technology
February 2021





Thesis Committee:
dr. Mariët Theune
dr. Lorenzo Gatti
prof. dr. Dirk Heylen
dr. ir. Wouter Eggink

Human Media Interaction
Faculty of Electrical Engineering,
Mathematics and Computer Science
University of Twente
P.O. Box 217
7500 AE Enschede
The Netherlands

Abstract

Emojis are used in Computer Mediated Communication (CMC) as a way to express paralinguistics otherwise missing from text, such as facial expressions or gestures. However, finding an emoji on the ever expanding emoji list is a linear search problem and most users end up using a small subset of emojis that are near the top of the emoji list. Current solutions such as the search bar, or simple recommendations still requires effort from the user or does not offer a wide range of emojis. In order to understand how people use emojis, a literature review was carried out for articles that categorise emoji functions. From these, 6 functions were mentioned repeatedly: emphasis, illocutionary, social, content, aesthetic, and reaction. Illocutionary and social emojis make up the bulk of emojis that accompany text.

Two main emoji recommendation models were built. One which recommends emojis similar in meaning to the text input (Related model), and another which recommends only the most common emojis (Most Used). The outputs of the two models were combined to form a third model (Combined). A between-within subjects text-based experiment was carried out over Discord. Participants' emoji user behaviour was compared between a without recommender and a with recommender condition (within subjects). Furthermore, the three models were tested against each other in the with recommender condition (between subjects).

The Related and Combined model were perceived well, while the Most Used did not always recommend appropriate emojis. Participants did use more emojis as well as a larger variety of emojis when an emoji recommender is present, however, this may be largely due to the design of the experiment. When a recommender is included on the phone emoji keyboard, the effect may be much smaller.

Acknowledgements

This thesis marks the end of my studies at the University of Twente. I would like take this opportunity to give a big thank you to my supervisors Mariët Theune and Lorenzo Gatti. Thank you for your patience and guidance throughout my long thesis process. I did not foresee myself writing an emoji thesis, and yet here we are. It makes me very happy that to have been in an environment where I could explore my own curiosities.

I would like to thank all the participants for their help, and all my friends who recruited their friends. I would also like to thank Sujay Khandekar for helping me figure out how to insert emojis with LaTeX. Finally, I would like to thank all my friends who supported me along the way, who listened to me rant about how Python and LaTeX don't like emojis, and bounced ideas with me back and forth.

Lilian Sung 🦫🦫
Enschede, February 2021

Contents

Chapter 1

Introduction

1.1 Background and Motivation

During face-to-face conversations, *how* something is said (the paralanguage) is just as important as *what* is being said. The paralanguage of speech includes aspects that can be vocalised such as intonation and volume, as well as aspects that are visible, such as facial expressions, gestures, and body language. During text-based communication, punctuation is traditionally used to mark much of how the text should be read. For instance, commas add pauses, exclamation marks convey emphasis and perhaps an increase in volume. With the rise of Computer Mediated Communication (CMC), such as emails and instant messaging, users have developed other ways to convey affect. The case, such as ALL CAPS, or aLTerNaTINg CAsE may be used to indicate shouting and exasperation respectively. Emoticons (short for "emotion icon"), which are emoting faces made from text, such as :) or >:(are also powerful tools in marking the mood of text. More recently, emojis have become a staple in instant messaging used to add flair as well as emotions to text.

Emoji, from the combination of the Japanese characters 絵 (e, meaning picture) and 文字 (moji, meaning character), are pictographs and ideographs that can be inserted into electronic text. Emojis take the spirit of emoticons and add to it colour and detail. Currently there are more than 3,000 unique emojis spanning smileys 😀🙄☹️😌, food 🍕🥐🧁🍅, nature 🌸🌲🍁, objects 🚺👠⚽🎷, symbols ❤️♻️✴️, and flags 🇪🇺🇳🇱. Each year more emojis are added[1], expanding the possible emoji vocabulary. However, as the number of emojis increases, the process of finding and inserting emojis become increasingly difficult for the user as this is a linear search task (Pohl, Stanke, & Rohs, 2016). New additions may go unused due to people not knowing about their existence. Emojis add nuance to a person's text, broadening someone's emoji vocabulary can deepen their potential for expression, similar to learning new words.

Some alternatives and additions to the current emoji keyboard have been explored. For instance, a zooming keyboard (Pohl et al., 2016), a gesture based insertion method (Alvina, Qu, Mcgrenere, & Mackay, 2019), and a facial expression emoji filter system (El Ali, Wallbaum, Wasmann, Heuten, & Boll, 2017). The original idea for this thesis was to design a novel method of emoji insertion inspired by affective language and metaphors. Perhaps the emoji keyboard could be categorised by affect (e.g. "happy", "angry", etc.) first? Or perhaps emojis could be explored based on related concepts (i.e. when you click on an emoji, it shows that emoji as well as related ones)? While designing a new method of emoji insertion is exciting, it was rather difficult to come up with potential concepts that were applicable to multiple emoji occasions. For example, not all emojis could be categorised into an emotion (what emotion would a canoe emoji 🛶 fall under?). Ultimately, I decided to look at ways to improve the current emoji keyboard instead of designing something completely new.

Current emoji insertion can be broken down into the following steps:

[1] 230 in 2019, 117 in 2020. Burge (2019a, 2020a)

1. Type text portion of message[2]
2. 'Emoji Moment', i.e. want for inserting an emoji, often with the desired emoji in mind
3. Switch to emoji keyboard
4. Find the emoji
5. Insert the emoji

The question, then, is how to decrease the amount of work the user has to do between their Emoji Moment, and emoji insertion. Some features have already been implemented by various keyboards, namely a "recently used" tab or section at the top of the emoji list, a search function where users can use keywords to find emojis, and sometimes a few recommended emojis that appear in the autocorrect or next-word recommendation space of the regular text keyboard. If this last recommendation feature worked perfectly, this would mean the user does not have to switch to the emoji keyboard at all. However, only a limited number of emojis can be shown in this space. Thus, for the current thesis, emoji recommendation was explored.

1.2 Research question and Approach

The goal of the thesis is to investigate emoji recommendation and how recommendations can impact user's emoji behaviour during text messaging. Formulated as a research questions it is **What makes a successful emoji recommender, and how can emoji recommendations influence users' emoji behaviour?**

First a literature review was carried out in order to understand how people use emojis, what different functions emojis serve within text messaging, and previous work regarding emoji prediction/recommendation. This provides some guidelines for the requirements of an emoji recommender as well as ideas for how to approach building a recommender.

From here two main recommender models were built, one which recommends a broad range of emojis related to the text input (Related model), and another which recommends only the most used emojis (Most Used model). A third model was also built that combines the results of the previous two (Combined model).

The three models were evaluated in an experiment where participants had two short text-based conversations with a simple chatbot. The first conversation was without recommender, while in the second recommendations were added to the participant's text messages. This allowed for between-subjects comparison of the three recommender models, as well as within-subjects comparison of the effects of an emoji recommender.

1.3 Structure of the Thesis

The thesis is organised as follows:

Chapter 2 (Literature review) provides further background information into emojis and their use, as well as outlining past work on emoji prediction and recommendation. The results of this chapter motivates the decisions in the next.

Chapter 3 (Building the Recommender) covers the creation of the recommender models. A survey was conducted to get an understanding of emoji variability between users as well as to collect an independent test set of text messages with emojis that can be used for offline evaluation of the models. In the end, three recommenders were made. The first is a model that recommends emojis related to the text input (Related). This model is based on emoji and word vectors as well as emoji senses. The second model recommends the most used emojis (Most Used). This model is trained on Twitter GIF data. The third model is the combination of two prior models (Combined), recommending both related as well as most used emojis.

[2] It is possible to have emoji moments without accompanying text (e.g. A: "I'll be there in 20 minutes" B: "🦁").

Chapter 4 (Evaluating the Recommender) presents the methodology, results, and discussion of the user evaluation. A Discord-based experiment was designed to compare emoji behaviour without and with an emoji recommender (within subjects), as well as compare emoji use between the three models (between subjects).

Chapter 5 (Conclusion and Future Works) summarises the work done throughout the thesis. Directions for future work are also presented.

Chapter 2

Literature Review

Emojis are interesting to study because they are relatively new in the timeline of human communication. Emojis are being used more often, slowly replacing their predecessor the emoticon (Pavalanathan & Eisenstein, 2015), but what exact communicative niche do they fill that existing text paralinguistics such as punctuation do not? In this chapter a more detailed background into the history and workings of emojis will be given. In addition, research spanning various disciplines, from communication, to linguistics, to psychology will be outlined. An analysis was carried out to summarise emoji papers that categorise their function in communication. From these a list of functions was formed which was used to motivate the priorities of the recommender. Emoji prediction and recommendation has also been touched on in the field of computer science, this will be outlined here too, giving rise to some of the approaches for building a recommender in the next chapter.

2.1 Makings of an Emoji

Emojis first appeared in the late 90's and were officially adopted into Unicode in 2010 (Burge, 2019b; Pardes, 2018)[1]. The Unicode Consortium maintains the Unicode standard for how written text is encoded. Each grapheme (one unit of writing, such as 'a', or '🌭') has a corresponding code point, for example, the code point for 'a' is U+0061, while '🌭' is U+1F32D. Although Unicode gives suggestions for how each grapheme can be rendered, this is not enforced. As such, each provider is able to implement their own emoji renderings, this is akin to the font of the emoji.[2]

Throughout the past decade, different providers have homogenised the emoji designs to a certain extent, though there is still room for creativity and style. It is possible that depending on which device two individuals are using, the emojis one person sends are rendered vastly different on the receiver's phone, causing miscommunication (Franco & Fugate, 2020; Miller et al., 2016).[3] For instance, in earlier years, the face screaming in fear emoji 😱 varied quite a bit across platforms in levels of shock (see Figure 2.1). In 2016, the Samsung emoji had a design where the emoji was so scared, its soul left their body. If this was the intention of the sender, but the receiver received the Google version, miscommunication was highly likely.

It was mentioned in the introduction that there are currently more than 3,000 emojis. The bulk of the current set of emojis, however, is made up of sequences of emojis joined together

[1]Very often, Shigetaka Kurita of Japanese phone carrier Docomo is cited as the creator or father of emoji in 1999. However, SoftBank actually released their emoji set in 1997!

[2]During this writing, a mixture of different renderings is used. JoyPixels https://www.joypixels.com/ is used throughout the text and tables, Windows emojis are found in screenshots from chrome, phone screenshots are either Google or Samsung depending on the keyboard used, finally, Discord screenshots feature Twitter emojis (Twemoji).

[3]Since emojis are constantly evolving, older papers may be basing their research on earlier emoji renderings.

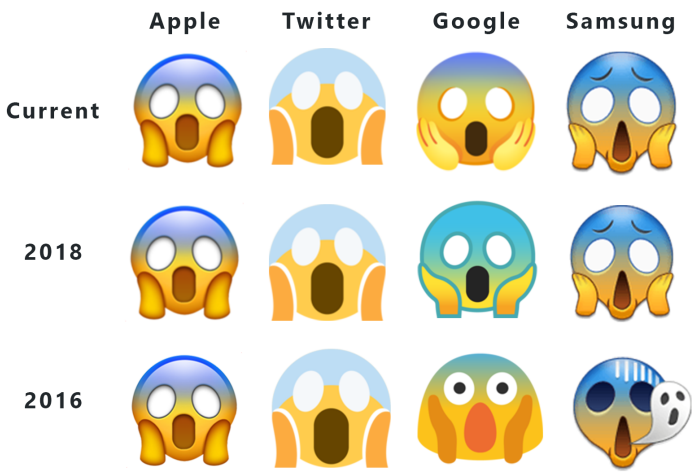


Figure 2.1: Current, 2018, and 2016 versions of face screaming in fear emoji for Apple, Twitter, Google, and Samsung. Apple and Twitter have remained largely the same, while Google and Samsung have made alterations. The current versions are much more similar across platforms compared to the 2016 versions.

by the Zero Width Joiner (ZWJ). For instance, the 'Family: Woman, Woman, Girl, Boy' emoji 👩‍👩‍👧‍👦, is actually made up of four people emojis 👩👩👧👦 , combined using the ZWJ character. Skin tone and gender modifiers are most common in emoji ZWJ sequences, although Unicode is introducing more non-human emoji ZWJ sequences such as the service dog 🦮 (🐕+🦺), polar bear 🐻‍❄️ (🐻+❄️), and black cat 🐈‍⬛ (🐈+⬛). ZWJ sequences allow for more emojis to be created without constantly creating code points. This comes into effect later in Chapter 3, when emoji vectors are concerned.

2.2 Importance of Emoji in Communication

Some of the more well known emojis are often those depicting exaggerated faces such as smiling face with heart eyes 😍 or the loudly crying face 😭. Facial expression is one of the main ways to communicate emotions during face-to-face communication, as such, it is easy to assume that face emojis are also used for emotion expression. However, there are many occasions where a person's facial expression does not match how they are feeling. For instance, in a scary situation, one might smile at their child to reassure them; in a restaurant, the waiter might smile to be polite. Similarly for emojis, there are more nuanced functions other than emotional expression.

2.2.1 Emojis and Emotions

Studies that link emojis to emotions tend to follow one of the two main groups of emotion theory: discrete or dimensional. Discrete models suggest that there are a finite number of emotions which are basic/core/universal to every human being (e.g. (Ekman, 1999)). Jaeger and Ares (2017) conducted a survey investigating how people attribute emotions to emojis. They found that some emojis are strongly associated with one emotion, for example 😠 with anger or 😘 with love. There were also emojis that are associated with multiple emotions, for example 😐 with neutral, not caring, no comments. Some emojis are associated with similar emotions, for example 😢😖😩 were all in the group that was associated with sad, depressed, disappointed, and frustrated. However, emojis in this group associated with the different emotions at different strengths, pointing to a level of nuance in their meaning and use. The other group of emotion theories suggest that emotions can be defined as existing on a point on one or more dimensions.

The most commonly used dimensions are valence (positive or negative) and arousal (the energy level) (Russell, 1980). Emoji are rated consistently on valence (Jaeger, Roigard, Jin, Vidal, & Ares, 2019; Novak, Smailović, Sluban, & Mozetič, 2015) and arousal (Jaeger et al., 2019). While much emoji research focus on the most used emojis, which consists mainly of smileys and common emotive symbols such as the heart ❤️ or fire 🔥, in the study of Novak et al. (2015), many non-face emoji were included which are also consistently rated on valence.

More recently, Barrett (2017) proposed the Constructionist theory of Emotion which theorises that there are no naturally occurring categories for emotions, and that the emotional words used are just what we conceptually use to define emotions. A helpful analogy is the words we use to describe colours: the western concept of "red" has no exact cut off wavelength in the electromagnetic spectrum, and each person may associated redness with different things. For instance, I associate it with apples, and Christmas, and Lightning McQueen from Pixar's Cars. Each time something affective (positive or negative) is experienced, we attribute an emotion with it, and thus this becomes an instance of the emotion. The connections between emotional words and emotions are subjective to each person, likewise, the connections between emojis and their meaning is subjective. With that said, there are some trends in how emojis are interpreted. Riordan (2017) conducted a study where ambiguous texts (″Got a shot″or ″Got a ticket″) were either on their own, or followed with a disambiguating emoji ([🏀, 🍹, 💉] for different types of shot or [🚔, 🎥, 🛫] for different tickets). The results showed that the emoji did manage to disambiguate the text. Additionally, although the message did not explicitly contain affective information, the affective connotations with each type of ′shot′or ′ticket′carried over and had an effect on the valence evaluation of the messages. For instance, "Got a shot 🏀" was rated positively while "Got a shot 💉" was rated negatively.

2.2.2 Emojis for additional information

Emojis are used to strengthen the emotive value of text, although most of the time they serve more subtle functions. Emojis often play the part of the paralinguistics of CMC. For example, in ″How are you? 🙂″, the slightly smiling face emoji 🙂 isn′t necessarily used to convey that the sender is happy, but a politeness/friendliness marker. Vidal, Ares, and Jaeger (2016) found that emojis are mostly used to convey information not expressed in words, and that emojis used to emphasise information expressed in words are much less common. Dresner and Herring (2010) analysed how emoticons (not emojis) were used to alter or indicate illocutionary force in CMC. Emoticons could express emotions that would usually be expressed though facial expressions (e.g. smiling for happy, raised eyebrows for surprise), convey non-emotional meaning expressed through facial expressions (e.g. winking to indicate joking), and convey illocutionary force (e.g. smiling to soften a demand). Herring and Dainas (2017) extended this research to include emojis as well as other ′graphicons′(GIFs, stickers, image, video) on Facebook discourse.

Holtgraves and Robinson (2020) found that emojis can be used to convey indirect meaning, particularly for instances where the indirect meaning is negative. For example, if someone asked: "What did you think of my presentation?", the reply: "😬 It's hard to give a good presentation" is more easily interpreted as negative only when the emoji is present (or if the emoji was given as the reply alone). Similarly, Rodrigues, Lopes, Prada, Thompson, and Garrido (2017) found that emojis can soften negative messages and increase perceived positivity, but only when the conversation is considered to be jokey or non-serious. In a serious conversation, using emojis along with a negative message may signal a lack of interest for the conversation and the sender may be perceived more negatively.

The meaning of emojis is flexible and can differ depending on situational or social context (Wiseman & Gould, 2018). For example, the rose emoji 🌹 may refer to the flower in one conversation, and to a person named Rose in another. Culturally, emojis have also taken on more than their literal meaning. For example, the eggplant 🍆 is often used to refer to the penis, and the trophy 🏆 to refer to the feelings of winning/being a champion and not the literal

tournament cup.

2.3 Emoji Categories of Use

Emojis reside on the continuum between language and nonlanguage. Sometimes they are used to replace words e.g. "I'm going 🏊 later, want to join?" (swimming), while other times they are used to convey information that in face-to-face communication would be expressed through facial expression or bodily language e.g. "😳😳😳" (in reply to a surprising message). There are a number of works analysing the functions of emoji, the categories of which vary depending on through what lens the researchers view emoji, as well as the type of data used for analysis. The aim of this section is to summarise existing literature and consolidate the results of different studies into one list that can guide the development of the emoji recommender. A summary of each study's categories can be seen in Table 2.1. The table also shows any counts of messages or emojis belonging to each category if available in the paper. Very often one emoji can serve multiple functions so the numbers do not necessarily add up to the number of messages/emojis in the dataset.

Gawne and McCulloch (2019) and Danesi (2016) both based their analysis of emojis on previous theories. Gawne and McCulloch discussed similarities between emoji use (mainly on Twitter) and gestures during face-to-face communication using McNeill's (1992) classification of gestures. Danesi (2016), on the other hand, analysed emojis based on Jakobson's (1960) theory on the functions of language. Within Danesi's writing, however, it is not very clear how emojis map onto each function. The definitions used for each function also seem to differ from other writings which utilise Jakobson's theory (e.g. Ismaeil, Balalau, and Mirza (2019)). For example, Danesi notes that the conative function includes "emoji with strong emotional content" (p. 103), while other interpretations of the conative function focus on language that requests, demands, or advices the addressee (Ismaeil et al., 2019).

Na′aman, Provenza, and Montoya (2017) sorted emojis into three categories: 1) function word stand-in (e.g. "I 🍩 like you"), 2) lexical word stand-in (e.g. "The 🔑 to success is 🍕"), and 3) multimodal (e.g. "Omg why is my mom screaming so early 🙄"). Their categories arose from "observation", though no concrete source was cited. The goal of their paper was to see if it is possible to train a model to automatically categorise emojis into these functions. Na′aman et al.'s 'multimodal' category seems to cover all emoji uses that are not a stand-in. The multimodal category was further broken down into four subtypes [attitude, topic, gesture, other].

The remaining three papers used a bottom-up approach where functions were derived from the analysis of collected data. Al Rashdi (2018) analysed group conversations with the same participants over a period of time, while Cramer, de Juan, and Tetreault (2016) analysed text messages collected through a survey. In addition to the text messages, Cramer et al. (2016) also asked survey participants to explain the meaning and context of the emojis. The functions were broken into two main groups: Sender intended, and Linguistic. Sender intended functions were broken down further into a) emojis that added additional information, 2) emojis that changed the tone of the text, 3) and emojis used for engagement and relationship maintenance. The linguistic functions were broken down into 1) emojis that repeated the text, 2) emojis that complemented text, and 3) emojis that replaced text.

Dainas and Herring's (2019) categories are largely based on a previous study in 2017 where comments from public Facebook groups were analysed (Herring & Dainas, 2017). The original study investigated the pragmatic functions of graphicons (emojis, emoticons, GIFs, images, stickers, and videos), which resulted in the following functions: tone modification, reaction, action, mention, riff, sequence, ambiguous. The original list of functions was modified when only considering emojis. Softening was added in addition to tone modification, decoration and physical action were added, and riff (joke/banter) was removed.

Table 2.1: Summary of papers categorising emoji functions

Authors	Data	Emoji Functions
Al Rashdi (2018)	WhatsApp messages from two group chats	1. Indicating emotions 2. Contextualization cues 3. Indicating celebration 4. Indicating approval 5. Response to thanking and compliments 6. Signalling openings and closings of conversations 7. As Linking device 8. As indicators of fulfilling a requested task
Cramer et al. (2016)	228 text messages collected through Mechanical Turk. 146 unique emojis, 480 emojis in total.	**Sender intended functions:** 1. Additional information (195) (a) Expressing emotion (139) (b) Situational context (56) 2. Changing tone (26) 3. Engagement and Relationship (20) (a) Engaging the recipient through novelty or flair (7) (b) Tool to adhere to social and conversational norms (8) (c) Relationship maintenance through e.g. shared tradition (5) **Linguistic functions:** 1. Repetition of text (40) 2. Complementary usage (155) 3. Text replacement (45)
Danesi (2016)	323 text messages provided by university students.	1. Emotive: conveys the intent, attitude, or mood of addresser (589) 2. Conative: produces an effect on the addressee (512) 3. Referential: refers to context of communication, often informative (456) 4. Phatic: establishes, maintains, or discontinues communication (412) 5. Poetic: draws attention to the form of the message (134) 6. Metalingual: reference to the 'code' (0)
Dainas and Herring (2019)	Analysis of Facebook comments	1. Tone modification 2. Softening 3. Reaction 4. (Virtual) Action 5. Mention 6. Physical expression (user actually carrying out action) 7. Decoration

Gawne and McCulloch (2019)	Observation of emoji use in online communication, mostly Twitter	1. Illocutionary: the intention of the speaker in saying a particular utterance 2. Illustrative: refer to concrete objects 3. Backchanelling: response of someone listening to the speaker 4. Metaphoric: refer to abstract concepts 5. Pointing: gesture that draws attention to something 6. Beat: repetitive; useful for adding rhetorical emphasis
Na′aman et al. (2017)	567 tweets; 878 emojis; 775 emoji spans. A span may include multiple emoji used in sequence.	1. Stand-in for function word (38) 2. Stand-in for lexical words or phrases (51) 3. Multimodal (686); enrich grammatically-complete text with markers of affect or stance

2.3.1 Emoji Functions in Text Messaging

Trends can be gathered from the various studies. Some functions appear repeatedly between papers, albeit under different names. Due to perhaps the variance in approach and data source, a category from one paper may be broken down into subcategories or did not appear in another. Data from Facebook or Twitter or even single text messages may not encompass all types of interactions emojis are used in (single emoji reactions, for example, would be excluded from collected single text messages that target emojis used with text). The methodology used here for consolidating the existing research is to first group similar functions together, then restructure the list so that it is concise and useful for the current project. The resulting list will be covered below, a summary of which can be seen in Table 2.2.

Emphasis Emojis

All the papers include a category for emojis that has to do with emotional information. This is also the way emojis are used most often (Cramer et al., 2016; Danesi, 2016; Na′aman et al., 2017). However, there is a difference in emojis that are congruent with the sentiment of the text alone ("I am so happy right now 😆"), and emojis that mark the intention of the speaker ("I am so happy right now 🙃🔫🍗"). Emphasis emojis only refer to the first case where the emotion of the emojis matches the text; they strengthen the emotions of the whole message. The second case where the emojis suggest an opposite meaning from the text, indicating the actual intention of the text, are illocutionary emojis instead (described next).

Gawne and McCulloch pointed out that repetition of the same emoji e.g. 😂😂😂 or emojis of the same theme e.g. 🍟🍔🌭🍕 are used to supply emphasis to either the emotions or topic of text messages, much like beat gestures. Thus, emphasis emojis also include those that repeat the non-affective information in the text, often nouns or verbs ("aaah I love spring 🌸🌿🌸🌿🥰").

Illocutionary Emojis

Dresner and Herring (2010) analysed the function of emoticons (:), :p, :(, >:(, etc.) under the framework of speech act theory (Austin, 1962). A speech act has three levels: 1) the locutionary act, or the apparent meaning, 2) the illocutionary act, the underlying intention of the sender, and 3) the perlocutionary act, the actual effect of the act which may or may not occur (e.g. the perlocutionary act may be to persuade, but whether the act is successful depends on word choice, mood of the receiver, etc.). Apart from the common use of emoticons

Table 2.2: List of Emoji Functions in Text Messaging

Function	Explanation	Example
Emphasis	Referring to concepts or objects mentioned previously in the text message. Strengthens emotional value.	I definitely want a pet when I move out 🐁🐱🐶🐹
Illocutionary	Clarifying/altering intention of text or adding emotional information otherwise missing.	She wants me to drive her again 🙄💤
Social	Performing social communicative acts such as backchannelling or conversation management (opening/ending the conversation).	👏👏 Heyy! How are you doing??
Content	Adding non-emotional information otherwise missing from the text. May be used to disambiguate the message. Also includes emojis used to spell.	1) Wanna grab 🍩 later? 2) I 🍩 like him!
Aesthetic	Adding decorative elements to the message .	Nice to meet you too 🌸🌸🌼🌼
Reaction	Replying to another person's prompt, usually a stand alone emoji.	A: can you buy eggs too? B: 👌👌

to convey emotions, emoticons are also often used to signify joking, flirting, or sarcasm which are not emotions. Emoticons are also used to indicate or modify the illocutionary force of the text message. In one of their examples "Since I've never worked on this kind of data before, I am writing for some suggestions. :)", it was pointed out that the :) here does not mean the sender is happy, but softens the request.

Similarly, emojis are not only used to code for emotions, they can also be used to alter the illocutionary as well as the locutionary force of the message. As seen in Figure 2.2, sometimes the emojis are crucial to the meaning of a message (high in locutionary; bottom right cluster), where without the emoji the message would not make sense. Other times, the meaning of the message is complete with the text alone, but the emojis enforce or alter the intention of the message, as seen in the top left cluster (high in illocutionary force). Emojis that are important for the meaning (locutionary force) of the message are content emojis (covered later), while emojis that are important for the intention of the message are illocutionary emojis.

Illocutionary emojis' main function is to clarify the emotion or intention of the message. For example, consider the different intentions in "I ran past and ignored him 🤣🤣" and "I ran past and ignored him 😱😠". The first may suggest the sender felt it was a funny event while in the second the sender seems to feel some sort of regret. Illocutionary emojis appear explicitly in Cramer et al. (2016) as "changing tone" and in Dainas and Herring (2019) as "softening". These types of emojis may not appear as often as emotional emphasis emojis, this may be due to subtleties in use that are harder to identify.

Social Emojis

These emojis map roughly onto the phatic function of Jakobson (1960), which mainly refers to 'small talk' or language that is used to keep the conversation pleasant 1960. Danesi (2016) found that emojis under the phatic function can further be classified as utterance opener, utterance ending, and silence avoidance. The first two were also observed by Al Rashdi (2018) as 'signalling openings and closings of conversations', while silence avoidance is also noted by Gawne and McCulloch (2019) as backchannelling. Social emojis are also what Cramer et al. (2016) observed as 'tool to adhere to social and conversational norms' and 'relationship

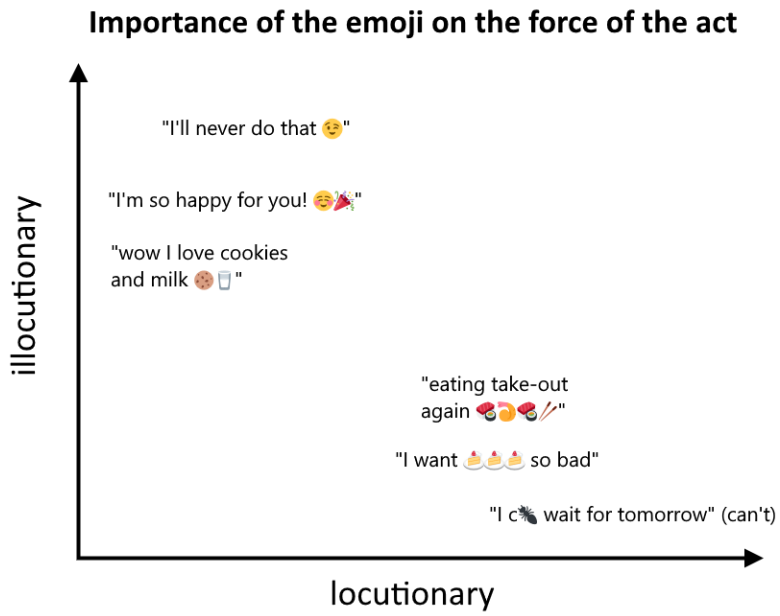


Figure 2.2: Graph showing the various roles of emoji within a speech act. Some emoji have a higher role in the semantic meaning of message while some have a higher impact on the illocutionary force.

maintenance'.

Emojis use to fulfil the social function may be arbitrary and idiosyncratic. Social emojis are also not necessarily accompanied by text. For example, "👋" may be used alone as a conversation opener, while single emoji replies (e.g. A: "guess what I ate today?" B: "😮😮😮😮") may be used as a form of backchanneling.

Content Emojis

Content Emojis add non-emotional information otherwise missing from just reading the text. The three messages in the bottom right cluster in Figure 2.2 are all examples of content emojis. There are two main sub-categories within this group: emojis that represent concepts, and emojis that replace sounds in some sort of visual pun.

Within the first sub-category, the emoji can appear in the middle of a text ("I want 🍰🍰🍰 so bad") or after a complete sentence ("eating take-out again 🍣🍤🍣🥢"). Without the emojis, we would be missing crucial information (cake and sushi/Japanese food respectively).

The second type of content emojis is what Solomon (2020) refers to as 'emoji spelling', where emojis are used to spell out words, for example im🍑ment (impeachment), 🐝 happy (be happy), and 🧢italism (capitalism).

Aesthetic Emojis

Aesthetic emojis are emojis which main function is to add colour to a message. They can be used to make certain words or phrases stand out (e.g. "🌸🌷 Family announcement🌸🌷"). Then can also be used instead of bullet points, or to add decorative "borders" at the top and bottom of a text (more often seen on social media posts than in text messaging). It is perhaps true to say that all emojis are aesthetic emojis, however, their main function may not be for the added appeal of the emoji.

Reaction Emojis

Reaction emojis are quite diverse. They can be used to express emotion, agreement, response to others' message, to name a few. The key feature of reaction emojis is that they are usually the main message and not the supplement. Their meaning is also highly dependent on prior conversation. The OK hand sign emoji 👌 can mean "I agree", or "that's cool", or "I have it under control" depending on context.

2.3.2 Implications for Recommendations

Communicating with emojis is an interpretive process. Meaning is perceived from the way they are rendered on your screen, but there is also the larger connotations made with the concepts emojis try to capture. Emojis can be imbued with metaphors, for example, the heart-eyes emoji 😍 is not a realistic rendition of a facial expression, yet the emoji is readily associated in western society with expressing love for something or someone.

The emoji functions discussed above are not necessarily mutually exclusive to each other. For example, in the short exchange A: "It was good to hear from you again!", B: "🦦🦦", the otter emojis serve the function of reaction, with an implied meaning based on the mutual understanding that otters are positive and they both had a good time. The reply also indicates that B read and acknowledges the message from A, thus acting as a social emoji.

The different emoji functions translates to different challenges and difficulty with regards to recommendation. Emphasis emojis, for example, are relatively straightforward as their content can be readily lifted from the text. Most emojis have a definition that is readily understood; a book emoji 📖 is a book before anything else. Although emojis are context dependent, some emojis are widely regarded as positive (e.g. 😊😚❤️) while some are negative (😫😟😭). Social emojis are also somewhat conventional and can be modelled given enough data (e.g. by learning the emojis that tend to follow 'good morning', for instance).

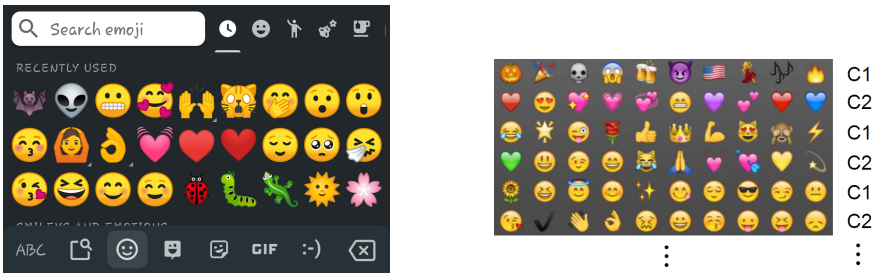
The other emoji functions (illocutionary, content, aesthetic, reaction) are more difficult as they deal with information unavailable from the text. These emojis are used because otherwise the message would be misunderstood or incomplete. Emojis used for these functions depend on the user's thought process and could be learned to a certain extent if enough user data is gained. Any emoji that is used at the start of a message poses a challenge for a recommender unless it has access to the conversational context. Disambiguating emojis used in the middle or end of a sentence are also difficult to predict, but may follow trends (e.g. "that's" could be followed by 🔥 or 💩; there may be certain text-emoji bigrams that are more common).

2.4 Emoji Prediction and Recommendation

From the computer science side, there have been a number of works investigating whether emojis can be predicted given text and sometimes image input. This research largely falls under one of two contexts: research on 'public' social media content such as tweets or research on private messages between people. It is important to keep in mind these contexts as findings might not always be generalisable to both.

2.4.1 Related works

Much of the work in emoji prediction approaches it as a classification problem where sentences using only one emoji out of a short list of up to 20 emojis are used as input, using the sole emoji as the label (Barbieri, Ballesteros, & Saggion, 2017; Liebeskind & Liebeskind, 2019; Xie, Liu, Yan, & Sun, 2016). Barbieri, Marujo, Karuturi, Brendel, and Saggion (2018) expanded the list of emoji to 300 as well as including the time of the year the emoji was used as a feature in their model. Lin, Chao, Wu, and Su (2019) used twitter data containing one to three emoji from



(a) Gboard recently used **(b)** Recommendations from Kim et al.

Figure 2.3: Left: The recently used section on the Gboard. Right: Recommended emoji keyboard layout, showing which cluster the emoji is from on the right, picture taken from Kim et al. (2019)

the top 500 most used emojis. Up to three emojis are returned by their model, additionally, by modelling the emoji as words/phrases, their output also takes the ordering of emoji into account.

The single class classification approach is not quite the same use case as a recommendation system. Although perhaps if more emojis with a high probability were given as an output as in Lin et al. (2019), it could act as a recommendation system too. The following is the list of 20 emojis used in Barbieri et al. (2017): 😂❤️😍🔥💯😊🙌😘🎄💕🎉😭💙✨❄️😎💪🙏👌💋 (the most frequent emojis in their dataset). Looking at this list, the first five overlap in their use, and could very well be recommended for the same tweets or text messages. Guibon, Ochs, and Bellot (2018) used a MultiLabel RandomForest classifier trained on instant messages that included a set of 169 sentiment-related emojis. They used both textual features (bag of words, word count, punctuation, n-grams) and sentiment features (positive/negative/neutrality scores as well as the current mood as selected by the user when sending the instant message). Their model was able to predict the emojis quite well.

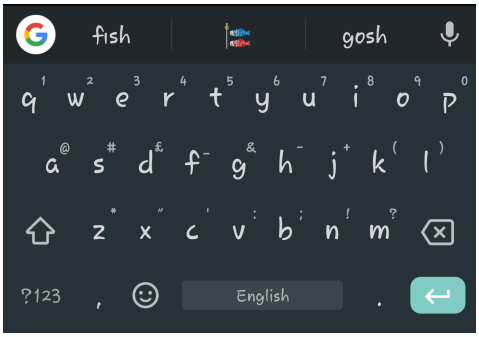
The goal of recommendation is to provide a broad selection of relevant emojis so that the user can pick the ones they like to use. Even if multiple emojis with high probabilities were to be returned, if the model was trained based on user data, the kinds of recommendation would always be limited to those that are widely or commonly used. A meaningful system would include both current ways of using emoji, as well as suggest potential novel ways of using emoji. Figure 2.3a shows the number of emoji that show up in the 'recently used' section of the Gboard emoji keyboard on my phone. With 27 emoji, there is room for recommendations that are more exploratory than usual use cases.

Kim et al. (2019) is the only work so far (to my knowledge) that trained a model to provide a large number of recommendations. Twitter data is used, although they specifically targeted series of replies in order to retrieve conversations. The past five sentences are converted into vectors and given to a Long Short-Term Memory (LSTM) model. The model outputs 'concept' vectors which may represent clusters of emotions/information represented in the conversation. These concept vectors are then further clustered, and each cluster forms a list of emoji closest to it. All the emoji from their list of 111 emoji (it is not very clear how this list is chosen) are attributed to one of the clusters. The clusters' lists of emoji are then displayed in alternating rows as in Figure 2.3b.

2.4.2 Current state of emoji recommendation in keyboards

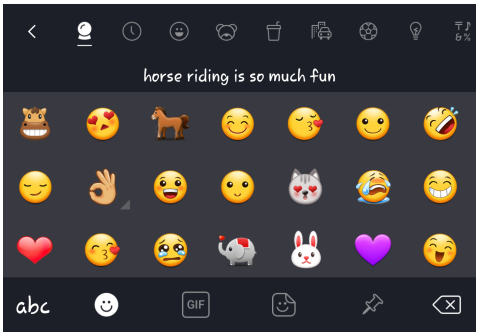
Most of the time emoji recommendations will appear in the same area as the auto-complete bar while typing text using Gboard and Swiftkey (two virtual keyboards I have experience with), these are usually for straightforward nouns. For 'fish', Gboard recommended the carp streamer emoji 🎏 (decoration in Japan for Children's Day; Figure 2.4a) even though it is not a typical fish emoji (i.e. 🐳🐟🐠🐡). For 'horse', Swiftkey recommended the horse face emoji 🐴 over the

horse emoji 🐎 (Figure 2.4b). It is not clear how the recommended emoji is chosen.

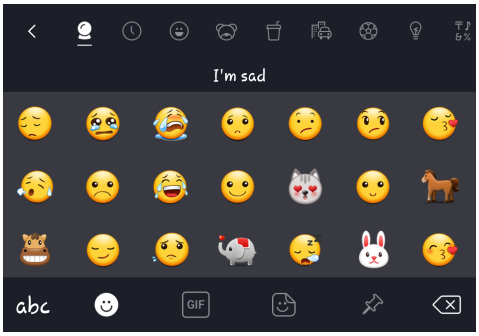


(a) Gboard 'fish'

(b) Swiftkey 'horse'



(c) Swiftkey example 1



(d) Swifkey example 2

Figure 2.4: Top: simple recommendations while typing on Gboard and Swiftkey. Bottom: Swiftkey recommended emoji tab after switching to the emoji keyboard.

Swiftkey also has a recommended/suggested tab when switching to the emoji keyboard that seems to give recommendations based on what has previously been typed. The recommendation for "horse riding is no much fun" (Figure 2.4c) include two horse emoji as well as an array of positive emoji, however, the actual horse riding emoji 🏇 is not included. The recommendations for "I'm sad" (Figure 2.4d) include many sad/negatively charged emoji. Swiftkey also seem to include recently used emoji in the same tab since they do not have a separate 'recently used' tab which other keyboards usually depict with a clock icon. This can be concluded from unrelated emoji such as 🐘🐱 appearing under both sentences.

2.5 Conclusion

Emojis are used in informal text conversations to provide something extra that would otherwise be missing. This can be to simply ✨decorate the text✨ or to drastically alter how the text is read. The latter function, i.e. the illocutionary function, is one of the more common ways emojis are used. Current studies investigating emoji recommendation or prediction have not really taken into account how the different functions or contexts can influence their models' performance. Emojis that turn a relatively positive text message (e.g. "Thanks") into a sarcastic one (e.g. "Thanks 🙄😒") are more difficult to predict than if the emojis were in the same affective space as the message (e.g. "Thanks 😻"). Additionally, it is interesting to look at emojis that people *didn't* use; just because the model did not accurately predict the emoji the user used, does not mean the ones it did predict are necessarily bad.

Designing a useful emoji recommender would have to take into account the difficulty of recommendations falling into the different functions, as well as their importance to the user. Existing emoji prediction and recommendation studies have primarily looked at the most used emojis which are often illocutionary emojis. Illocutionary emojis are used a lot, but require a

certain level of “mind reading” and knowledge of usage norms. On the other hand, emphasis emojis, which refer to concepts already present in the text, are a lot easier to predict and have yet to be investigated. It would be interesting, for the present thesis, to compare how users react to recommenders that focus on illocutionary or emphasis emojis. If emphasis emojis are received well when recommended, this could be an easy way to improve user’s emoji insertion experience.

Chapter 3

Building the Recommender

This chapter outlines the thought and developing process of the emoji recommender models. As stated in the previous chapter, it is difficult to create a recommender that covers all functions of emojis in all contexts. The present models will target emphasis emojis as well as some illocutionary and social emojis. In order to have a better understanding of emoji variability between people, a short survey was carried out. The survey was also used to obtain a small test set for offline evaluations of each model's performance.

Two main models were built, the Related model, which recommends emojis relevant and associated with the content of the text, and the Most Used model, which recommends most commonly used emojis based on the tone of the text sent. The recommendations of the two models were combined to form a third Combined model. The three models will be evaluated in the next chapter.

3.1 Emoji variability survey

There are currently no available corpora of text messages that contain emojis (to my knowledge). In other works which made use of text messages, the messages were collected specifically for the study. Due to privacy concerns for the participants, the data have not been made publicly available. Having a test set of text messages containing emojis allows for the offline evaluation of the models, which will provide some insight into how the models might perform in action.

For this project, text messages were collected for offline evaluation of the recommender using an online survey. The survey consists of two parts, the first asked each participant to copy and paste three messages including emojis they have recently sent. Participants were instructed to avoid messages containing sensitive information or to use placeholders otherwise (i.e. [name], [university], [address], etc instead of the actual information). The text messages collected here forms the test set used in Section 3.6 for offline evaluation. The second part of the survey asked the participants to input emojis they might add to a set of text messages. This second section of the survey gives insight to the variance of how emojis may be used under the same circumstances. Ten messages were taken from my own chats that originally contained emojis (the emojis were removed for the survey), three of which were randomly selected for each participant.

32 people filled in the survey with an average age of 24.272 (SD=2.597). In the first part of the survey, 78 text messages were collected in total after removing messages that were not in English or didn't use any emojis. On average, people used 1.87 emojis per text, with a slightly lower number of unique emojis at 1.55 due to repeated emojis within one text (e.g. "Thank you so much 💕💕💕"). Overall, 119 unique emojis were used with the most commonly used emojis being: 😂 (11), 😅 (7), 🙃❤️ (4), and 💕🙈😏😍 (3).

For the second part of the survey. each of the 10 text messages received between 12 and 22 answers, the results can be seen in Table 3.1. Some of the emojis were chosen more than

once. For example, for the one about ghost stories, the ghost emoji 👻 was used in 13 of the 22 messages. As can be seen, there is a high variability in the emojis used for any situation, suggesting that a large number of emojis are suitable for the same situation. For an initial recommender, it may be a good idea for there to be a large number of recommendations so that the user can choose the one(s) they want to use. Over time, it is possible for a recommender to learn each user's preference, leading to a more precise and concise list.

Table 3.1: Emoji responses for the second part of the survey. The number of responses (N) and mean number of emojis used per response (M) are shown in parentheses respectively.

Text Message	Unique emojis
Hello peeps, I'd like to be captain if no one objects? (N = 15, M = 1.53)	😊🧑🚿👻🖐️😀😁😀🤠🤓😎👩‍💼🙋‍♂️😎😉😊
Also, would've been really good ghost story timesss (N = 22, M = 1.68)	👍👌🤡👺😆📒👂⌚😄😕😮😜😰🤩👀🙉😂😉👻
Did you have to touch cursed paper currency? (N = 17, M = 2.35)	😅😈👺💵💶💷🧛‍♀️😍😐😒😟🤔😢👉👈😡💵💀🤑🤢🤮🙄😰😱💸💶
I want to go home but it's raining and I don't want to bike through the rain (N = 14, M = 1.43)	🌱😅😐😐😑😖😭😩😔😥😫🙃🙄😒🤗☔🌧️
Enjoy the concert!! (N = 15, M = 1.80)	✨💕🎉🎸🎻🕺🏌️‍♂️😁🎶🎵🎵👨🎤🙃🤗🤝😘🥰🥳😊
I only have good ideas (N = 13, M = 1.46)	👨🔬🤡🙅‍♂️😊😉😎🙄💯💭🤗🤪🤓😎
ye but not sure what to do otherwise (N = 19, M = 1.16)	😁😑😌😑😒🙃😅😕🤔🤷
yiss i wait for vegetables (N = 12, M = 3.0)	🍆👍😊😍🥳🕐🕑🕒🕓🕔🙂👈🙄😳☺️🌶️😁🥕🍟🥦🥒🍅🍆🫑🥬
Christmas songs are the only thing keeping me above water through November hahah (N = 17, M = 1.88)	❄️🎅🎷💪🌝😭😎🤩🎁😀✨😩⛄🎶🎄🎄😅😂
lmk when you leave campus (N = 16, M = 1.06)	😊✌️🦮👍👣🎻💕💪😁😁😅😇😊😎🙂🚲🚶

3.2 Designing an emoji Recommender

The second part of the survey shed some light on the variability of emoji use. The ideal emoji recommender is one that provides emojis the user wants to use, as well as introducing novel emojis that they are less familiar with. Novel recommendations might encourage users to use a broader range of emojis, increasing their vocabulary as well as potentially increasing the complexity of emoji use. The whole set of recommendations should be around 25 emojis. This number is approximately how many emojis can fit on one 'page' of an emoji keyboard on a phone. Figure 3.1 shows the Google keyboard's 'Recently used' section on their emoji keyboard which spans three rows of nine emoji for a total of 27.

As the analysis of different emoji functions from the previous chapter shows, some emojis are almost impossible to predict without being able to read the user's mind (e.g. the doughnut emoji 🍩 in "Wanna grab 🍩 later?"). On the other end of the spectrum, emphasis and social emojis follow conventions and are easier to predict given the preceding sentence (e.g. the smiling cat face with heart-shaped eyes emoji 😻 in "Your cat is so cuteee I love her! 😻😻😻"). For this project, the recommender should be able to cover emphasis as well as social emojis.

Table 3.2: Preprocessing outcome.

Text Message	Tokens
Hello peeps, I'd like to be captain if no one objects?	['hello', 'peeps', 'like', 'captain', 'one', 'objects']
Did you have to touch cursed paper currency?	['touch', 'cursed', 'paper', 'currency']
I want to go home but it's raining and I don't want to bike through the rain	['want', 'go', 'home', 'raining', 'bike', 'rain']

In messages including emojis, emojis are most often placed at the end of a message (Al Rashdi, 2015; Cramer et al., 2016). Emojis are also often used alone as reactions (Al Rashdi, 2015). The recommender for this project will be for end-of-sentence emojis only. This constrains the type of input the recommender will accept while still being largely relevant to normal usage of emojis.

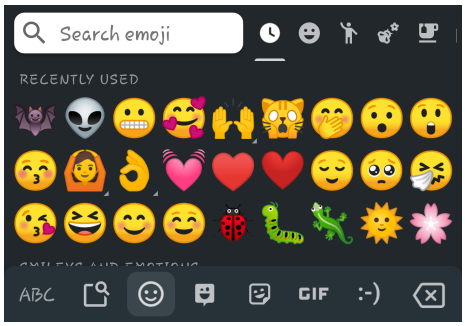


Figure 3.1: Recently used emoji section on the Gboard emoji keyboard. 27 emojis are displayed.

Three main approaches were taken to build the recommender: vector embeddings, word/emoji senses, and a categorical model trained on tweet responses. Vector and senses both mainly cover object and verb emphasis emojis, as well as some social emojis if keywords such as "hello" or "good night" are used. In the end, two were combined as they complement each other. The categorical model aims to predict the type of message, and covers mainly illocutionary as well as social emojis.

3.3 Preprocessing

The outputs of the recommenders can be compared with the emojis participants selected during the second part of the survey. Thus the same set of 10 sentences are used as input for the recommenders. In terms of preprocessing, each text message was first converted into lower case. Punctuations, duplicate tokens, and stop words were removed. The stop words list is based on the nltk English stop words corpus which can be found in Appendix A.[1] Some examples of the preprocessing output can be seen in Table 3.2.

3.4 Related Model

3.4.1 Vectors

The first approach is to utilise word embeddings, which are vectors that represent the meaning and context of words (Mikolov, Yih, & Zweig, 2013). The intuition for vectors is that words which appear in similar contexts will have similar meanings, and thus have similar vectors in the vector space of the corpus. Word2vec is an algorithm that trains vectors based on the task: "given a word, what is the probability of other words appearing near it?" If emojis are included in the training of embeddings, then for a given word, the most similar emojis can be returned (Barbieri, Ronzano, & Saggion, 2016).

[1] https://www.nltk.org/book/ch02.html

emoji2vec

Eisner, Rocktäschel, Augenstein, Bošnjak, and Riedel (2016) used a similar approach to train emoji vectors in the same space as 300-dimensional word2vec embeddings trained on Googlel News. However, instead of using emojis as the words, for each emoji, its vector was based on the emoji's Unicode description. For example, the "person in suit levitating" (Unicode description) emoji 🕴's vector would be the sum of the word vectors for 'person', 'suit', and 'levitating'. Eisner et al. (2016) referred to their emoji embeddings as emoji2vec, and it includes emojis up to Unicode 9.0 (there have been five more Unicode releases since). Each skin tone variation of people emojis were given their own vectors too. Using an emoji's Unicode description means that this method does not rely on users to use the emoji in order for the emoji to have a reliable embedding. This means that for newly released emojis, this is a good starting point to base the recommendation model on.

Recommendations

For a proof of concept, the emoji embeddings from Eisner et al. (2016) were used to see whether they could be leveraged for recommendations. For each text message, the vectors of all tokens were summed. The summed vector was then used to retrieve the 20 closest emojis based on their similarity to the emoji vectors. Depending on the set of tokens, some recommendations were related while some others are not. The recommended emojis can be seen in Table 3.3. Since each skin tone variation of an emoji has its own vector, sometimes multiple variations of an emoji are returned. These are removed in the table, so for some, the top 20 isn't a full 20.

Looking at the cosine similarities of the top 20 emojis, almost all emojis are above 0.35. Using 0.45 as the cut-off point, the recommendations become more concise. Table 3.3 shows the contrast between the initial results and the results after filtering using a 0.45 cut-off (third column). Using the cut-off generally decreased the number of recommendations, only leaving the ones which are related. In some cases, only a few emojis remain.

A different approach is to return recommendations based on each token's closest emojis. Giving each token a chance to affect the outcome ensures that the recommendations cover everything said. The rightmost column in Table 3.3 shows some example recommendations at 0.50 cut-off. Some tokens are close to a large number of emojis (e.g. vegetables) while some do not have any (e.g. leave). The full results for both methods can be seen in Appendix B. Looking at the per word recommendations gives an idea of which tokens have the most impact on the recommendations based on the sum of the vectors.

3.4.2 Emoji Senses

Apart from using emoji embeddings, another approach to finding related emojis is to look at the senses of a word and its related concepts. For example, "bank" could refer to both the financial institution as well as a river bank. Carrying this example to emoji recommendations, "bank" could result in both money related emojis 🏦💰 $ as well as river related emojis ♒🏄🦆. It is not so important for an emoji recommender to figure out which sense of the word the user is referring to, since recommending emojis that relate to both could encourage interesting uses of emoji. By combining Emojinet (a dictionary of emoji senses made by Wijeratne, Balasuriya, Sheth, and Doran (2017)), and an ontological approach based on BabelNet (a database of semantic relations), a new dictionary of emoji senses was made which links each emoji to a variety of concepts.

Table 3.3: Example vector recommendations. Columns show top 20 recommendations based on token sum, top recommendations based on token sum filtered by 0.45 cut-off, and recommendations based on each token.

Tokens	20 top Sum	Sum > 0.45	Per token
[yes, wait, vegetables]	[😊🍅🥜🍝🍆🧔🍠🌿🍌🍰🤦👅🍶🍜🎃🍦🍭🍥🌰]	[😊🍅🥜🍝🍆🧔]	['🧔', '🤦', '👎', '🙌', '😆', '😅', '😜', '🕴', '😔', '🤦', '👌'], [], ['🍅', '🌿', '🍝', '🥜', '🍏', '🍐', '🍤', '🥔', '🍤', '🍓', '🍆', '🌰', '🍇', '🍌', '🌾', '🍄', '🍉', '🥒', '🍚', '🍡']
[also, really, good, ghost, story, times]	[👍😔🤦👤🙋💭🎋💛🙈📰😕🎅👆🙂💚❤️😔🎃]	[👍😔🤦👤🙋💭🎋💛🙈📰😕🎅]	[], [💭, 🤦, 😔, 👍, 🙈, 🧔, 💛, 😂, 🤦, 🎎], [👆, 🤦, 🤦, 👍, 🤦, 🤦, 🤦, 👌, 💯, 💛, 😔], [👻, 🐊, 🎃, 👺], [🎭], []
[let, know, leave, campus]	[🙏🤦🙈🈂🧔😰🏫🏮👎😔🤦😔👌🙋💭]	[🙏🤦🙈🈂🧔]	[🙏, 🤦, 👎], [💭, 🤦, 🙈, 😂, 🧔, 🙏, 😰, 🤦, 😔, 😞], [], [🏫, 🚌]

EmojiNet

EmojiNet[2] is a resource connecting information from Emoji Dictionary (a crowd sourced emoji dictionary), Emojipedia (an emoji reference website that lists common uses for each emoji. It also keeps an archive of emoji versions across platforms), and BabelNet (a multilingual dictionary with a semantic network). Within EmojiNet, each emoji is linked to a set of BabelNet synset IDs. A synset is similar to a concept, for example, the word "bank" appears in 49 synsets on BabelNet (each with their own ID), the top two of which are "sloping land" and "a financial institution that accepts deposits and channels the money into lending activities".

In order for EmojiNet to be used as recommendations, for each token in a given text message, I fetched the BabelNet synset IDs that applied to it. If an emoji's set of synset IDs contained one of the token's IDs, then the emoji is added to the list of recommendations. Figure 3.2 shows an example for the tokens ['italian', 'noodles', 'rice'] (with made up synset IDs). When looked up on BabelNet, each token was linked to two synset IDs (concepts), then this was checked with the EmojiNet dictionary for matches. Indeed, 'italian' matched with 🍕 and 🍝, while 'noodles' matched with 🍝 and 🍜, note that 'rice' did not match with anything (only for this arbitrary example. In actuality, 🍚 would have likely matched with rice). Thus, for the tokens ['italian', 'noodles', 'rice'], the following emojis would be recommended: [🍕, 🍝, 🍜].

Some of the recommendations for the trial sentences can be seen in the first column of Table 3.4 (recommendations for all 10 sentences can be found in Appendix C. Some tokens do not return any emoji (e.g. 'touch') while some tokens return a large list of emojis (e.g. 'go'). The recommended emojis all 'make sense' and are related to the token in some way. Sometimes this is more metaphorical, for instance, the emojis for 'go' cover those related to movement/transit

[2] http://emojinet.knoesis.org/ though the website is currently down.

Table 3.4: Recommendations based on EmojiNet senses and BabelNet ontological relationships

EmojiNet	**Ontologies**
touch: [], cursed: [], paper: [🗞️, 📰, 📘, 📑, 🖨️, 📝, 📜, 📋, 📄, 📃], currency: [💹, 💷, 💶, 💵, 💱, $, 💸]	touch: [🤗, 💆, 👯, 👫, 👬, 💑, 🏉], cursed: [], paper: [🎉, 📃, 📄, 📰, 🗞️], currency: [💱]
want: [💚, 🙏], go: [🏃, 🚶, 🚳, 👣, 👼, 🧿, 🌭, 💀, 🚾, 🚹, 🚩, 🚤, 🚌, 🚎, 🚞, 🚚, 🚛, 🚕, 🚙, 🚐, 🚒, 🚆, 🚇, 🚉, 🚝, 🚄, 🚅, 🚈, 🚂, 🚀, 🛰️, ➡️, 📦, 💸, 🎒, ⬅️, 🌐, 🌍], home: [🏠, 🚪, ⛪, 🏘️, 🏰, 🏯, 🏪, 🏚️, 🏡, 🏠, 🏘️], raining: [💦, 🌧️, ☔, ☂️, 🌧️, 🌦️, 🌂], bike: [🚵, 🚴, 🚲, 🚲, 🏍], rain: [💦, 🌧️, ☔, ☂️, 🌧️, 🌧️, 🌂, 💧]	want: [], go: [🏃, 🥁], home: [🛏️, 🏘️, 🏠, 🏡, 🏫, 🏢, 🏬, 🏰, 🚤, 🚪, 🛋], raining: [🌧️, 🌦️, 🌧️, 🌧️], bike: [🚲, 🚵, 🛵, 🚲], rain: [🌧️, 🌦️, 🌧️, 🌧️]
good: [📈, 👍], ideas: [💭, 💬, 💡]	good: [🎁], ideas: [🤔]

as well as those related to death. The emojis for 'want' includes the green heart emoji 💚, green often being the colour of jealousy.

```
for each token in tokens:
        lookup BabelNet synset IDs for token
        check if any of the synset IDs match EmojiNet dict

tokens = ['italian', 'noodles', 'rice']
lookup token BabelNet IDs
        'italian'   : [bn0001, bn0005],
        'noodles' : [bn0003, bn0006],
        'rice'      : [bn0007, bn0008]

EmojiNet dictionary = {
        🍕 : [bn0001, bn0002],
        🍝 : [bn0001, bn0003],
        🍜 : [bn0003, bn0004]
}
matched emojis
        'italian'   : [ 🍕, 🍝]
        'noodles' : [🍝, 🍜]
        'rice'      : [none]
```

Figure 3.2: Example dictionary look-up process for each token

Ontologies

BabelNet connects concepts through ontological relationships. For example, banana *is a*: herb and a berry, and it *has part*: banana peel. Some emojis can be searched in BabelNet, for example, when the potato emoji 🥔 is entered into the search bar, the page for potato is returned (Figure 3.3 left). The BabelNet API seems to be using version 4.0 which includes less emojis than their live version. For version 4.0, most emojis also only return one result, while the live version sometimes returns multiple senses for the same emoji. This means that recommendations relying on the BabelNet API would not cover as many emojis as there currently exist. However,

the present method will likely still be useful in the future since BabelNet is actively expanding emoji entries.

The four types of ontological relationships given by BabelNet are: hypernymy, hyponymy, meronymy, and holonymy. Figure 3.3(right) shows the four relationships using the potato 🥔 example. 'Root vegetable' is a hypernym of potato while Red Pontiac, a kind of potato, is a hyponym of potato (therefore, potato is a hyponym of root vegetable and hypernym of Red Pontiac). The other two relationships have to do with wholes and parts. Baked potato contains potatoes; baked potato is a meronym of potato. A jacket is the outer skin of a potato, therefore it is a part of a potato; jacket is a holonym of a potato.

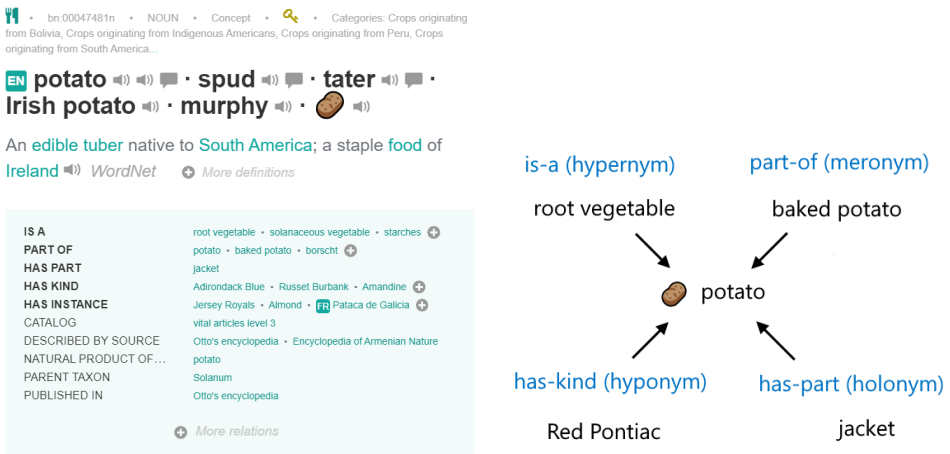


Figure 3.3: a) Result page when searching the potato emoji 🥔 on BabelNet. b) Relationships between potato 🥔 and other concepts.

A dictionary was created for each emoji that maps onto the BabelNet IDs of its set of hypernym, meronym, hyponym, and holonym relations, as well as itself (i.e. the ID list for 🥔 includes the ID of 'potato'). Since there is a limited number of keys/requests that can be made to BabelNet per day, building an emoji dictionary means that each emoji does not have to be looked up each time, saving the keys for looking up the tokens. Some emojis do not have a BabelNet entry, so would not be in this model's library.

A similar search process is then carried out for each token as previously done for the EmojiNet senses where each token's potential senses are checked against each emoji's related senses (see Figure 3.2 again for look-up process). This time, however, the ontological dictionary created here is used for the matching process instead of the EmojiNet dictionary. Initially, a large number of tokens resulted in the following emojis: [🎞, 🎮, 📕, 📗, 📘, 📙, 📚]. Turns out, these tokens were also names of movies, video games, or books, falling under the hypernym/hyponym relation (e.g. the taxi emoji 🚕 is linked to the Taxi movie from 2004). These two emojis: [📰, 🗞] appear less often than the previous set, but still occurred in the recommendations of unrelated tokens, likely for names of news outlets. For these two sets of emojis, the hyponym relations were removed when building the dictionary.

Some example results from the trial sentences can be seen in the second column of Table 3.4 (the full recommendations for the 10 test sentences can be seen in Appendix C). The recommendations using ontologies give different results than those of EmojiNet, although sometimes the recommendations do overlap (e.g. 'paper', 'rain'). The two approaches utilising emoji senses give overlapping results for some tokens (e.g. 'rain', 'paper'), while completely different results for others (e.g. 'ideas'). Sometimes one approach provides recommendations while the other does not (e.g. 'touch').

3.5 Most Used Model

So far the emoji recommendations are mostly objects and verbs. However, a large portion of emojis in daily use are the smileys. Smiley emojis are more nuanced, as they can both be used to strengthen the sentimental value of the text message, as well as to change the illocutionary force entirely (i.e. sarcastic emojis such as the upside-down face 🙃). The main goal of this section is to create a recommender model that covers the illocutionary emojis. An important heuristic used to create this model is: most emojis are used to serve the illocutionary function, thus if most used emojis are recommended, it is likely that some of these are useful for serving said function.

In order to recommend appropriate most used emojis given a text message, there still need to be some understanding of the text. For this purpose, the EmotionGIF2020 challenge[3] is used as the training set. The original aim of this challenge is to correctly predict the categories of GIF used given a tweet. GIFs are often used alongside text in a tweet to mark the general intention (not unlike emojis). For the present purpose, these GIF categories can be seen as the intention tags of the text and will be used to train a classifier that predicts these categories. The tweets in the training data often also include emojis, which will help with the assignment of emojis to each category. This will allow a prediction of a text message's intentions, and consequently emoji recommendations fitting that category.

3.5.1 Data and Preprocessing

The EmotionGIF2020 challenge features 32,000 labelled two-turn Twitter threads. The response tweet always includes an animated GIF, sometimes including text (12,762), sometimes with no text and only the GIF response (19,238). Figure 3.4 shows an example of an original tweet, and the response tweet including text. Only the responses with text from this dataset were used for the present model. The original task of the challenge is to predict the GIF label(s) given the original tweet and the reply (which can be empty). The GIF responses are categorised into 43 categories, these are the labels assigned to each thread.

Figure 3.4: Example of a Twitter thread, showing the original tweet, and the response which features the text "Hell yeah", as well as a GIF of a man clapping.

Categories and Factor Analysis

There were 43 possible categories that the GIFs could be labelled as. Some GIFs were labelled with multiple categories. 43 is a really high number of categories, which results in higher difficulty of training a high-accuracy model. Thus the first step of the process is to simplify the categories. This will be done using factor analysis, which is a method that combines similar, correlated variables into a lower number of factors. Some categories co-occurred with others more frequently, which suggests that the labels were used in similar situations and are correlated. A co-occurrence matrix can be made by counting the number of times a pair of categories were used for the same GIF. Two categories, popcorn and thank_you, did not co-occur with any other category and were removed for factor analysis.

Exploratory factor analysis was carried out using the python factor_analyzer module [4]. The categories and which factors they load on can be seen in Appendix D. 15 factors had eigenvalues

[3] https://sites.google.com/view/emotiongif-2020/home
[4] https://github.com/EducationalTestingService/factor_analyzer

Table 3.5: New and final categories for the Most Used model, and the old categories that belong to them. An empty "Old categories" entry means the "New category" was its original label.

New category	Old categories	Emojis
sigh	[sigh, eye_roll, smh, facepalm, seriously, no, thumbs_down, yawn]	🤦🙄😆😡😢
sorry	[deal_with_it, sorry, oops, yolo]	😅😩✌️👇
good	[win, awww, agree]	😇😔❗💛
love	[kiss, want, you_got_this]	❤️😄😁💋😌
support	[good_luck, hug]	💜🤗💗🌸🙂
flirty	[wink, hearts]	😏❤️🌹👉
idk	[idk, shrug]	😅🤷🤔😏😑
shock_disgust	[eww, omg]	😱💔😟🙈
applause	[applause, slow_clap]	👏🎉✨🤪
dance	[dance, happy_dance]	😍💃🔥😃🎶
yes		🙋😜😊🤩❣️
high five		💪😎💯😀
popcorn		🤣👀😉😊
please		🙏😭😌💥
thank you		😘💙💖💞💚
fist bump		❤️👊💕😊😄
thumbs up		👍😊👌🖤🙌
removed	[ok, mic_drop, do_not_want, oh_snap, scared, shocked]	

larger than one, however, factors 8 to 15 did not have multiple categories that loaded strongly on it. Factor 1 can be broken into two subgroups of categories. The first set has loadings higher than 0.9, the second set has loadings between 0.6 and 0.8. The second set also loads on other factors as well. Based on the secondary loadings of the second set, they can be further broken down into [idk, shrug] and [ewww, omg]. This resulted in 17 new combined labels or categories.

The reply tweets were relabelled with the new combined labels according to the factors (for example, tweets with the labels 'idk' or 'shrug' were both relabelled into 'idk'). Then the emojis that accompanied each reply tweet and (new) categories was counted. From here, some categories were removed due to having a low number of emojis (e.g. 'ok'), or a low occurrence in the dataset (e.g. 'mic_drop'). The final set of categories and which original labels formed them can be seen in Table 3.5. From here, the emojis were recounted according to the final set. For each category and emoji, its tfidf value is calculated according to the Formula 3.1, e for emoji, c for category.

$$tfidf(e,c) = tf(e,c) * (log\frac{1+n}{1+df(c,e)} + 1) \tag{3.1}$$

The most used emojis (according to The Unicode Consortium (2020b)) were assigned to the category where they had the highest tfidf value, unless that category already has five emojis. If this was the case, emoji already assigned with the lowest tfidf value and the emoji to-be-assigned are compared. The one with the higher value "makes the team" for that category, and the one that didn't make it looks for a space in the category where it has the second highest tfidf value.

Training the model

Since the categories were greatly imbalanced, the categories with a large number of cases were under sampled while the categories with a low number of cases were oversampled. The Keras

library[5] was used to train a sequential LSTM model.

3.6 Final Recommender Models and their Offline Performance

3.6.1 Final models

Related

Emojis that match with each word from both EmojiNet and the Ontologies are returned. If this results in a set of emojis larger than 10, then the sum of the word vectors from the input is used to rank the emojis with regards to cosine distance. Then the top 10 closest emojis are returned. On the other hand, if the original set from EmojiNet and Ontologies is lower than 10, this is supplemented by emojis that are closest (but not already in the list) to the sum of the input's word vectors.

Most Used

Given a piece of text, the Most Used outputs the two categories with highest probabilities. This translates into 8 - 10 emoji recommendations per input, since 4 to 5 are attributed to each category.

Combined

The Combined model is simply the combination of the outputs of the other two models, removing any duplicated emojis. This results in a varied set of between 15 and 20 emoji recommendations.

3.6.2 Offline evaluation metrics

The outputs of this chapter are three recommender models (Related, Most Used, Combined). These are the ones that will be evaluated in the next chapter with an experiment. Some expectations for how the models will perform could be gained from offline evaluation. The offline evaluation will be carried out with the test set collected during the first part of the emoji variability survey. Offline evaluation of recommender systems have traditionally revolved around some sort of accuracy score, however, accuracy is not necessarily the goal of the present task (or many recommender systems!). Aside from accuracy, being able to recommend emojis the user is unfamiliar with is also important, especially for broadening the user's emoji vocabulary.

Accuracy-related measures may be tricky to define for the present problem as everything apart from true positives are not so straightforward. For instance, if the original text included this set of emojis [🥰, 😊, 🙌], and this is recommended [🥰, 😍, 😘, 😚, ❤️, 🎉], how would true negative, false positive, false negative be defined? For every emoji that was not used and also not in the recommendations, does this count as a true negative? Since there are over 3000 emojis, looking at negatives does not feel very impactful. Instead, an adaption of precision was used here, which, for each input-emoji(s) pair, it is counted as a true positive if the set of recommendations contain the original emoji(s) used. Thus for the example above, this would be a precision of 0.33.

$$ILS_{user} = \frac{1}{2} \sum_{i_j \in L} \sum_{i_k \in L} sim(i_j, i_k) \tag{3.2}$$

As stated previously, measures other than accuracy are also important. Here I looked at diversity and coverage. Diversity measures the variety of the recommendations for each user. For example, [🍕, 🍔, 🍟] is less diverse than [🍕, 🤠, 🦵]. Diversity is calculated

[5] https://keras.io/

Table 3.6: Model performance in terms of precision, diversity, and coverage on test data

Model	Precision	Diversity	Approximate Coverage
Most Used	0.014	0.302	0.027
Related	0.006	0.263	0.204
Combined	0.020	0.243	0.215

as the mean ILS (Intra-List Similarity; see Equation 3.2) of each set of recommendations (Anna, 2016; Ziegler, McNee, Konstan, & Lausen, 2005), the higher the score, the more similar each recommendation is to each other, and thus the less diverse the recommendation. Since the number of recommendations varies depending on the input, the ILS for each set of recommendations is normalized by the number of comparisons. Another measure is coverage, which is the percentage of emojis that the model can recommend. Since a large number of emojis are the various skin tone and gender combinations, removing these from the whole set of 3304 leaves a rough estimate of 2000, this smaller number is used for calculating coverage(The Unicode Consortium, 2020a). The performance of the models can be seen in Table 3.6. The preprocessing of the text messages are the same as for the example recommendation output messages used previously.

3.6.3 Precision

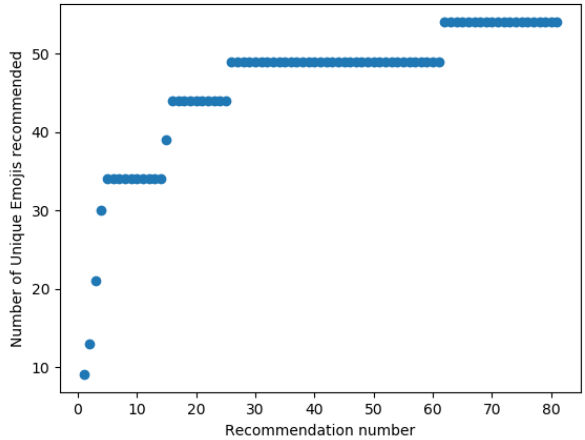
All three models did not perform very well on precision, although the Most Used and Combined did do a little better than Related. This makes sense as these two models include common emojis in their recommendation, so by chance should recommend emojis the original messages contained more often. The Related model does add a few more correct hits to the Combined model.

3.6.4 Diversity

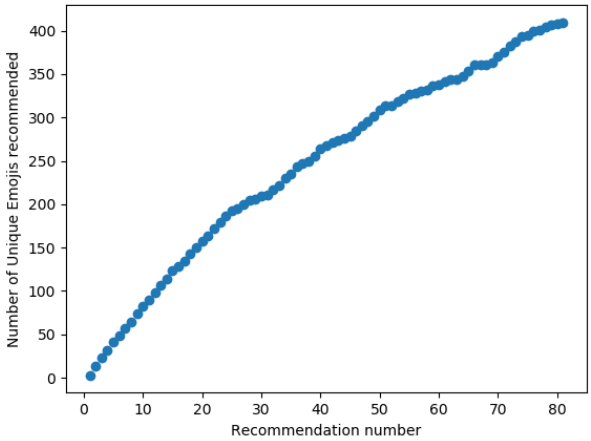
The higher the diversity score, the more similar the recommendations are for each text message. The Most Used model scored the highest, meaning the recommendations are the least varied, while the Related and Combined model score quite similar. Including Most Used emojis in the Combined model does improve the diversity score a little bit compared to the Related model. This suggests that by combining the models, a wider range of emojis can be covered, increasing the chance of a recommendation the user will find useful.

3.6.5 Coverage

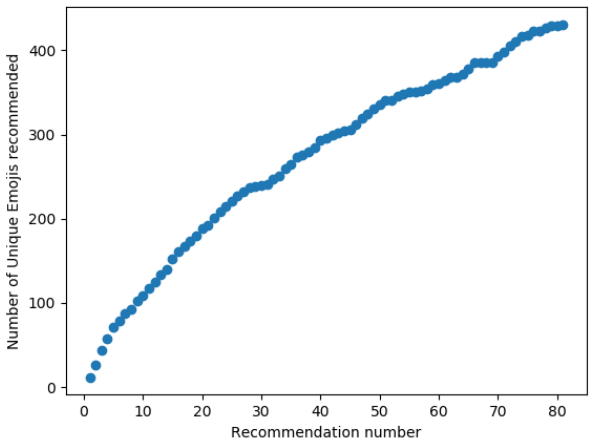
The coverage is the number of emojis that the model can potentially recommend. Since the related model is a combination of emoji embeddings from Eisner et al. (2016), EmojiNet's emoji sense dictionary (Wijeratne et al., 2017), and my own ontology based sense dictionary, I don't know for certain which emojis are covered. Thus, the coverage for the related model was calculated by taking the unique emojis the recommender recommends after recommendations have been made for the entire test set. The Most Used model, as designed, only covers 54 emojis. The Related and Combined model in the end covers 409 and 430 emojis respectively. Figure 3.5 shows the number of emojis after each input. From this it can be seen that some of the categories in the Most Used model are not recommended very often (only appearing after many messages). This could also mean the test set is not very varied.



(a) Most Used



(b) Related



(c) Combined

Figure 3.5: Coverage of the three models after making 81 recommendations.

3.7 Conclusion

In this chapter, 3 final emoji recommendation models were made: Related, Most Used, and Combined. The Related model makes use of emoji embeddings made by Eisner et al. (2016), as well as two emoji sense dictionaries. The Most Used model outputs categories that the text falls under. Emojis that were most used according to The Unicode Consortium (n.d.) have been distributed to each category depending on occurrence. The categories that the Most Used model predicts are thus turned into emoji recommendations. The Related model targets emphasis emojis, doing a good job at recommending concepts that have occurred in the text. The Most Used model, on the other hand, recommends emojis that are more likely to be used for the illocutionary function of altering the tone of the text. Finally, the Combined model adds the recommendations made from the previous two, forming the best of both worlds, a well rounded model. The three models can now be put into the world to see how they perform.

Chapter 4

Evaluating the Recommender

The emoji recommender is evaluated in a within as well as between subjects experiment. During the experiment, participants carry out two short conversations with a chatbot, one without emoji recommendations where they are prompted to respond to messages as they normally would, and a second one with recommendations where they are asked to refrain from using emojis. In this second conversation, emoji recommendations are provided, participants are asked to select the ones they would use from this list. This allows within-subjects comparison of their emoji use with and without recommender. This also allows between subjects comparison for the different recommender models (Most Used, Related, and Combined). After the conversations, participants fill in a survey which has Likert scale questions for the evaluation of the recommender performance, as well as open questions to probe how they feel about emoji recommendation in general.

4.1 Research questions

The main research questions of the user evaluation are as follows:

RQ 1 What is the added value of an emoji recommender?

RQ 2 How do the three models compare with each other? Are both the Related and Most Used model important?

Research question 1 investigates whether an emoji recommender influences the emoji use behaviour of users. This will be answered through a comparison of participants' emoji use with and without a recommender, as well as some open questions. Some questions will be targeted to gauge opinion of the potential of having "the ideal recommender".

Research question 2 investigates how the two main recommendation models (Most Used and Related) perform and work together (Combined). The creation of the three models was discussed in Chapter 3. This question will be answered by comparing emoji use between the participants during the with recommender conversation. Emoji count and number of unique emojis (emoji variability) are the main measurements during the experiment. In order to see how the models work together, within participants in the Combined model condition, the number of emojis chosen from each model is counted and compared. The evaluations of each model in the survey is also compared.

4.2 Hypotheses

Hypothesis 1.1 Recommendations will increase the number of emojis used.

Hypothesis 1.2 Recommendations will increase the number of unique emojis used.

Hypothesis 2.1 When a broader range of emojis are recommended (i.e. with the Related and Combined model), users will use a broader range of emojis.

Hypothesis 2.2 Users will use fewer emojis when presented with recommendations from the Related model than from the Most Used model, since Related model recommends emojis users may be less familiar with and thus less likely to use.

Hypothesis 2.3 Users will rate the Combined model higher, followed by the Most Used model. This is because the combined model offers what the users commonly use, while the Related model might be a hit or miss.

Hypothesis 2.4 Within the Combined model, users will use the most common emojis more often.

4.3 Experiment Design

4.3.1 Explanation of Discord

The experiment was carried out over Discord, a free text messaging and VoIP (Voice over Internet Protocol) platform that facilitates organisation of communities. Discord was chosen because of its support for bots, which can be designed to perform tasks such as adding the emoji recommendations, and guiding participants through the user test.

A Discord server was created for the experiment. Within the server, there are multiple text channels which help provide structure to the experiment. A hashtag '#' will be used to denote text channels. The various channels and general flow of the experiment can be seen in Figure 4.1. A large amount of the participant's interactions within the experiment is done through reactions. Reactions are emojis added to a message and shows up under a message. For example, in Figure 4.2, there is one ✅ reaction added to the message already. The participant can also react with ✅ to indicate their agreement. Once the participant has done this, they are given the role "Participant", which grants them permission to see the other channels.

Three bots were created for the user evaluation process: Admin Bot, Recommender Bot, and Conversation Bot (Connie). Admin Bot takes care of processing reactions appropriately so that the participant has access to channels and information when they need it. Recommender Bot adds reactions to the participant's messages during the With Recommender conversation. Lastly, Connie carries out the short structured conversations with the participant which allows showcasing of the recommendations. Bots are used to ensure that each participant is given the same conversation and context, and allowing the experiment to move smoothly.

4.3.2 Experimentation flow

The flow of the experiment can be seen in Figure 4.1. When participants first join the experimentation server, they will only be able to view the #information-and-consent text channel, which will show an overview of the evaluation procedure as well as ask for their consent for participation. They are asked to react to the message if they consent (Figure 4.2). After the appropriate reaction is added by the participant, they are given permission to view #introduction and #conversation.

In #introduction, participants are first given a link to the survey. The survey has two parts: demographics and after-conversation evaluation. Participants are asked to only fill in the demographics portion for now, which asked for their age, how often they use emojis, and how often they use Discord. Connie is introduced, along with a brief explanation for how to type emojis within Discord. During the conversation with Connie, participants were instructed to reply as they normally would, using emojis or not as usual. At the end of this first conversation, participants are guided to the #recommendation channel.

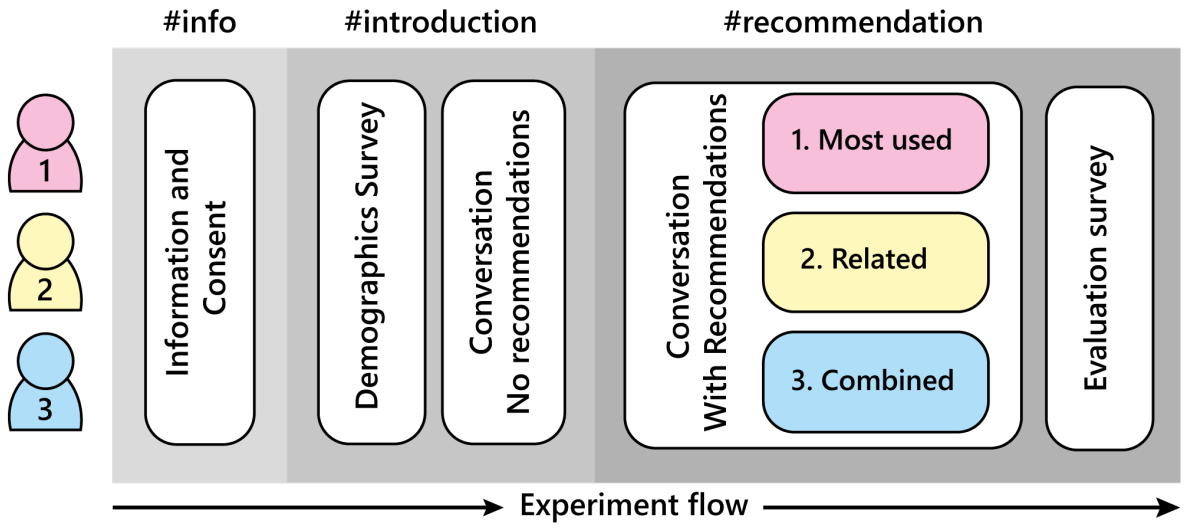


Figure 4.1: Flow of the experiment from left to right. The name of the channel where it takes place is at the top of each section. Participants are put into one of three conditions: Most used, Related, and Combined; this changes the recommender's model.

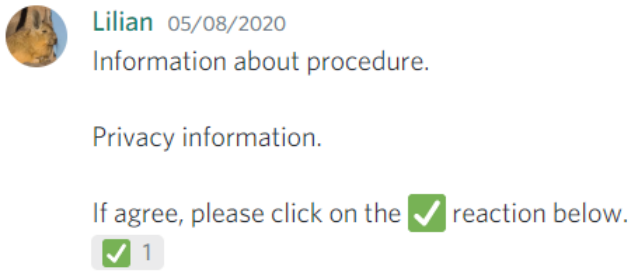


Figure 4.2: The first message the participant are presented with when they join the server.

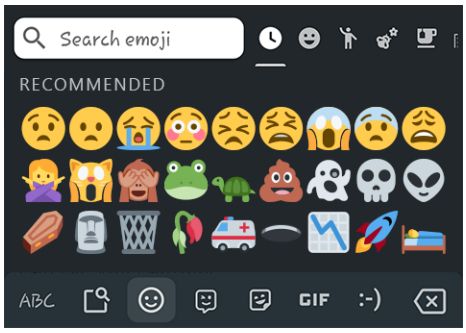


Figure 4.3: Example picture used to demonstrate to the participant where and how the recommendations may appear on a smartphone emoji keyboard.

The #recommendation channel hosts the conversation with emoji recommendations. Participants are introduced to the recommender and recommendations by Admin bot using a short demonstration. Emoji recommendations are added to the text messages the participant sends by Recommender bot via reactions. Participants are asked to imagine the recommendations as if they showed up on their phone's emoji keyboard as in Figure 4.3. The whole demonstration conversation can be seen in Appendix G. Once participants understand the recommendation process, Connie will engage in another short conversation with the participant. The three recommender models (i.e. Most Used, Related, and Combined) are assigned randomly, with the first participant given Most Used, the second Related, the third Combined, fourth Most Used again, and so on.

Connie's script and the questions she asks are designed to be open and to encourage a variety of answers and thus recommendations. Two sets of questions/prompts were written, the order of which set is used in the first and second conversations alternates per participant similar to how the recommender model was assigned. The two sets of conversations can be seen

Table 4.1: Evaluation questions

Quality	1. The recommender gave me good suggestions 2. Finding an emoji to use with the help of the recommender was easy 3. The recommender effectively helped me find the ideal emoji 4. The recommender is useful 5. Overall, I am satisfied with the recommender 6. If a recommender such as this exists, I will use it to find emojis to use
Novelty	8. The recommender bot helped me discover new emojis 9. The emojis recommended to me are novel and interesting 10. (Most of the recommended emojis are familiar to me) 11. (The emojis recommended to me are similar to each other)
Emoji Overload	12. There were too many recommendations
Open Questions	1. How do you currently experience the process of inserting emojis on your phone? Please describe any difficulties you encounter, if any. 2. Would you rather have a "Recommended" or "Recently used" section in your emoji keyboard? Maybe both? Why or why not? 3. Any other thoughts?

in Appendix F. Some examples of the prompts Connie asks are:

- Is there something you're scared of? Why?
- What did you do yesterday?
- Look at this rad croc! (shows video of a crocodile jumping into water with people cheering)

After the second and final conversation, participants fill in the second part of the survey. When participants have completed the survey, they are instructed to return to Discord and react to the last message Admin Bot sent with the appropriate emoji, at which point they are removed from the Discord server, and their Discord conversation is logged.

4.3.3 Recommender Evaluation questions

In the evaluation section after the conversations, there is a combination of 11 questions rating the recommender on a Likert scale (5 point, strongly disagree to strongly agree), as well as more general open questions about emoji recommendations. The Likert scale questions are based on Pu, Chen, and Hu's (2011) work about user evaluations of recommender systems, adapted for the current study. The order of the Likert scale questions is randomised. The questions can be seen in Table 4.1, and are grouped into questions about recommender quality, novelty of recommendations, and a single question on whether there were too many recommendations. The questions in brackets are inverted. There were also three open questions, the aim of these were to gather general thoughts about emoji input and prospects of a recommender.

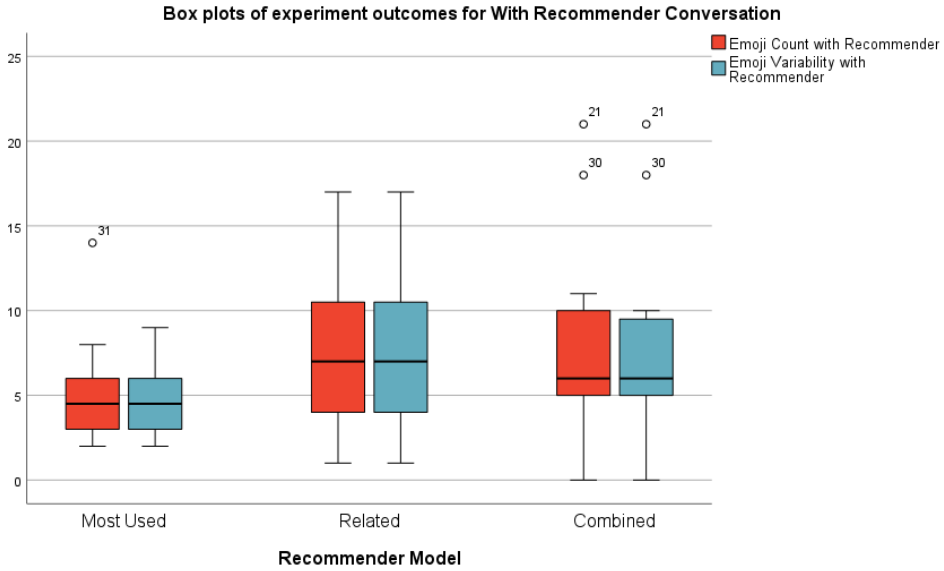


Figure 4.4: Boxplots showing outliers for Emoji Count and Emoji Variability during the With Recommender conversation

4.4 Results

4.4.1 Preprocessing, descriptive statistics, and potential covariance variables

49 people participated in the experiment, of these, three people (participants 10, 45, 47) did not complete the evaluation survey so were removed from analysis. Furthermore, the instructions were not always clear for everyone. During the first conversation with no emoji recommendations, three participants only used emojis with no text, perhaps because they were introduced to the experiment thinking it's "An Emoji Thesis", and a previous survey sent out by the researcher did ask for emoji input. This did not result in any outliers in emoji count or emoji variability for that conversation. A few participants started out answering with only emojis, realised there might have been a misunderstanding and messaged the researcher to check. Either way this did not affect the results, and they did use text for the second conversation. For the second conversation with an emoji recommender (especially for Related and Combined conditions), some people selected all emojis that were relevant, instead of the ones they would actually insert in the situation. This resulted in three outliers (participants 21, 30, 31 see Figure 4.4), who were removed. Of the 43 remaining, the mean age was 26.07 (SD=3.158), mean emoji use was 3.35 (between 'a moderate amount' and 'a lot'; SD=0.783), mean Discord use was 2.84 (between 'a little' and 'a moderate amount'; SD=1.430). 13 participants were in the Most Used condition, 16 in Related, and 14 in Combined.

First, for each conversation, the number of emojis used as well as the number of unique emojis used were counted for each participant. The two sets of questions asked by Connie may elicit different emoji counts, which may affect the results, but upon plotting and visual inspection, this did not seem to be the case (see Figure 4.5). Second, correlations between the demographic information (age, frequency of emoji and Discord use) and emoji use within the conversations is checked. No correlations were found so this did not need to be kept in mind during the rest of the analysis.

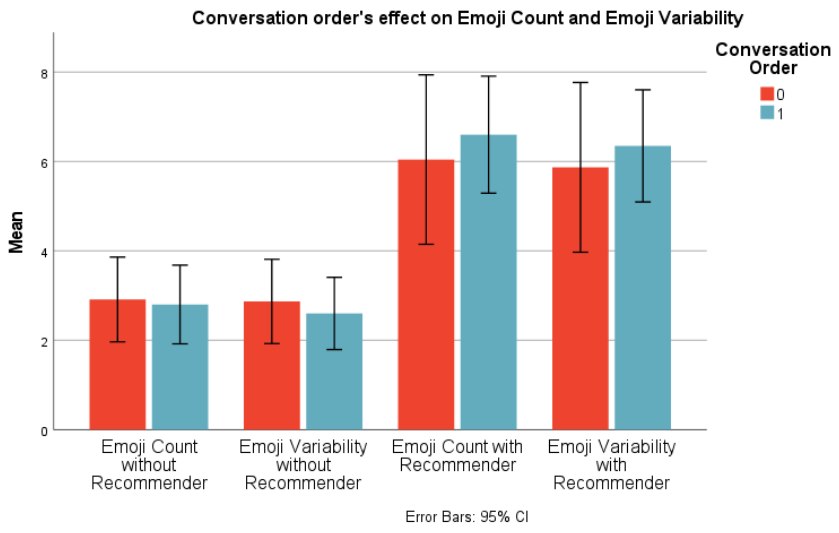


Figure 4.5: Bar chart of Emoji Count and Variability depending on Connie's conversation order

4.4.2 Differences between With and Without Recommender Conversations

Hypothesis 1.1: Recommendations increases the number of emojis used

A paired-samples t-test was conducted to compare emoji count in the without recommender and with recommender conditions (i.e. conversation 1 vs conversation 2). In the with recommender condition, participants used significantly more emojis (M=6.30, SD=3.700) than in the without recommender condition (M=2.86, SD=2.030); t(42)=5.235, p<.001. See left two bars in Figure 4.6.

Hypothesis 1.2: Recommendations increases the variety of emojis used

A paired-samples t-test was also used to compare emoji variability between the with and without recommender conditions. Participants used a significantly larger variety of emojis in the with recommender condition (M=6.09, SD=3.663) than in the without recommender condition (M=2.74, SD=1.965); t(42)=5.261, p<.001. See right two bars in Figure 4.6.

4.4.3 Evaluation of the different models

Hypothesis 2.1: Broader range of recommendations encourage higher emoji variability

Planned contrasts showed that emoji count was significantly higher in the related condition (M=7.94, SD=4.697) compared to the most used condition (M=4.62, SD=1.895); t(40)=2.532, p=0.015. This is in the opposite direction as the hypothesis. See left cluster in Figure 4.7.

Hypothesis 2.2: Lower emoji count with Related model compared to Most Used model

Planned contrasts showed that when a broader range of emojis are recommended (as in the related and combined models), participants used a significantly broader range of emojis than in the related model; t(40)=2.399, p=0.021. See right cluster in Figure 4.7.

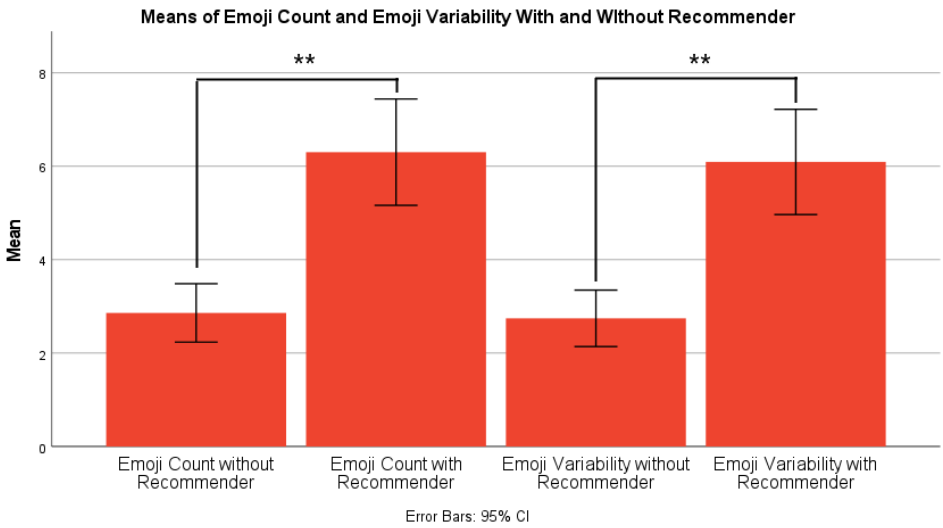


Figure 4.6: Bar graph showing the difference in emoji count and emoji variability between the with and without recommender conditions. * signifies $p<0.05$, ** signifies $p<0.01$

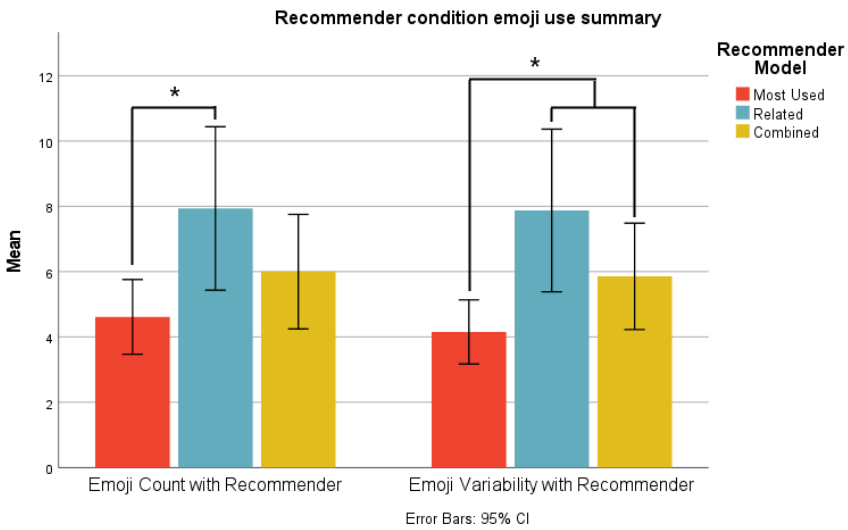


Figure 4.7: Bar graph showing the emoji counts and variability between the different recommender conditions. * signifies $p<0.05$, ** signifies $p<0.01$

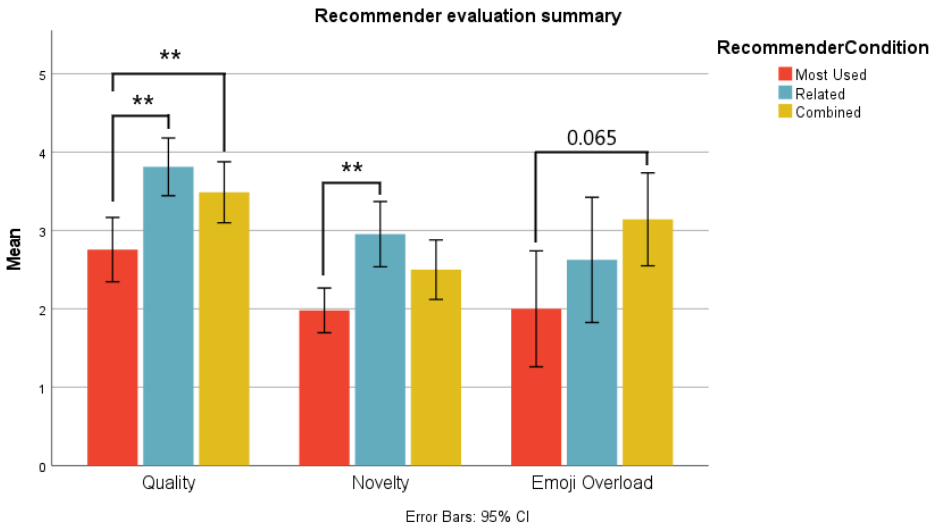


Figure 4.8: Bar graph showing the evaluation survey outcome means for the different recommender conditions. Scores are averages of questions according to Table 4.1. * signifies p<0.05, ** signifies p<0.01

Hypothesis 2.3: Combined model rated highest, followed by Most Used, then Related

There was a statistically significant difference between the groups as determined by one-way ANOVA for quality (F(2,40)=8.823, p=0.001) and novelty (F(2,40)=7.771, p=0.001) but not for whether there are too many emojis presented (F(2,40)=2.693, p=0.081).

A Tukey post hoc tests showed that for quality the Most Used model (M=2.756, SD=0.679) was rated significantly lower than the Related (M=3.813, SD=0.691, p<0.001) and Combined (M=3.488, SD=0.675, p=0.022) model. There is no significant difference between the Related and Combined models (p=0.404). For novelty, the Related model was rated highest (M=2.953, SD=0.781), which was significantly higher than the Most Used model (M=1.981, SD=0.473, p=0.001), and higher but not significant than the Combined model (M=2.500, SD=0.658, p=0.159). The Combined model was also rated higher in novelty than the Most Used model, though this was not significant (p=0.116). See Figure 4.8.

Hypothesis 2.4: Most common emojis used more with Combined model

For participants in the Combined model condition, the emoji counts can further be broken down to emojis chosen from the Most Used model or Related model. A paired-samples t-test was then conducted to compare the emoji counts between the two. Participants used a significantly higher number of emojis from the Most Used model (M=4.21, SD=2.751) than the Related model (M=1.79, SD=1.369); t(13)=2.925, p=0.12 (Figure 4.9).

4.4.4 Summary of the Open Question answers

Current experiences and difficulties of emoji insertion on the phone

Some people stated that they did not have difficulties with emoji insertion. Most emoji keyboards already having a recently/frequently used section, which is enough those who only use a small set of emojis. For some, emoji insertion is not a problem because they don't use emojis at all, or prefer emoticons. Though one person said they use emoticons because they did not want to bother switching to the emoji keyboard. Some people who use Discord often, find

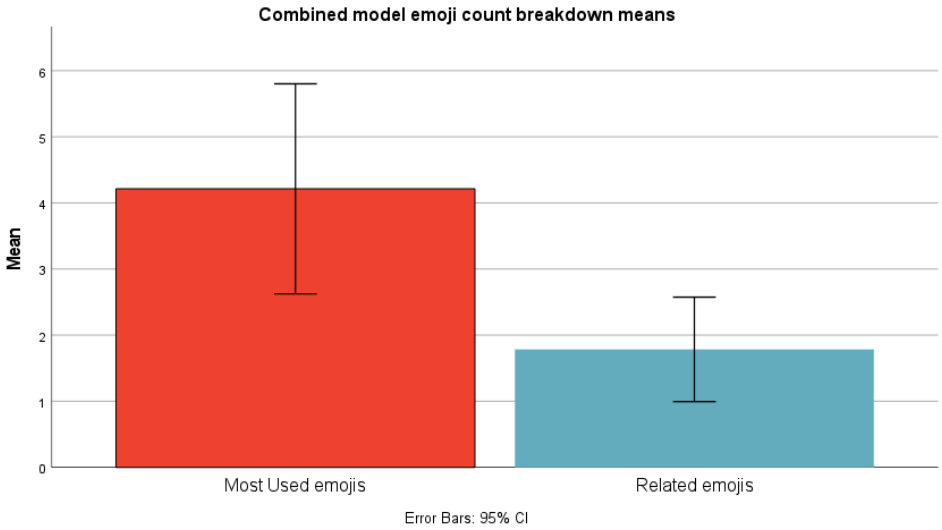


Figure 4.9: Bar graph showing the model attribution breakdown of emojis counts within the Combined condition.

that inserting emojis by short code (e.g. Discord shows various options when you begin typing ":smile" as in Figure 4.10) is fast, especially once they're familiar with the short codes for the emojis they use most often.

When the recently/frequently used section is not enough, people use a combination of the search bar function, simple recommendations provided by the keyboard (e.g. iPhone, Gboard), and looking through the emoji keyboard. The search function is enough for some people and used often, but unhelpful when the names of the emoji or search terms to use are unknown, as well as when the search function is not "smart". Looking through the emoji keyboard is often tedious, with the whole emoji list being a bit overwhelming, this is also mentioned in Pohl et al. (2016). Although someone did say their most used emojis are often near the top of the emoji list (most likely faces), so they did not find looking through the emoji keyboard too bad. People appreciated when their phone suggests/recommends emojis. These recommendations are generally based off of the last word typed, and acts as autocomplete (e.g. if I typed "fox", then use the suggested fox emoji, the emoji would replace the "fox") or next-word-insert (usually after a space has been inserted after the last completed word, emojis suggested here would appear after the last word you typed).

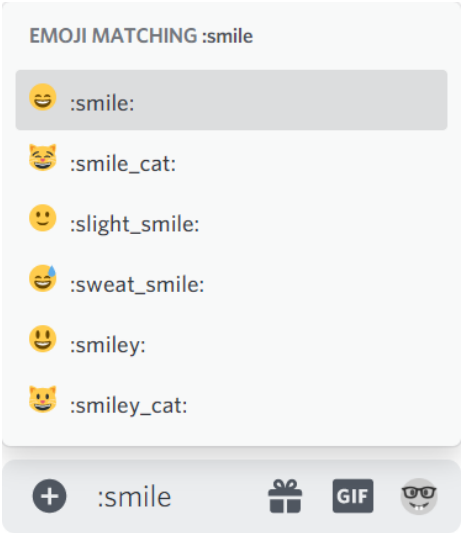


Figure 4.10: Example of Discord insert by short code.

Other comments about emoji insertion included that a slow phone makes the entire process longer, and that the emojis are not the same across apps. Someone said that they enjoy the customisable "thumb-up" emoji on Facebook Messenger (Figure 4.11). This emoji can be quickly accessed and chosen to be different for each conversation.

Thoughts on implementation of recommended or recently used section into emoji keyboard

Most people answered that if given the opportunity to have both, they would prefer both recently used and recommended. Noting that if the recommender algorithm is well designed, it should take into account their frequently used emojis too. People especially liked the idea of a recommender because it may show new and diverse emoji they don't usually use, thus expanding their emoji vocabulary. It also functions as an initial search for related emojis that are quite straightforward but not easy to find. It is helpful when the most recently used consists of random emojis that are not related to the current topic, and when the user doesn't know what emoji to use. Someone noted that it depends on the interface of the recommender, stating that they would not enjoy it if the implementation was as it were in the present experiment.

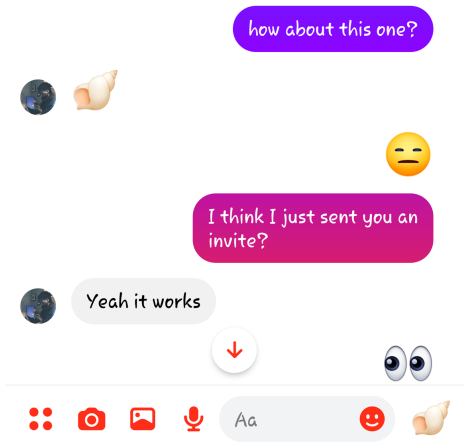


Figure 4.11: Example of customisable "thumb-up" emoji in Facebook Messenger. In this conversation, the custom emoji has been set to the spiral shell emoji 🐚.

Some people would rather only have the frequently/recently used section, especially when they use a small set of emojis regularly and their frequently used section covers them all. Some say that they use emojis uniquely, and that the meaning attributed to emojis is subjective so a recommender would have trouble fitting to them. Another recurring theme was privacy; people did not like the idea that their messages could be used to train a recommender, or to have personalised recommendations as they feel it may be manipulating their behaviour.

Other comments

A few people said they do not like emojis in general, because of how they look, or because of who the people using them are associated with (Multilevel Marketers were mentioned in particular[1]). With regards to emoji recommenders, due to emojis being informal, recommendations may not be suitable in all conversations, and may be intrusive if recommendations always popped up. Emojis used sarcastically would also be difficult for the recommender to pick up. Someone noted that they would prefer if the recommender could eventually recommend a more concise list of emojis after learning their specific preferences.

4.5 Discussion

4.5.1 Differences between With and Without Recommender Conversations

Both emoji counts and variability were higher in the With Recommender conversation compared to the Without Recommender conversation. When typing on phone keyboards, people are not always presented with emoji options (although some keyboards do have limited autocorrect and next-word emoji recommendations). The results suggest that when people are shown emoji options, they might be more likely to use emojis. In the case where people would have wanted to insert an emoji but were too lazy to switch keyboards, a concise selection of recommendations might be useful. On the other hand, constant recommendations take up precious screen space on handheld devices, and may not be suitable for all conversations.

There are a few experimental design factors to take into account when interpreting the results between with and without recommender conversations. First, people were mainly instructed to

[1]Multilevel Marketing (MLM), which are often pyramid schemes, is a business model where participants are paid based on how many new members they can recruit rather than sale commissions.

carry out the experiment on their computers so it is easier to navigate between Discord and the survey website. Although participants were instructed on how emoji insertion works within Discord, they may still be unfamiliar with it thus use emojis less than usual during the without recommender condition. Second, during the with recommender conversation, especially for the Related model, instead of choosing the emojis people would actually send, in some cases, all related emojis were chosen instead. Usually when composing a text message, emoji insertion has direct feedback where you can see how the message looks with the emojis. For the present recommendation implementation, participants did not have this feedback which might have caused the misunderstanding of the task. This would have inflated both emoji count and variability for the with recommender conversation.

4.5.2 Emoji use and variability between different models

Emoji count for the Related model was higher than the Most Used model. While emojis from the Most Used model may mostly be those used to mark illocutionary force or perform a social function, the Related emojis mainly fall under emphasis and aesthetic. The Most Used model generally recommended different emojis that cannot be used interchangeably, while the Related model often recommends many emojis that are thematically or conceptually similar. For example, for the input sentence: "Zoo!", the Most Used model recommends [😇, 🤩, 😊, 😜, 😔, ❣️, 💛, 🙋, ❗], while the Related model recommends [🐆, 🐯, 🐘, 🦬, 🐧, 🦉, 🐊, 🚤, 🎠, 🚝]. While some of the emojis from the Most Used model are similar (💛, ❣️), nuanced differences are likely to be attributed to each, especially for the face emojis. Thus it is unlikely for the participant to choose multiple from the Most Used model. On the other hand, for the Related model, the various animal emojis do not alter the emotional value of the message in different ways, so can be sent together without confusing the sentiment of the message.

It was originally hypothesised that the Most Used model will have a higher emoji count because people generally use these more versus other emojis. As seen from the open questions, the reason people use these more may be because they are near the top of the emoji keyboard, or in their frequently used section, thus easier to find. In other words, some people might want to accompany "I love elephants!" with both 😍 and 🐘, but since the elephant emoji is more difficult to find, they only use 😍. It is also possible that the Related model more reliably recommends at least one emoji that was applicable, while for the Most Used model, there is a higher chance the participant didn't like any of the recommendations. Additionally, the type of conversation and prompts made by Connie may have imitated a usual conversation poorly. Emojis are used most often during informal conversations between friends or family, instead, here Connie is a stranger, and a bot at that.

It was shown that people used a higher variability of emojis in the Related and Combined model than in the Most Used model. One important thing to keep in mind here is that the two models where the participants used a higher variability of emojis, also recommended more varied emojis (due to the Related model being present in both). Looking at the emoji variability of the recommender outputs per conversation (right cluster in Figure 4.12), it can be seen here that the Related and Combined models both recommended a wider variety of emoji than Most Used. Thus it is possible that the variability is due solely to the recommenders. An additional Tukey's post-hoc test was carried out here in order to see if there is a significant difference between selected emoji variability of the Related (M=7.88, SD=4.674) and Combined model (M=5.86, SD=2.825; p=0.249), but no significant difference was found, which does not mirror the variability pattern of the recommenders. As mentioned previously, the added variety in recommendations made by the Most Used model may not translate to an increase in the number of emojis the participant chose because at most participants would pick one or two from the Most Used emojis.

It is interesting then, that for participants given the Combined model condition, emojis which "came from" the Most Used model were chosen more often than those from the Related model.

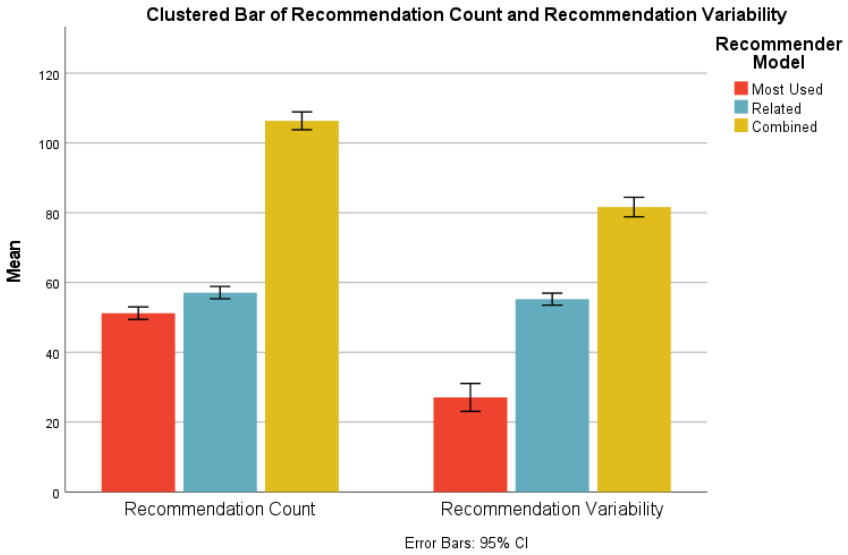


Figure 4.12: Bar graph showing the number of emojis recommended, and the number of unique emojis recommended per recommender condition.

It is possible that the behaviour of the recommenders changed the participants' understanding of the experiment. When more recommendations were given (i.e. around 20 as in the Combined model), participants might be able to better imagine the recommendations appearing on the phone as in Figure 4.3, so that they did not choose all related emojis, but only the ones they would use, in this case those that they commonly use. For the Related model, participants might have been more inclined to choose all that were related. For Most Used, participants might have chosen only those they would use too, but the model recommended less emojis that they would use than Combined. This would support the trend seen for emoji counts too.

4.5.3 Evaluation of recommender between different models

The Related model was rated the highest in quality, followed by the Combined model, with the Most Used model coming in significantly lower that the other two. The Combined model was hypothesised to be rated highest, since it included emojis from both models. The participant might have perceived the Related model as having better recommendations because they could see why the emojis were recommended as they have a connection to the input text. On the other hand, the connection between text and recommendations was not so direct for the Most Used model and some of the emojis for the Combined model. Thus the ratings for the Combined model may have been brought down by the recommendations of the Most Used model.

For novelty, the Related model was rated the highest, followed by the Combined and the Most Used model. It makes sense that the Most Used model is rated the lowest, as the recommendations draw from those emojis that people used most often in general. The Combined model was rated slightly lower (though non-significantly) instead of the same as the Related model, which is interesting considering it had the highest variability in its recommendations and includes all that the Related model would recommend. Perhaps when the Related emojis are alone, participants pay more attention to those that they are less familiar with, while if they are mixed in with emojis that are highly familiar as in the Combined model, participants scan the recommendations quicker, only looking for those they are familiar with and would use.

With regards to whether there were too many recommendations, all models did not have very high means. The Combined model, which generally recommended double the number of emojis as the other two (Figure 4.12 left cluster), still had a mean of 3.14 (SD=1.027), which was "neither agree nor disagree". This suggests that around 20 emojis is a good number

of recommendations. The Most Used model's mean is around the "slightly disagree" choice, suggesting that participants did not mind the number of recommendations, and were perhaps open to a higher number of recommendations too.

4.5.4 Open questions

From the open questions, it can be gathered that participants would use an emoji recommender if it performs just as well or better than the frequently/recently used section current phone keyboards support. Unlike text insertion, emoji insertion often involves additional steps if the emoji you want to insert is not immediately suggested in the autocorrect or next-word suggestion space. Most of the time, it seems that people know what emojis they want to use, the frustration coming from not being able to find it if it's not used often.

4.5.5 Limitations

One of the main caveats when interpreting the data is taking into account the way the experiment was implemented. Having the recommendations appear as reactions, added one by one, is not quite representative of its potential phone keyboard equivalent. Because Discord does not allow you to edit other people's messages, it was not possible to implement the bots to 'insert' the emojis into the participants' message.

Secondly, people were asked to carry out the experiment on a desktop/laptop and not their phone, which means their usual typing experience is not quite the same either. Participants did not have access to their usual "recently used" section as they would have on their phone, or their familiar emoji keyboard. This means that the results for the first conversation might differ from the way participants would normally act when on their phone.

Lastly, the simulated conversation with Connie may not have been a good step-in for a real informal conversation. The ecological validity of the experiment is not the best, so the results here cannot be generalised with certainty to the influence of emoji recommendations during real conversations.

Implementation of Models

The Most Used model was trained on Twitter replies, which are a specific subtype of text messages. The messages sent by participants during the experiment may not have matched up to the training data, and so not all possible categories were recommended. This might have negatively impacted the perception of the model. The types of questions asked by Connie might have been more favourable to the Related model, allowing it to showcase its abilities better than the Most Used model. The maximum number of unique reactions that can be added to any message on Discord was 20. It was decided for the Most Used and Related models to recommend a maximum of 10 emojis, so that the Combined model does not exceed the maximum number of reactions allowed. Each category of the Most Used model recommends five emojis. If a larger number of categories was recommended each time, this might have resulted in a higher variety of recommendations.

A limitation of models based on user data (Most Used and the emoji2vec portion of Related) is that they would not recommend new emojis added by Unicode when they are first introduced. For example, users may not be aware of the pleading face emoji 🥺 when it was first added in 2018, even though it is currently one of the top used emojis on Twitter (Burge, 2020b). A good model should have mechanisms in place so that users can be made aware of new emojis when they are applicable, especially when they are initially added to keyboards.

The models also did not take into account the computation performance. Sometimes it took a while for the recommender to produce recommendations which would not be ideal in an actual

emoji keyboard. Having to wait for recommendations would likely make the users less willing to even have a recommended section of the phone.

4.6 Conclusion

RQ 1 What is the added value of an emoji recommender?

An emoji recommender influences the emojis a person may use, increasing both the number and range of emojis used. Although the current implementation of the emoji recommender is rather different from how it may be integrated into a phone keyboard, the present experiment does give an indication to the effects of emoji recommendations. Most participants indicated that they would probably make use of recommendations if the algorithm performs well. Since many people already make use of the "recently used" section on their emoji keyboards, it is likely that a "recommended" section which accounts for emojis that the user frequently uses would further improve the user's emoji insertion experience, as well as exposing them to emojis that they are unaware of or otherwise difficult to find.

RQ 2 How do the three models compare with each other? Are both the Related and Most Used model important?

The Most Used model on its own was not received as well as the other two models. However, when given emojis from both the Most Used and Related models, participants still chose those from the Most Used model more often. This means that both models are important for the whole set of recommendations to be useful. Recommendations that do not include common emojis at all would mean users have to scroll past the recommendations and still find the emojis they want to use. On the other hand, only including most used emojis does not provide the user with novel emojis and does not have the potential to expand the user's vocabulary.

Chapter 5

Conclusion and Future Work

The purpose of this thesis was to explore emoji recommendation in the context of text messaging. In order to understand emojis better, a literature review was carried out. The literature review resulted in a list of emoji functions which helped structure and motivate the approach for creating the recommendation models. The different emoji functions each have their own implications for how a recommender might be made for them. Two main models were created, one which recommends emojis related to the text input (Related model), and another which recommends commonly used emojis (Most Used model). A third model was also created which combines the outputs of the two models (Combined model). These three models were evaluated in a Discord-based experiment.

So, what are the main findings with regards to the research question: **What makes a successful emoji recommender, and how can emoji recommendations influence user's emoji behaviour?**

A summary of the literature on emoji functions resulted in six main functions: Emphasis emojis which repeat concepts mentioned in the text (e.g. "Please do the dishes 🍽️"), illocutionary emojis which clarify the intent of the text (e.g. "Please do the dishes 😤" versus "Please do the dishes 😇"), social emojis that perform communicative acts such as conversation management or backchannelling (e.g. "👋 Please do the dishes"), content emojis which replace text (e.g. Please 🧽🧼 the 🍽️), aesthetic emojis which add decorative elements (e.g. "Please do the dishes 🌻"), and finally reaction emojis which are often stand alone emoji replies (e.g. A: "Please do the dishes" B: "🙏🙏"). Often an emoji might serve multiple functions at once, for instance, aesthetic emojis may not be related to the text, but by their presence alone they may make a message more positive and serve an illocutionary function.

The perfect emoji recommender would be able to cover all the functions, however, some of the functions are difficult to account for without being able to read the user's mind. This is especially the case for content emojis, social emojis at the start of a sentence, and reaction emojis if the recommender cannot read the context messages. Illocutionary emojis are also not so straightforward. The two examples given for illocutionary emojis above, for instance, read differently when considering the emojis. The face with steam from nose emoji 😤 makes the sender sound frustrated, while the smiling face with halo emoji 😇 may suggest a genuine request. Illocutionary emojis are the most common function for emojis, however, so a recommender that does not attempt to cover these would not be useful a large portion of the time.

The two main models built for the thesis aims to cover illocutionary emojis (Most Used model) and emphasis emojis (Related model). The Most Used model was trained on tweet replies that included a GIF. In the dataset, the GIF was categorised into broad categories (e.g. "popcorn", "thumbs up"). Although these are GIF categories, in the context of tweet replies, they also mark the illocutionary force of the text portion of the tweet. The Most Used model recommends emojis that are most commonly used by people across social media platforms. The Related model, on the other hand, uses semantic relations between words as well as emoji

vectors to give recommendations. This model recommends emojis which are close in meaning to the text input, or otherwise connected to it in some way (e.g. “bird” includes [🦢, 🐤, 🦉, 🦆]). The third Combined model outputs emojis from both of the previous models.

A Discord text-based experiment was designed to investigate how emoji recommenders may influence the way participants used emojis, as well as comparing the three models against each other. Participants held two short conversations with a chatbot, one without emoji recommendations, and one with recommendations provided by one of the models. Having an emoji recommender did increase the participants’ emoji use and the variety of emojis they used. However, since the mode of emoji insertion was unequal between the without and with recommender conversations, this should be interpreted carefully, as much of the effect may be explained by the availability of emojis as well as participants’ interpretation of the instructions and their own hypothesis about the goal of the experiment.

The Most Used model was not perceived as well as the Related and Combined models, whereas the latter two performed similarly well. The two latter models were evaluated as having higher quality than Most Used. The Related model was rated as the highest in novelty too. On the other hand, even with 20 recommendations, it was not rated as too many.

Including a recommended section on an emoji keyboard would be beneficial for the user, especially if the section was personalised for the user, although this should be something the user could choose as not everyone is comfortable with machine learning being carried out on their activity. The user’s preference for privacy should be taken into account when implementing recommendations. It is possible to be explicit about the type of data used for the models. For the Related model, for example, user data does not necessarily have to be involved. Perhaps the participants who were worried about their privacy would still be comfortable with related suggestions in addition to a recently used section.

5.1 Future work

5.1.1 Improvements to the models

High amounts of data are available on public platforms such as Twitter or Facebook, however, these virtual spaces are not equal to spaces where text messaging takes place. The mode of interaction is different so the way people talk (or rather, type) is different. A good place to start with improving models would be to collect more valid training data. One of the reasons why Twitter data is so commonly used is because users do not need to give explicit consent for their tweets or replies to tweets to be scraped. On the other hand, it would be a even larger invasion of privacy if the same were done to private messages.

Since a large number of recommendations can be given at a time, it is not crucial for recommendations to be highly accurate as long as some of the recommendations are useful. Thus it would be interesting to look at different language modelling approaches where a large amount of text message data is not needed. The emoji2vec models trained by Eisner et al. (2016) were based on emoji descriptions provided by Unicode, this alone was quite successful in producing related (object) emojis. Implementing (emotional) metaphors into the emoji senses dictionary could be a good place to start. For example, the volcano emoji 🌋 could be linked to words related to anger. On the other hand, linking a concept to too many emojis is also not ideal. For example, while there are numerous emojis related to trains 🚆🚄🚅🚃🚂🚝, it is unnecessary for a recommender to recommend all of them all the time. Knowing which train to recommend for the situation is a big problem to tackle.

For the Most Used/illocutionary emojis, the present model seem to bias certain categories. This may be due to an imbalance in the training data. There are possibly other machine learning models or approaches that can deal with data imbalance better. Alternatively, understanding illocutionary emojis better would also be an interesting approach. Carrying out an analysis on

when and in what context the most frequently used emojis are used can provide insight into what features are important to look at in order to train a better recommendation model.

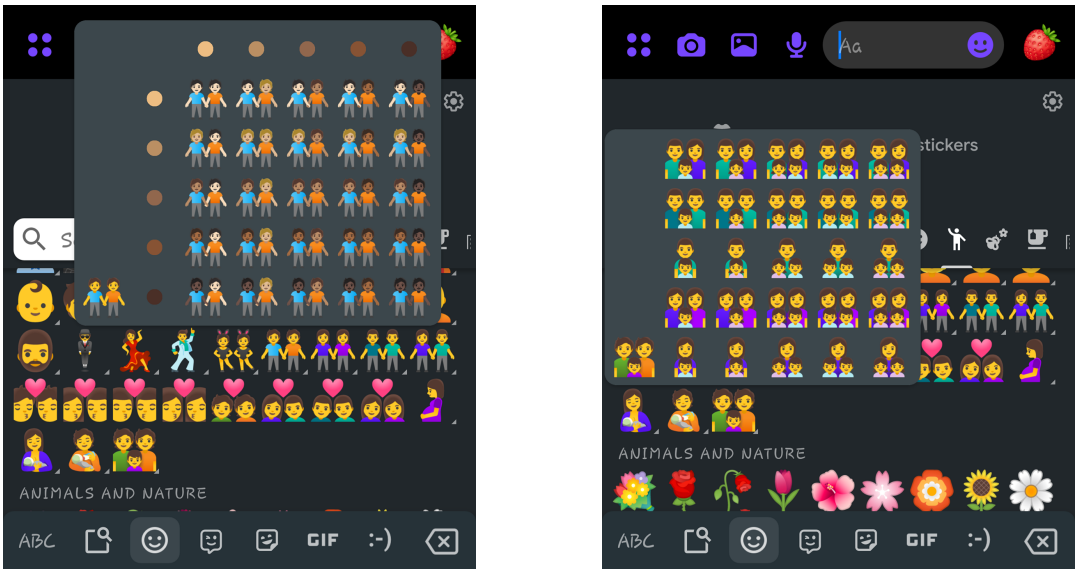
5.1.2 Implementing recommendations within keyboard

One of the major limitations of the present thesis is the implementation of the recommendations not being ecologically valid. It would be valuable to see how emoji behaviour would change if the recommendations were to be implemented into a functional phone keyboard. One of the motivations for recommendations was to expand the user's emoji vocabulary. This could be further investigated over a longitudinal study where an emoji keyboard with extensive recommendations are compared with a usual emoji keyboard.

5.1.3 Other methods of emoji insertion

Some participants mentioned that they liked Discord's method of emoji insertion where they type out the emoji instead of selecting it. It might be interesting to consider a "mode switch" key on a phone keyboard that allows you to search for emojis using the alphabetical keyboard straight away without having to 1. switch to the emoji keyboard, 2. click on the search bar, 3. type search terms.

Currently on Gboard, if you want to insert emojis that may be close to each other conceptually but far away on the keyboard (i.e. in different sections), you must browse through the keyboard or try different search terms. It may be interesting to explore having a next-emoji suggestion bar while on the emoji keyboard. So for instance, if I insert the vampire emoji 🧛, the blood emoji 🩸, bat 🦇, wolf 🐺, and moon 🌕 emojis could be suggested.



(a) Gboard People holding hands emoji

(b) Gboard Family

Figure 5.1: Gboard user interface for left: selecting various combinations of skin tones and right: family composition.

Unicode has been adding more ZWJ emoji sequences recently. With the large number of skin tone and gender combinations possible, the user interface can quickly become overwhelming. Figure 5.1 shows the Gboard interface for skin tone and family composition selection. Currently family emojis do not have options for skin tones of each family member, but the current interface also has no space for something like this. Additionally, Unicode has been adding more emojis that are ZWJ sequences instead of having a code point of their own, such as the black cat 🐈‍⬛ being a combination of 🐈 and ⬛, it is not a stretch to imagine more coloured cats being

added. Furthermore, Gboard has been experimenting with custom emoji combinations. This "Emoji Kitchen" feature allows the user to combine different emojis together to create new emojis (Daniel, 2020). Figure 5.2 shows the combination 🐢 and ☕, as well as other possible combinations. Emoji kitchen creations are inserted into messaging apps as images instead of as emojis since currently these combinations are not in Unicode. However, it is possible that platforms might be willing to support more ZWJ sequences even when they're not in Unicode. In that case, the current emoji keyboard will not be the best way to browse through or explore these combinations.

5.2 Final Note from the Author

Emojis continue to be on the rise. While they may not make their way into all avenues of digital writing, their place in informal social networking websites and instant messaging can only become more complex. As an emoji enthusiast, this is all very exciting and I'm looking forward to the addition of new emojis and tools to assist users in exploring them.

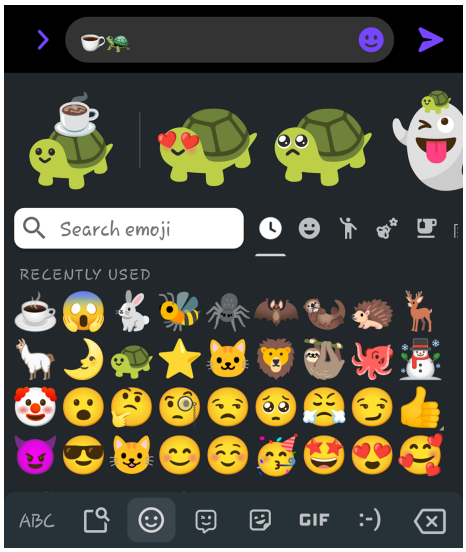


Figure 5.2: An example of Gboard's Emoji Kitchen. Typing the coffee emoji and the turtle emoji in sequence ☕🐢 results in a combination of a turtle carrying a cup of coffee.

References

Al Rashdi, F. (2015). *Forms and functions of emojis in whatsapp interaction among omanis* (Unpublished doctoral dissertation). Georgetown University.

Al Rashdi, F. (2018). Functions of emojis in whatsapp interaction among omanis. *Discourse, context & media*, *26*, 117–126.

Alvina, J., Qu, C., Mcgrenere, J., & Mackay, W. E. (2019). Mojiboard: Generating parametric emojis with gesture keyboards. In *Extended abstracts of the 2019 chi conference on human factors in computing systems* (pp. 1–6).

Anna, B. (2016). Recommender systems-it's not all about the accuracy. *Gab41, Gab41*, *52*.

Austin, J. (1962). How to do things with words, 2nd edn, jo urmson and m. *Sbasa (eds)*.

Barbieri, F., Ballesteros, M., & Saggion, H. (2017). Are emojis predictable? *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers*.

Barbieri, F., Marujo, L., Karuturi, P., Brendel, W., & Saggion, H. (2018). Exploring emoji usage and prediction through a temporal variation lens. *arXiv*, arXiv–1805.

Barbieri, F., Ronzano, F., & Saggion, H. (2016). What does this emoji mean? a vector space skip-gram model for twitter emojis. In *Calzolari n, choukri k, declerck t, et al, editors. proceedings of the tenth international conference on language resources and evaluation (lrec 2016); 2016 may 23-28; portorož, slovenia. paris: European language resources association (elra); 2016. p. 3967-72*.

Barrett, L. F. (2017). The theory of constructed emotion: an active inference account of interoception and categorization. *Social Cognitive and Affective Neuroscience*, *12*(1), 1–23.

Burge, J. (2019a, February 5). 230 new emojis in final list for 2019. *Emojipedia*. Retrieved 14-01-2020, from `https://blog.emojipedia.org/230-new-emojis-in-final-list-for-2019/`

Burge, J. (2019b, March 8). Correcting the record on the first emoji set. *Emojipedia*. Retrieved 04-12-2019, from `https://blog.emojipedia.org/correcting-the-record-on-the-first-emoji-set/`

Burge, J. (2020a, January 29). 117 new emojis in final list for 2020. *Emojipedia*. Retrieved 14-01-2020, from `https://blog.emojipedia.org/117-new-emojis-in-final-list-for-2020/`

Burge, J. (2020b, April 24). A new king: Pleading face. *Emojipedia*. Retrieved 14-12-2020, from `https://blog.emojipedia.org/a-new-king-pleading-face/`

Cramer, H., de Juan, P., & Tetreault, J. (2016). Sender-intended functions of emojis in us messaging. In *Proceedings of the 18th international conference on human-computer interaction with mobile devices and services* (pp. 504–509).

Dainas, A., & Herring, S. C. (2019). Interpreting emoji pragmatics. *Internet pragmatics: Theory and practice. Amsterdam: John Benjamins. Prepublication version: http://ella. ils. indiana. edu/~ herring/Interpreting_Emo ji_Pragmatics. pdf*.

Danesi, M. (2016). *The semiotics of emoji: The rise of visual language in the age of the internet*. Bloomsbury Publishing.

Daniel, J. (2020, December 3). Emoji kitchen cooks up a new batch of mashups. *Google Blog*. Retrieved 15-12-2020, from `https://blog.google/products/android/emoji-kitchen-new-mashups-mixing-experience/`

Dresner, E., & Herring, S. C. (2010). Functions of the nonverbal in CMC: Emoticons and illocutionary force. *Communication theory*, *20*(3), 249–268.

Eisner, B., Rocktäschel, T., Augenstein, I., Bošnjak, M., & Riedel, S. (2016). Emoji2vec: Learning emoji representations from their description. *arXiv preprint arXiv:1609.08359*.

Ekman, P. (1999). Basic emotions. *Handbook of cognition and emotion*, *98*(45-60), 16.

El Ali, A., Wallbaum, T., Wasmann, M., Heuten, W., & Boll, S. C. (2017). Face2emoji: Using facial emotional expressions to filter emojis. In *Proceedings of the 2017 chi conference extended abstracts on human factors in computing systems* (pp. 1577–1584).

Franco, C. L., & Fugate, J. M. (2020). Emoji face renderings: Exploring the role emoji platform differences have on emotional interpretation. *Journal of Nonverbal Behavior*, 1–28.

Gawne, L., & McCulloch, G. (2019). Emoji as digital gestures. *Language@ Internet*, *17*(2).

Guibon, G., Ochs, M., & Bellot, P. (2018). Emoji recommendation in private instant messages. In *Proceedings of the 33rd annual acm symposium on applied computing* (pp. 1821–1823).

Herring, S., & Dainas, A. (2017). ″Nice picture comment!″Graphicons in Facebook comment threads. In *Proceedings of the 50th hawaii international conference on system sciences.*

Holtgraves, T., & Robinson, C. (2020). Emoji can facilitate recognition of conveyed indirect meaning. *PloS one*, *15*(4), e0232361.

Ismaeil, Y., Balalau, O., & Mirza, P. (2019). Discovering the functions of language in online forums. In *Proceedings of the 5th workshop on noisy user-generated text (w-nut 2019)* (pp. 259–264).

Jaeger, S. R., & Ares, G. (2017). Dominant meanings of facial emoji: Insights from chinese consumers and comparison with meanings from internet resources. *Food Quality and Preference*, *62*, 275–283.

Jaeger, S. R., Roigard, C. M., Jin, D., Vidal, L., & Ares, G. (2019). Valence, arousal and sentiment meanings of 33 facial emoji: Insights for the use of emoji in consumer research. *Food Research International*, *119*, 895–907.

Jakobson, R. (1960). Linguistics and poetics. In *Style in language* (pp. 350–377). MA: MIT Press.

Kim, J.-G., Gong, T., Huang, E., Kim, J., Lee, S.-J., Kim, B., … Ko, J. (2019). Bringing context into emoji recommendations (poster). In *Proceedings of the 17th annual international conference on mobile systems, applications, and services* (pp. 514–515).

Liebeskind, C., & Liebeskind, S. (2019). Emoji prediction for Hebrew political domain. In *Companion proceedings of the 2019 world wide web conference* (pp. 468–477).

Lin, W., Chao, T., Wu, J., & Su, T. (2019). Predict emoji combination with retrieval strategy. *arXiv*, arXiv–1908.

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought.* University of Chicago press.

Mikolov, T., Yih, W.-t., & Zweig, G. (2013). Linguistic regularities in continuous space word representations. In *Proceedings of the 2013 conference of the north american chapter of the association for computational linguistics: Human language technologies* (pp. 746–751).

Miller, H., Thebault-Spieker, J., Chang, S., Johnson, I., Terveen, L., & Hecht, B. (2016). ″Blissfully happy″or ″ready to fight″: Varying interpretations of emoji. In *Tenth international AAAI conference on web and social media.*

Na′aman, N., Provenza, H., & Montoya, O. (2017). Varying linguistic purposes of emoji in (twitter) context. In *Proceedings of ACL 2017, student research workshop* (pp. 136–141).

Novak, P. K., Smailović, J., Sluban, B., & Mozetič, I. (2015). Sentiment of emojis. *PloS one*, *10*(12), e0144296.

Pardes, A. (2018, February 1). The WIRED guide to emoji. *Wired*. Retrieved 04-12-2019, from

https://www.wired.com/story/guide-emoji/

Pavalanathan, U., & Eisenstein, J. (2015). Emoticons vs. emojis on twitter: A causal inference approach. *arXiv preprint arXiv:1510.08480*.

Pohl, H., Stanke, D., & Rohs, M. (2016). Emojizoom: emoji entry via large overview maps😄🔍. In *Proceedings of the 18th international conference on human-computer interaction with mobile devices and services* (pp. 510–517).

Pu, P., Chen, L., & Hu, R. (2011). A user-centric evaluation framework for recommender systems. In *Proceedings of the fifth ACM conference on Recommender Systems* (pp. 157–164).

Riordan, M. A. (2017). The communicative role of non-face emojis: Affect and disambiguation. *Computers in Human Behavior*, *76*, 75–86.

Rodrigues, D., Lopes, D., Prada, M., Thompson, D., & Garrido, M. V. (2017). A frown emoji can be worth a thousand words: Perceptions of emoji use in text messages exchanged between romantic partners. *Telematics and Informatics*, *34*(8), 1532–1543.

Russell, J. A. (1980). A circumplex model of affect. *Journal of personality and social psychology*, *39*(6), 1161.

Solomon, J. (2020, February 5). Emoji spelling is peachy🍑. *Emojipedia*. Retrieved 01-05-2020, from https://blog.emojipedia.org/emoji-spelling-is-peachy/

The Unicode Consortium. (n.d.). *The most frequent emoji*. Retrieved 16-01-2020, from https://home.unicode.org/emoji/emoji-frequency/

The Unicode Consortium. (2020a, July 23). *Emoji counts, v13.0*. Retrieved 19-01-2021, from https://www.unicode.org/emoji/charts-13.0/emoji-counts.html

The Unicode Consortium. (2020b, January 29). *Full emoji list, v13.0*. Retrieved 06-03-2020, from https://unicode.org/emoji/charts/full-emoji-list.html

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

Wijeratne, S., Balasuriya, L., Sheth, A., & Doran, D. (2017). Emojinet: An open service and api for emoji sense discovery. In *Eleventh international aaai conference on web and social media*.

Wiseman, S., & Gould, S. J. (2018). Repurposing emoji for personalised communication: Why🍕 means "I love you". In *Proceedings of the 2018 CHI conference on human factors in computing systems* (p. 152).

Xie, R., Liu, Z., Yan, R., & Sun, M. (2016). Neural emoji recommendation in dialogue systems. *arXiv*, arXiv–1612.

Ziegler, C.-N., McNee, S. M., Konstan, J. A., & Lausen, G. (2005). Improving recommendation lists through topic diversification. In *Proceedings of the 14th international conference on world wide web* (pp. 22–32).

Appendix A

Stop Words list

i	which	at	both	aren
me	who	by	each	aren't
my	whom	for	few	couldn
myself	this	with	more	couldn't
we	that	about	most	didn
our	that'll	against	other	didn't
ours	these	between	some	doesn
ourselves	those	into	such	doesn't
you	am	through	no	hadn
you're	is	during	nor	hadn't
you've	are	before	not	hasn
you'll	was	after	only	hasn't
you'd	were	above	own	haven
your	be	below	same	haven't
yours	been	to	so	isn
yourself	being	from	than	isn't
yourselves	have	up	too	ma
he	has	down	very	mightn
him	had	in	s	mightn't
his	having	out	t	mustn
himself	do	on	can	mustn't
she	does	off	will	needn
she's	did	over	just	needn't
her	doing	under	don	shan
hers	a	again	don't	shan't
herself	an	further	should	shouldn
it	the	then	should've	shouldn't
it's	and	once	now	wasn
its	but	here	d	wasn't
itself	if	there	ll	weren
they	or	when	m	weren't
them	because	where	o	won
their	as	why	re	won't
theirs	until	how	ve	would
themselves	while	all	y	wouldn
what	of	any	ain	wouldn't

Appendix B

Emoji2Vec results

Tokens	20 top Sum	Sum > 0.35	Per Token
[hello, peeps, like, captain, one, objects]	[👋 😜 😂 🤦 😚 😉 😆 👤 🤦 😃 😱 👅 🎅 ♏ 😗 💑 😆 👎 😀 🙌]	[👋, 😜, 😂, 🤦, 😚, 😉, 😆, 👤, 🤦, 😃, 😱, 👅, 🎅, ♏, 😗, 💑, 😆, 👎, 😀, 🙌]	[[👋, 😂, 🤗, 😃, 😀, 😆, 😅, 😊, 😚, 😻, 😉, 😊, 😏, 😹, 😄, 💑, 😗, 😹, 🤝], [🤦], [👍, 💭, 🤦, 🙈], [🏉], [1 , 1 , 👨👨👧👦, 🔞, 👩👩👦, 👨👨👧, 👩👩👧, 👨👨👦, 🙋, 🙋, 🙋, 🙋, 👨👩👧👦, 🕐, 🙋, 👨👩👧, 🍀, 👨👨👧👧, 3 , 3], []]
[also, really, good, ghost, story, times]	[👍 😔 🤦 👤 🙋 💭 🤦 🎋 💛 🙈 📰 😕 🎅 🙋 👆 🙂 💚 ❤️ 😞 🎃]	[👍, 😔, 🤦, 👤, 🙋, 💭, 🤦, 🎋, 💛, 🙈, 📰, 😕, 🎅, 🙋]	[], [💭, 🤦, 😔, 👍, 🙈, 🎃, 💛, 😂, 🤦, 💑], [👆, 🤦, 🤦, 👍, 🤦, 🤦, 🤦, 👌, 💯, 💛, 😔], [👻, 🐉, 🎃, 👺], [😴], []]
[touch, cursed, paper, currency]	[💵 💶 💷 💴 $ 💱 💸 🐼 👎 🦾 👊 ☸ 👊 🙏 🇬🇧 🐵 👋 😟 🚾 😞]	[💵, 💶, 💷, 💴, $, 💱]	[[], [😟, 😨, 👺, 👹], [💵, 💶, 💷, 💴, 📒, 📝], [💱, $, 💷, 💶, 💴, 💵]]
[want, go, home, raining, bike, rain]	[🚌 🐫 🌧 🏠 🐫 🚳 🌄 🚧 💸 🔪 🚵 🌉 🚻 🚴 😅 🌉 🎃 🏃 📼 💧]	[🚌, 🐫, 🌧, 🏠, 🐫, 🚳, 🌄, 🚧, 💸, 🔪, 🚵, 🌉, 🚻, 🚴]	[[🙏, 🚳, 🚮, 🤦], [📼, 🚦, 🎽], [🏠, 🌏, 🏠, 🌍, 🏡, 🌐, 🌎], [🚌, 🐫, 🏠, 🌧, 😅, 🌋], [🚴, 🚲, 🚵, 🏍, 🛴, 🚳, 🛵, 🚙, 🚌, 🚗, 🚘, 🛥, 🚤, 🎡, 🎠, 🚐], [🚌, 🌧, 🌧, 🔪, 🚌, 🏠, 🌦, 🌨, 🌀, 🌧, ☃]]
[enjoy, concert]	[🎻 🎋 🥂 🍾 🎡 🎁 🎧 🎷 🎪 🎂 🎅 🎻 🎶 🎑 🎭 😄 🎵 🎰 🎺 🎆]	[🎻]	[[😊], [🎷, 🎸, 🎻]]
[good, ideas]	[🎋 🤦 👆 😔 👍 🤦 🤦 🤦 💛 🤦 🤦 💡 💭 👅 😃 😆 🎆 😹 😊 👏]	[]	[[👆, 🤦, 🤦, 👍, 🤦, 🤦, 🤦, 👌, 💯, 💛, 😔], []]
[ye, sure, otherwise]	[🙏 🤦 🍀 🎅 🎅 🎃 😈 😦 😅 🎈 🎭 🏰 🎎 🎅 🆗 👌 🐼 👌 ℹ 🙂]	[🙏]	[[], [💭, 🆑, 🤦], []]

[yes, wait, vegetables]	[😋🍅🥓🍜🍆🧑🍣🌿🍉🍰🤦👅🍶🍤🎃🍦🍭🍥🎋🌰]	[😋, 🍅, 🥓, 🍜, 🍆, 🧑]	[[🧑, 🤦, 👎, 🙌, 😆, 😂, 😜, 🕴, 😉, 🤦, 👌], [], [🍅, 🌿, 🍜, 🥓, 🍏, 🍐, 🍤, 🥔, 🍤, 🍓, 🍆, 🌰, 🍇, 🍉, 🌾, 🍄, 🍉, 🥒, 🍚, 🍢]]
[christmas, songs, thing, powering, november, hahah]	[🎄〽️🎃🤦🎅🙌🎮🕴🤣🎅💿📀😂🌊🎧🍀🏈🌲🎌🇬🇧]	[🎄, 〽️, 🎃, 🤦, 🎅, 🙌, 🎮, 🕴, 🤣, 🎅, 💿, 📀, 😂, 🌊, 🎧, 🍀, 🏈, 🌲, 🎌, 🇬🇧]	[[🎄, 🎅, 🌐, 🌲, 🎅, 🎅, 🎅, 🎅, 🎃, 🌷, 🎅, 🎌, 🎅, 👻, 📀], [🎶, 🎵, 🎷, 🎻, 〽️, 🎧, 🎺, 🎹, 🎸, 🎤, 💿], [🗨️, 🈯], [], [📅, 🖥️], [🤦, 🕴, 🤣, 🌊, 😂, 🙌, 👺, 💩, 🤦, 😈, 🇳🇴, 🤦, 🤦, 😋, 😉, 👽, 🧑, 🍟, 🆗, 🤦]]
[let, know, leave, campus]	[🙏🤦🐵🈯🧑😥🧑🏫🥊👎😔🤦😔👌🤦👌🤦🤦🙋🗨️]	[🙏, 🤦, 🐵, 🈯, 🧑]	[[🙏, 🤦, 👎], [🗨️, 🤦, 🐵, 😅, 🧑, 🙏, 😰, 🤦, 😔, 😞], [], [🏫, 🚌]]

Appendix C

Emoji Senses results

EmojiNet Recommendations	Ontological Recommendations
hello: [😊, 👋], peeps: [], like: [😄], captain: [], one: [🕐, 1], objects: []	hello: [], peeps: [], like: [], captain: [👮‍♀️, 🧑‍✈️, 🏏], one: [], objects: []
also: [], really: [], good: [📈, 👍], ghost: [👻, 😱, 👺], story: [👼, 😇, 🔮, 😈, 👾, 👽, 👻, 👺, 👹, 👩‍🎨], times: [❌, ✖️]	also: [], really: [], good: [🙆‍♀️], ghost: [👻, 🧑‍🎤, 🧑‍🎨, 🏔️, ⛰️, 🚡], story: [], times: []
touch: [], cursed: [], paper: [🗞️, 📰, 📘, 📑, 🖨️, 📝, 📜, 📋, 📄, 📃], currency: [💹, 💷, 💶, 💵, 💶, $, 💱]	touch: [🤗, 🤹‍♀️, 👭, 👭, 👬, 💏, 🏉], cursed: [], paper: [🎊, 📃, 📄, 📰, 🗞️], currency: [💶]
want: [💚, 🙏], go: [🏃‍♀️, 🚶‍♀️, 🚳, 👣, 👼, 😳, 🍆, 💀, 🚆, 🚆, , 🚡, 🚌, 🏭, 🏗, 🚚, 🚛, 🚕, 🚐, 🚌, 🚋, 🚆, 🚄, 🚅, 🚇, 🚝, 🚞, 🚂, 🚈, 🚅, 🚂, 🚀, 🛫, ➡️, 📦, 💱, 🗾, ⬅️, 🌐, 🌍], home: [🏠, 🚪, 🏛️, 🏘️, 🏰, 🏯, 🏪, 🏫, 🏡, 🏠, 🏚️, 🏚️], raining: [💦, 🌧️, ☔, ☂️, 🌧️, 🌦️, 🌂], bike: [🚵‍♀️, 🚴‍♀️, 🚳, 🚲, 🏍], rain: [💦, 🌧️, ☔, ☂️, 🌧️, 🌦️, 🌂, 💧]	want: [], go: [🏃‍♀️, 🥁], home: [🛏️, 🏘️, 🏠, 🏡, 🏢, 🏬, 🏰, 🏯, 🚢, 🚪, 🛋], raining: [🌧️, 🌦️, 🌧️, 🌧️], bike: [🚴‍♀️, 🚵‍♀️, 🛵, 🚲], rain: [🌧️, 🌦️, 🌧️, 🌧️]
enjoy: [], concert: [🎟, 🎫]	enjoy: [], concert: []
good: [📈, 👍], ideas: [💭, 💬, 💡]	good: [🙆‍♀️], ideas: [🤔]
ye: [, sure: [💛], otherwise: []	ye: [], sure: [], otherwise: []
yes: [👍], wait: [🚦], vegetables: [🥜, 🥕, 🥔, 🥒, 🍀, 🍆, 🍅, 🍄, 🍀, 🍋]	yes: [], wait: [], vegetables: [🌿, 🍏, 🍅, 🍆, 🥔, 🌶️, 🥬, 🥦]
christmas: [🎅, 🎅, 🎄, 🎁], songs: [], thing: [📿, 👢, 🥿, 👠, 🥾, 🥾, 🥿, 👜, 👛, 👚, 👽, 👔, 👗, 👖, 👕, 👔, 👓, 🧢, 👑, 🎩, 🎓], powering: [], november: [], hahah: []	christmas: [], songs: [🗣️, 🥪, 📣, 🎶], thing: [🧑‍🎨, 🥊, 🧤, 🧥, 👗, 👔, 👚], powering: [], november: [], hahah: []
let: [], know: [👌, ♋, 🍆], leave: [🍁, 🏃‍♀️, 🚶‍♀️, 🚳, 👣], campus: []	let: [], know: [🤔], leave: [🏃‍♀️], campus: [🏕]

Appendix D

Most Used model Factor Analysis

Category	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7	Factor 8	Factor 9	Factor 10	Factor 11	Factor 12	Factor 13	Factor 14	Factor 15
sigh	0.962	0.004	0.012	-0.108	0.071	0.049	0.054	0.067	-0.050	-0.132	-0.018	-0.023	-0.056	-0.027	0.081
eye_roll	0.961	-0.094	-0.077	-0.044	0.069	0.047	0.053	0.037	-0.033	0.093	-0.045	-0.030	-0.028	-0.030	-0.053
smh	0.963	0.008	-0.023	-0.100	0.085	0.062	0.060	0.035	-0.041	0.082	-0.083	-0.050	0.005	-0.043	0.005
facepalm	0.935	0.032	0.038	-0.111	0.092	0.069	0.067	0.034	-0.068	0.054	-0.069	-0.053	-0.058	-0.050	0.008
seriously	0.956	0.010	0.044	-0.104	0.092	0.068	0.068	0.028	-0.077	0.042	-0.060	-0.046	-0.002	-0.044	-0.006
no	0.959	-0.129	0.058	-0.099	0.072	0.051	0.053	0.026	-0.068	0.020	-0.042	-0.034	-0.002	-0.036	-0.033
thumbs_down	0.950	-0.095	-0.085	-0.077	0.066	0.046	0.047	0.026	-0.051	0.098	-0.056	-0.030	-0.015	-0.030	-0.082
yawn	0.913	-0.120	-0.053	-0.079	0.064	0.044	0.047	0.019	-0.007	0.091	-0.078	-0.041	0.025	-0.036	-0.018
deal_with_it	-0.083	0.939	-0.040	-0.045	0.100	0.063	0.084	0.100	-0.030	0.075	-0.066	-0.042	0.041	-0.042	0.048
sorry	0.045	0.928	0.038	0.051	0.103	0.072	0.105	0.006	0.043	0.005	-0.070	-0.040	0.038	-0.053	-0.306
oops	-0.114	0.808	-0.136	0.102	0.153	0.120	-0.176	-0.012	-0.145	0.086	-0.065	-0.107	0.067	-0.081	0.422
yolo	-0.116	0.922	-0.037	-0.063	0.113	0.079	0.080	-0.195	-0.041	0.039	-0.032	-0.040	0.044	-0.045	-0.011
oh_snap	-0.061	-0.132	0.962	-0.068	0.067	0.043	0.077	0.085	-0.029	0.052	-0.029	-0.013	0.015	-0.014	0.063
scared	-0.086	0.209	0.908	0.006	0.071	0.073	0.084	0.060	0.000	-0.131	-0.093	-0.070	0.026	-0.071	-0.264
shocked	-0.009	-0.162	0.699	-0.131	0.096	0.075	0.076	-0.336	-0.140	-0.096	-0.045	-0.063	0.053	-0.064	0.085
win	-0.464	-0.248	-0.236	0.664	0.094	0.127	0.190	0.140	-0.047	0.060	0.186	0.103	0.120	0.223	0.070
awww	-0.309	0.319	0.023	0.795	0.042	0.076	-0.029	0.067	0.040	-0.166	0.011	-0.070	0.030	-0.065	-0.343
agree	-0.260	-0.110	-0.148	0.912	0.080	0.043	0.103	0.021	-0.001	0.034	0.137	0.006	0.046	0.028	0.149
kiss	-0.196	-0.121	-0.076	-0.030	-0.728	0.038	-0.181	0.054	-0.081	-0.088	-0.020	-0.088	0.005	-0.030	-0.061
want	-0.153	-0.100	-0.045	-0.105	-0.965	0.036	0.066	0.041	-0.044	-0.040	-0.039	-0.051	0.014	-0.039	0.012
you_got_this	-0.255	-0.161	-0.120	-0.015	-0.786	0.047	0.135	0.020	0.005	0.045	-0.072	0.177	0.030	0.001	0.042
good_luck	-0.221	-0.126	-0.100	-0.056	0.065	-0.948	0.059	0.031	0.028	0.008	-0.015	0.009	0.026	0.112	-0.001
hug	-0.184	-0.122	-0.085	-0.069	0.052	-0.963	0.040	0.027	-0.041	-0.030	-0.033	-0.032	0.017	0.028	-0.005
wink	-0.262	-0.111	-0.119	-0.077	0.073	0.030	-0.843	0.070	-0.007	0.059	-0.068	0.023	0.023	-0.038	-0.036
hearts	-0.217	-0.053	-0.108	-0.084	-0.062	0.070	-0.934	0.064	0.011	0.139	-0.049	-0.076	0.031	-0.038	0.056
applause	-0.318	-0.265	-0.204	0.160	0.148	0.136	0.156	0.155	0.245	-0.121	-0.109	-0.174	0.094	-0.155	0.052
dance	-0.264	-0.241	-0.152	-0.067	0.113	0.121	0.101	0.146	0.010	-0.110	0.243	-0.043	0.069	-0.008	-0.034
happy_dance	-0.288	-0.272	-0.013	0.245	0.135	0.123	0.126	0.124	-0.198	-0.069	0.280	-0.051	0.086	-0.043	0.033
thumbs_up	-0.388	-0.037	-0.112	0.189	0.071	-0.201	0.101	0.074	0.467	0.357	0.509	0.270	0.035	0.153	0.024
idk	0.791	0.428	0.059	-0.128	0.116	0.094	0.097	0.062	-0.096	-0.044	-0.089	-0.080	0.021	-0.076	-0.001
yes	-0.373	-0.134	-0.170	-0.025	0.147	0.054	-0.035	0.059	0.840	0.044	0.209	0.094	0.093	0.050	-0.035
slow_clap	-0.253	-0.162	-0.118	0.149	0.079	0.046	0.072	0.059	0.163	-0.027	0.690	-0.042	0.025	-0.067	-0.005
high_five	-0.300	-0.177	-0.143	-0.002	-0.003	0.034	0.053	0.059	0.096	0.030	-0.026	0.912	0.047	0.038	-0.004
do_not_want	0.019	-0.124	-0.067	-0.098	0.045	0.040	0.049	0.050	-0.067	-0.042	-0.046	-0.038	-0.975	-0.038	-0.003
fist_bump	-0.299	-0.192	-0.154	0.071	0.082	-0.159	0.086	0.038	0.060	0.018	-0.058	0.043	0.049	0.888	-0.002
eww	0.706	0.068	0.548	-0.155	0.111	0.079	0.069	0.037	-0.110	-0.205	-0.033	-0.051	0.035	-0.050	0.235
please	0.250	0.234	-0.219	-0.059	0.115	0.054	-0.315	0.032	0.063	0.836	-0.025	0.029	0.060	0.009	0.018
shrug	0.738	0.531	0.010	-0.099	0.115	0.086	0.070	0.025	-0.063	-0.082	-0.077	-0.063	0.011	-0.056	0.066
omg	0.641	-0.096	0.550	-0.160	0.090	0.080	0.093	-0.003	-0.061	-0.112	-0.085	-0.068	-0.024	-0.051	0.092
ok	-0.218	-0.049	-0.304	0.422	0.121	-0.053	0.192	-0.480	0.251	0.282	-0.179	0.055	0.105	0.137	0.057
mic_drop	-0.213	0.096	0.074	-0.125	0.089	0.068	0.095	-0.888	-0.081	-0.057	-0.066	-0.061	0.034	-0.052	-0.011

Appendix E

Introduction and Consent form

Welcome to
Emoji Thesis

This is the beginning of this server.

Researcher Lilian 12/08/2020

Hello, and welcome to the Emoji Thesis server. This Discord server was created to carry out user evaluations for an emoji recommendation system that will suggest relevant emojis for your texts. The aim of the user evaluation is to gather feedback for the performance and usefulness of the emoji recommender.

The entire evaluation will take around 10-20 minutes. There are no risks or benefits to you that can be reasonably expected from the user evaluation.

First you will be directed to a survey and asked to fill in the first part about demographics information (age, familiarity with emojis, etc). Then there will be two short conversations with our Conversation Bot, Connie. During one of these conversations, you will be guided through the capabilities of the emoji recommender. Finally, you will be asked to fill in the last part of the survey about your experiences during the conversations.

If you have read the above and have decided to participate in this experiment, please understand your participation is voluntary and you may withdraw at any time without giving any reasons (by leaving the server).

If you have any further questions, you can contact the researcher @Researcher Lilian by sending her a direct message, or emailing her at l.sung@student.utwente.nl
If you are not satisfied with how this study is being conducted, or if you have any concerns, complaints, or general questions about the research or your rights as a participant, please contact the Ethics Committee, Faculty of Electrical Engineering, Mathematics and Computer Science (EEMCS), University of Twente, at +31534896719, or email ethics-comm-ewi@utwente.nl (edited)

Researcher Lilian 12/08/2020

Please click on the corresponding emojis below this message to indicate your consent

✅ I hereby declare that I have been clearly informed about the nature and method of the research and agree to participate.

👍 I hereby give consent for my answers to be quoted and used in presentations or publications. If they are used, they will be completely anonymous. (edited)

✅ 1 👍 1

Appendix F

Connie's Conversations

For the first conversation with no recommendations, Connie always starts with: "Hello 😊🌸 Nice to meet you! How are you doing today?" For the second conversation with recommendations, Connie always starts with: "Hello again :> What's the weather like right now?" After these initial sentences, she follows with either one of the blocks.

Block A

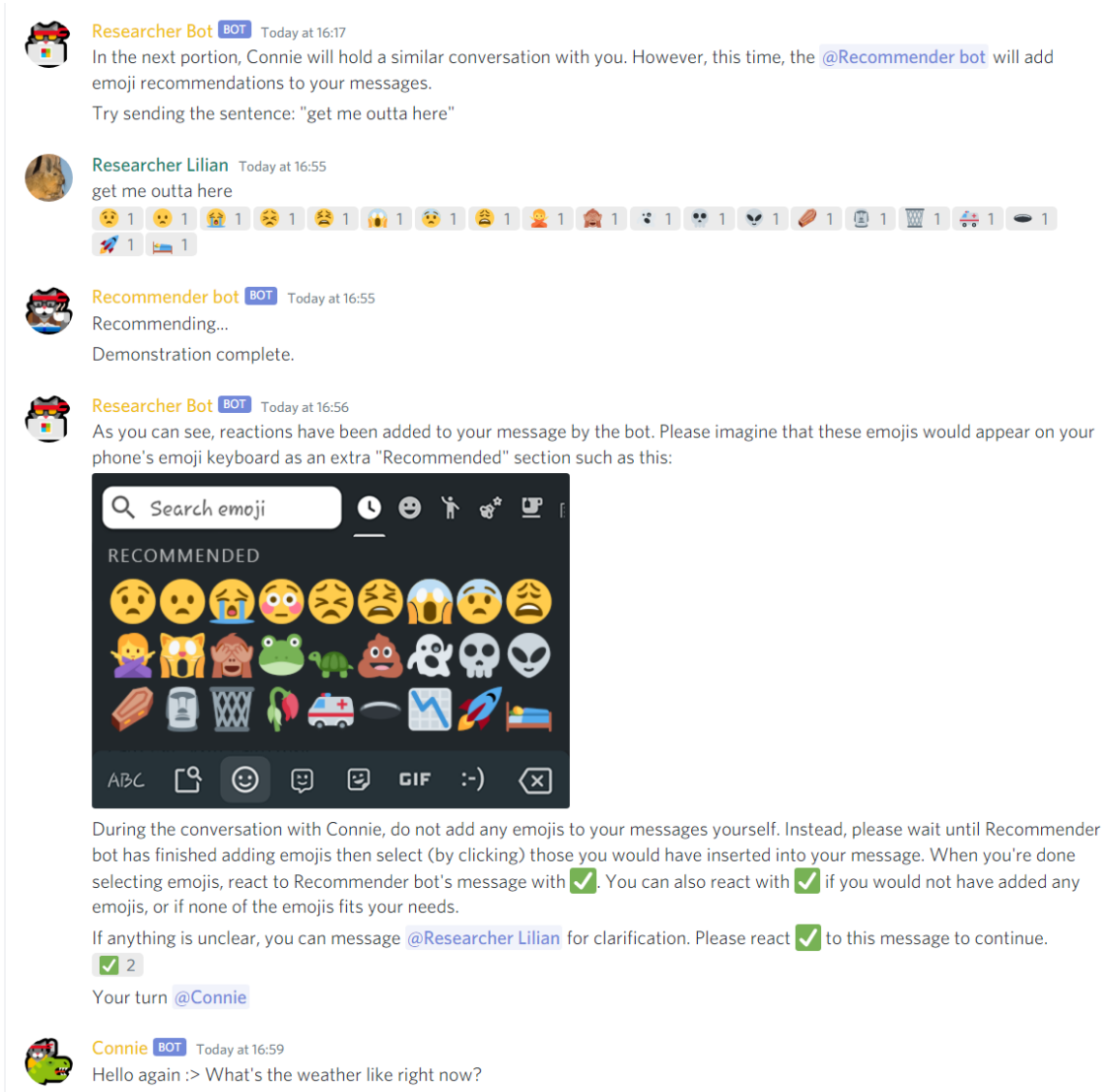
1. What did you do yesterday?
2. What is your favourite animal?
3. Would you rather swim in the sea or climb a mountain?
4. Is there something you're afraid of? Why?
5. Look at this rad croc! *shows video of crocodile jumping into water from a rock, people cheer*

Block B

1. Do you have plans for tomorrow?
2. What is your favourite food?
3. Would you rather go to the zoo or to the cinema?
4. Is there something that makes your sad? Why?
5. Lol he snores *shows sleeping cat snoring into a toy microphone*

Appendix G

Recommendation Demonstration



Appendix H

Consent form corpus survey

UNIVERSITY OF TWENTE.

This short survey is part of a master's thesis on creating an emoji recommendation system that will suggest relevant emojis while you are texting. The aim of this survey is to collect data to help evaluate the quality of the emoji recommendations during development.

It is highly preferred that you take this survey on your phone. The survey will take around 5 to 10 minutes. There are no risks or benefits to you that can be reasonably expected from this questionnaire.

Simple demographics information will be collected. There will be questions asking you to copy recent text messages you have sent that include both text and emoji(s). The text messages will be of your choosing, any sensitive information (such as names) should be omitted in your submissions. Additionally, there will be questions asking you to input emoji(s) you would use given a sample text message.

If you have read this and have decided to participate in this project, please understand your participation is voluntary and you have the right to withdraw from and discontinue the survey at any time without giving any reasons.

I would really appreciate it if you would share this survey further within your social circle, however, it would be appreciated if you did so with people who are above 18 years old due to protocol reasons.

If you have any further questions, you can contact the researcher Lilian Sung at l.sung@student.utwente.nl

If you are not satisfied with how this study is being conducted, or if you have any concerns, complaints, or general questions about the research or your rights as a participant, please contact the Ethics Committee, Faculty of Electrical Engineering, Mathematics and Computer Science (EEMCS), University of Twente, at +31534896719, or email ethics-comm-ewi@utwente.nl

- [ ] I hereby declare that I have been clearly informed about the nature and method of the research and agree to participate.
- [ ] I hereby give consent for my answers to be quoted and used in presentations or publications. If they are used they will be completely anonymous.