Emoji Grammar as Beat Gestures

Gretchen McCulloch
Lingthusiasm
gretchen.mcculloch@gmail.com

Lauren Gawne
La Trobe University
l.gawne@latrobe.edu.au

Abstract

Emoji are popularly characterized as a "language", but languages have grammar. What does an emoji grammar look like? Drawing from sequences of the most common two, three, and four emoji in a large corpus of real emoji use, we find that top emoji sequences have a high level of repetition (∼50%), whereas the equivalent top sequences of words from a large corpus have zero repetition. We argue that emoji are best analogized to "beat" gestures, a well-established type of co-speech gesture characterized by its high level of repetition.

1 Introduction

The use of emoji, small pictures encoded as text (chiefly faces, handshapes, and common objects), is often characterized as "language" or "linguistic" in popular writing (e.g. [Tho16]). Language is comprised of multiple levels. At a simple approximation, we can say that a language has phonemes, which combine to create lexical items (colloquially, words), which combine again to create phrases.

If emoji correspond to any of these levels, it is that of the word. For example, the ❤️ emoji stands for the word "heart" or "love", not the /h/ phoneme or a phrase like "my dog loves pizza." To convey "my dog loves pizza" in emoji, one would need, at minimum, emoji corresponding to "dog," "love," and "pizza," again reinforcing that words are the clear level of comparison.

In: S. Wijeratne, E. Kiciman, H. Saggion, A. Sheth (eds.): Proceedings of the 1st International Workshop on Emoji Understanding and Applications in Social Media (Emoji2018), Stanford, CA, USA, 25-JUN-2018, published at http://ceur-ws.org

But language isn't just a list of words. Language has structures such as subjects and predicates, verbs and objects, nouns and adjectives that modify them. If emoji are truly linguistic, they should also show similar structural properties as words do. In other words, if emoji are language, emoji must have a grammar. This paper searches for a "grammar of emoji" by looking at sequences of emoji from a corpus of over 1 billion emoji uses [McC16] in comparison to the expected sequences based on a large corpus of English words [Dav16] and to an alternate hypothesis from the field of gesture studies, the beat gesture.

Emoji have also been analysed as a strategy for indicating the emotional effect of written speech, which is usually born by prosody and facial expression in spoken language [Miy07] [Wag16]. For example, Face With Tears Of Joy 😂 can indicate a message is intended to be humorous. While we agree that this is one important function of emoji, we do not believe that this accounts for them fully: many common emoji, such as the heart, and all of the objects like food and animals, do not have straightforward effects on prosody. Even those emoji that do have emotional or prosodic functions also have lexical correlates: [Dim15] found that 😂 is used in similar contexts as "lmao," while 😭 is used like "ugh." Analyzing the structure of emoji in terms of words is thus not inconsistent with them having a range of functions, as words do.

1.1 Option 1: Words

When analyzing large corpora of language for structural recurrences, it is common to analyze them in terms of ngrams: recurring sequences of the same n number of words, such as bigrams (2 words), trigrams (3 words), and quadrigrams (4 words). Perhaps the most well-known tool for analyzing ngrams is Google Books Ngrams, where one can find that, for example, the sequence *'telephone operator'* had a sharp rise in the 1910s, and has been decreasing steadily in frequency since the 1940s. Here, as we're looking for a

snapshot of the most frequent ngrams in contemporary English as a whole rather than a historic view of particular ngrams, we computed the top 200 bigrams, trigrams, and quadrigrams from COCA, the Corpus of Contemporary American English [Dav16], which contains around 500 million words from a variety of English texts such as news outlets and websites.

As expected, there are many overlaps between the bigram, trigram, and quadrigram lists. For example, "of the" is a common bigram, while "end of the" is a common trigram and "the end of the" is a common quadrigram. Within these top 200 of each, however, there are zero instances of purely identical sequences, i.e. where the same word is repeated to form the entire ngram. Such identical sequences are possible in English (e.g. "had had" and "very very very"), but they are rare and thus not found in the top 200 lists. Drilling down further into the data, we see that when the same word appears more than once in a trigram or quadrigram, it is at the edges of complex constructions such as "as well as", "the end of the" and "the rest of the".

While COCA is a corpus of formal English, and emoji are often used in informal contexts, pure repetition is not common in any variety of English. For example, COCA has 585,083 instances of "very" of any kind, of which 442 (0.076%) are "very very" or longer [Dav16]. In comparison, the Corpus of Global Web-Based English (GloWbE) has instances of 14,493 "very very" or longer versus 2,345,058 "very" of any kind, a ratio of 0.061% [Dav13].

1.2 Option 2: Beat gestures

In comparison to grammar, repetition is common in the gestural domain. There is no such large public corpus of gestures for numerical comparison, but a particular gesture type, the "beat" gesture, is regularly defined as one that contains a repetitive up-down or side-to-side rhythm [McN92] [McN05]: 40-41; [Ken04]: 103-104, see also [Efr72] 'baton'; [Fri69] 'rhythmic'. The beat gesture is readily observed in both regular conversation, often for emphasizing the rhythm of the accompanying speech (e.g. one might gesture rhythmically in a circle while saying, "You just keep going on and on and on") and oratory (e.g. a confident speaker might thump rhythmically on a podium to emphasize their words, while a nervous speaker might jiggle their hands while talking).

Because a beat refers to the repetitive iteration of a gesture but all gestures must also involve some sort of hand shape in some sort of location, beats readily overlap with other categories of gestures [McN92] [McN05]: 38, 41. For example, a pointing index finger and the thumbs up are each classified in other categories (deixis and emblems, respectively), but either can be produced by moving the hand slightly back and forth for emphasis, i.e. in the style of a beat. We argue that repetition of emoji does not have to distract from its other functions (e.g. representing prosodic information), but can co-occur.

Top 10 Bigrams	Top 10 Trigrams	Top 10 Quadrigrams
1. 😂😂	1. 😂😂😂	1. 😂😂😂😂
2. 😘😘	2. 😘😘😘	2. 😘😘😘😘
3. 😍😍	3. 😍😍😍	3. 😍😍😍😍
4. 😊😊	4. 😊😊😊	4. 😊😊😊😊
5. ❤❤	5. ❤❤❤	5. ❤❤❤❤
6. 😭😭	6. 😭😭😭	6. 💋💋💋💋
7. 😁😁	7. 👍👍👍	7. 👍👍👍👍
8. 👍👍	8. 😁😁😁	8. 😭😭😭😭
9. 😢😢	9. 💋💋💋	9. 😁😁😁😁
10. 😍😘	10. 😢😢😢	10. 😳😳😳😳

Figure 1: Most common emoji combinations

.

2 The SwiftKey Corpus

To decide between these two options, we look at emoji ngrams in a corpus we'll call the SwiftKey Corpus. This corpus was collected from real-life emoji use by users of the SwiftKey smartphone keyboard app on both iOS and Android between January 2016 and April 2016 who had opted into the use of SwiftKey cloud data for more accurate predictions and had their language set to US English, containing over a billion instances of emoji use by English speakers. The most frequent sequences of emoji were programmatically extracted from the data as a whole and analyzed as a list by frequency. So as to preserve user privacy and anonymity, no individual examples of emoji use were examined. The SwiftKey Corpus was initially created for a talk at South by Southwest by Medlock and McCulloch [McC16] and subsequently re-analyzed with additional theoretical framework contributed by Gawne for this paper.

3 Results

The 10 most common sequences of two, three, and four emoji (bigrams, trigrams, and quadrigrams) in the SwiftKey Corpus are listed in Table 1 [McC16]; we analyzed up to the top 200 of each, with and without identical emoji sequences.

Validating the emoji ngram approach, there is considerable similarity between the most common emoji on all three lists, similar to what we saw with the overlap between top bigram, trigram, and quadrigram lists for the word corpus. However, unlike the word corpus, it is immediately evident that there is a very high degree of repetition in the SwiftKey Corpus, which is consistent with anecdotal evidence reported for other

emoji datasets. Looking at the top 200 most common sequences each of two, three, and four emoji in the SwiftKey Corpus, roughly half of each are completely identical repeats (53%, 52%, and 39.5%), with the proportion of non-identical sequences of emoji increasing as one progresses further down each list.

The first non-repeating emoji sequences show up at #10 on the bigram list (😍😘) and #23 on the trigram (😍😍😘) and quadrigram (😍😍😘😘) lists. Within the non-identical sequences, there remains a high degree of internal repetition. Looking only at the top 200 non-identical trigrams and quadrigrams, over half contain a partial repetition, in sequences such as aab, abb, and aba for trigrams (75.5%), and aabb, abab, aaab, abbb for quadrigrams (67.5%). (Non-identical bigrams were not counted, as they must consist of ab.)

Even within entirely heterogeneous sequences (i.e. abc for trigrams and abcd for quadrigrams), all of the top 200 non-identical sequences were thematically similar. Such sequences are heterogeneous at a Unicode character encoding level, but not to a human observer, containing hearts of different colours or shapes (such as 💗💕💖💗), several different monkey faces (🙈🙉🙊), faces of similar emotional valence (such as 😄😂😀, and related clusters of objects (such as 🍸🍺🍻🍷 and 🎂🎁🎈). The only sequence in the top 200 non-identical bigrams, trigrams, and quadrigrams that could possibly depict a scene is 👉 and 👌, but this is more plausibly a depiction of the two-handed gesture that it resembles (both used to represent coitus). No sequences containing simultaneously an attitudinal emoji (such as a face or a heart) and an object emoji (such as food or birthday items) were in the top 200 lists at all.

4 Analysis

Repetition is abundant in emoji sequences, and is rare in speech or written text. It is not impossible to repeat identical words in English, such as "very very very" or "I love love love love it" for emphasis, and salad-salad (in contrast to, say, pasta salad or potato salad) for contrastive focus [Rus04]. Similarly, one could, in principle, write heterogeneous sequences of emoji containing subjects and predicates (e.g. 🐕❤️🍕 to mean "my dog loves pizza" or 🍺😀 to mean "I am happy when I drink beer"). However, neither is a prototypical use, as no such attitude/object pairings are found on the top 200 ngram lists. (In this case, one might ask, what is a happy face emoji indicating an attitude towards? We point to the accompanying words.) Further, many emoji appear on the ngrams lists in both orders, something that is very much atypical for words in English: "birthday happy" is not the same thing as "happy birthday" and yet both "🎂🎈" and "🎈🎂" are common emoji sequences, or occur on larger strings of emoji.

In contrast, the prototypical use of both emoji and beat gestures is one of repetition. In fact, the thumbs up emoji directly appears in the top 10 emoji ngrams lists, just as repeating the thumbs up emblem gesture serves as a beat. Gestures also have the desired flexibility in terms of sequence ordering: like with 🎂🎈 and 🎈🎂 one could equally well point at a person and then a cake to ask if the other person wanted some cake, or to the cake and then the person for the same meaning.

[McC16] further reports that most (85%) of SwiftKey sessions containing any emoji do so alongside words, and of the sessions containing only emoji, the majority are only one to two emoji long, presumably a reply to a previous message. This reinforces another characteristic of the beat gesture, which is its close relationship with words spoken at the same time, although further research of a more fine-grained nature is necessary in order to determine what the details of that relationship.

5 Conclusion

When examining sequences of emoji in use, we have found the most illuminating analysis to be that of emoji as digital gestures, rather than as a grammar with hierarchical structure. In the same way that gestures do not have the same grammatical structure as speech, but act in concert with it, emoji are not taking on the function of grammar, but acting in relation to written text. In particular, repetition of emoji serves an emphatic function that parallels the use of beat gestures in spoken discourse.

While we have focused on beat gestures in this analysis of emoji sequences, we see many other parallels between the use of gestures with spoken language, and emoji with written language. Other gestural categories also show promise for understanding the remainder of the emoji paradigm, which we plan to explore in upcoming work [McC]. In particular, many popular handshape emoji directly represent the category of emblem gestures, and some extended emoji-only sequences parallel the gesture category of pantomime (see Emoji Dick, [Ben10] for one of the most elaborate manifestations of "emoji pantomime").

It is, perhaps, unsurprising that people use emoji in digital communication in ways that parallel use of

co-speech gesture, given that gesture has important functions both for communication [Hey75] [Coe18] and cognition [GM98] [Chu17]. Treating emoji as gesture makes it clear that emoji are unlikely to become a language in their own right. Languages that draw on the same modality as gestures are Signed Languages, and have structural properties that are are more similar to spoken languages than to co-speech gesture, i.e. precisely the structural regularities that we have demonstrated that emoji do not currently have. If emoji do ever emerge as a language proper, we will find it by seeing these same structural regularities emerge in a large corpus study like the one in this paper.

5.0.1 Acknowledgements

Gretchen McCulloch would like to thank the team at SwiftKey, including Jennifer Kutz and Nicky Budd-Thanos, for running data queries and facilitating the talk at South by Southwest. Lauren Gawne would like to thank La Trobe University and the David Myers Fellowship Program. Both authors would like to thank three anonymous reviewers for their helpful comments.

References

[Ben10] F. Benenson. *Emoji Dick*. Lulu, 2010.

[Chu17] S. Kita M. Alibali & M. Chu. How do gestures influence thinking and speaking? the gesture-for-conceptualization hypothesis. *Psychological review*, 124(3):245, 2017.

[Coe18] T. Jenkins M. Coppola & C. Coelho. Effects of gesture restriction on quality of narrative production. *Gesture*, 16(3):416–431, 2018.

[Dav13] M. Davies. *Corpus of Global Web-Based English: 1.9 billion words from speakers in 20 countries (GloWbE)*. https://corpus.byu.edu/glowbe/, 2013.

[Dav16] M. Davies. *The Corpus of Contemporary American English (COCA): 560 million words, 1990-present*. Available online at https://corpus.byu.edu/coca/, 2008-2016.

[Dim15] T. Dimson. Emojineering part 1: Machine learning for emoji trends. *Instagram Engineering blog https://instagram-engineering.com/emojineering-part-1-machine-learning-for-emoji-trendsmachine-learning-for-emoji-trends-7f5f9cb979ad*, May 1, 2015.

[Efr72] D. Efron. *Gesture, race and culture; a tentative study of the spatio-temporal and 'linguistic' aspects of the gestural behavior of eastern Jews and southern Italians in New York City, living under similar as well as different environmental conditions*. Mouton, 1941/1972.

[Fri69] P. Ekman & W.V. Friesen. The repertiore of nonverbal behaviour: categories, origins, usage, and coding. *Semiotica*, 1:49–98, 1969.

[GM98] J. Iverson & S. Goldin-Meadow. Why people gesture when they speak. *Nature*, 396(6708):228, 1998.

[Hey75] J.A. Graham & S. Heywood. The effects of elimination of hand gestures and of verbal codability on speech performance. *European Journal of Social Psychology*, 5(2):189–195, 1975.

[Ken04] A. Kendon. *Gesture: Visible action as utterance*. Cambridge University Press, 2004.

[McC] L. Gawne & G. McCulloch. Emoji are digital gesture (in prep).

[McC16] B. Medlock & G. McCulloch. The linguistic secrets found in billions of emoji. *Presented at SXSW, March 11-20, Austin, Texas*, 2016.

[McN92] D. McNeill. *Hand and mind: What gestures reveal about thought*. The University of Chicago Press, 1992.

[McN05] D. McNeill. *Gesture and thought*. University of Chicago Press, 2005.

[Miy07] K. Miyake. How young Japanese express their emotions visually in mobile phone messages: A sociolinguistic analysis. *Japanese Studies*, 27(1):53–72, 2007.

[Rus04] J. Ghomeshi R. Jackendoff N. Rosen & K. Russell. Contrastive focus reduplication in English (the salad-salad paper). *Natural Language & Linguistic Theory*, 22(2):307–357, 2004.

[Tho16] C. Thompson. The emoji is the birth of a new type of language (no joke). *Wired*, 24.04.16, 2016.

[Wag16] M. Wagner. How to be kind with prosody. *Speech Prosody 2016*, pages 1250–1253, 2016.