Brandeis University
Master's Program in Computational Linguistics

MojiSem: Varying Linguistic Purposes of Emoji in (Twitter) Context

Noa Na'aman, Hannah Provenza, Orion Montoya 💡 Brandeis University, Waltham, MA

Emoji serve different linguistic functions on different occasions

- Pipelines that ignore emoji, or bucket them as punctuation, ignore key aspects of computer-mediated communication
- Emoji analysis that looks only at frequency or distribution ignores the distinctive communicative potentials of non-textual characters

Identifying where an emoji is replacing textual content allows NLP tools the possibility of parsing emoji as any other word or phrase. Recognizing the import of non-content emoji can be a a significant part of understanding a message; in this, humans have a distinct advantage over computers.

Recent work (Miller et al., 2016) has explored the cross-platform ambiguity of emoji renderings; (Eisner et al., 2016) created word embeddings that performed competitively on emoji analogy tasks; (Ljubešic and Fišer, 2016) mapped global emoji distributions by frequency; (Barbieri et al., 2017) used LSTMs to predict them in context. (Solomon 2017) recently looked at implicit syntax in directional emoji.

We feel that a **lexical semantics of emoji characters** is implied in these studies without being directly addressed. Words are not deployed randomly, and neither are emoji. Even when they replace a word, emoji are used for different purposes than words. We believe that work on emoji would be better informed if it **made explicit accommodation of the varying communicative functions that emoji can serve in expressive text**. The current project annotated emoji in tweets by linguistic and discursive function. A model trained on this corpus predicted the communicative purpose of emoji characters in novel contexts.

We find that it is possible to **train a classifier to tell the difference between emoji used as linguistic content words and those used as paralinguistic or affective multimodal markers**, even with a small amount of training data; but that accurate sub-classification of these multimodal emoji into specific classes like attitude, topic, or gesture will require more data and more feature engineering.

Collect tweets with `tweepy`; annotate tweets with linguistics students

We pulled tweets from the public Twitter streaming API using the `tweepy` Python package. Tweets were automatically filtered to include only tweets with characters from the Emoji Unicode ranges and only tweets labeled as being in English. We excluded tweets with embedded images or links. Redundant/duplicate tweets were filtered by comparing tweet texts after removal of hashtags and @mentions; this left only a small number of mutant-clone duplicates. After that, tweets were hand-selected to get a wide variety of emojis and context in a small sample size — therefore, our corpus does not reflect a true distribution of emoji uses or context types.

The analytical **tasks of annotators** were:

- Identifying each emoji in the tweet
- Deciding whether multiple contiguous emoji should be considered separately or as a group
- Choosing the best tag for the emoji (or sequence)
- Providing a translation or interpretation for each tagged span.

Inter-Annotator Agreement

Dataset	# taters	span rem	total	mm	content
Set 1	4	78	0.2071	0.4251	0.1311
Set 2	4	49	0.8743	0.7158	0.8531
Set 3A	2	11	0.9096	0.4616	0.792
Set 3B	2	6	0.7436	0.3905	1.0
Set 4A	2	3	0.8789	0.4838	0.7435
Set 4B	2	1	0.3954	0.5078	1.0
Total/mean	4	150	0.6681	0.4974	0.7533

Table 1: Fleiss's κ scores and other annotation/agreement variables

We calculated agreement with Fleiss's κ, which requires that annotators have annotated the same tokens. Rather than impute disagreement in the case of an incompletely-annotated batch, we removed from our IAA-calculation counts any spans that were not marked by all annotators. There are many of these in the first dataset, and progressively fewer as the annotators gained facility. A total of 150 spans were excluded from Fleiss's κ calculations for this reason.

Content words are easy to label; our multimodal subtypes are too subjective

Content words. Part-of-speech identification is a skill familiar to most of our annotators, so we were not surprised to see excellent levels of agreement among emoji tagged for part of speech. These content words, however, were a very small proportion of the data (51 out of 775 spans) which may be problematically small.

Multimodal. Agreement on multimodal sub-labels was much lower, and did not improve as annotation progressed. Multimodal emoji may be inherently ambiguous, and we need a labeling system that can account for this. A smiley face might be interpreted as a `gesture` (a smile), an `attitude` (joy), or a `topic` (for example, if the tweet is about what a good day the author is having) — and any of these would be a valid interpretation of a single tweet. A clearer typology of multimodal emojis, and, if possible, a more deterministic procedure for labeling emoji with these subtypes, may be one approach.

Worst overall cross-label agreement scores were for Set 1, but all following datasets improved on that baseline after the annotation guidelines were refined.

Objective: Distinguish content tokens from multimodal uses

Key intuition: **content emoji are pronounceable**, while **non-content emoji must be described or performed.**

We attribute this to different motivations in using emoji. Annotators read tweets aloud to themselves in order to demonstrate the category of each use.

Pronounceable:

Emoji as function words

"I 🍩 like u"

Subtypes: `prep, aux, conj, dt, punc`

Emoji as content words:

"The 🔑 to success is 🍕"

Subtypes: `noun, verb, adj, adv`

Performative or topical:

Emoji as affect, topic, or gesture

`attitude`: "Let my work disrespect me one more time... 🙃"

`topic`: "Mean girls 🎬"

`gesture`: "Omg why is my mom screaming so early 🙄"

Gold-Standard Counts

Label	count
Multi-modal (mm)	total 686
attitude	407
topic	184
gesture	93
other	2
Content (cont)	total 51
noun	40
adj	6
verb	4
adv	1
Functional (func)	total 38
punct	34
aux	2
dt	1
other	1
emoji spans	total 775
words	6174
punctuation	668

Application: CRF-tagging the linguistic function of emoji tokens

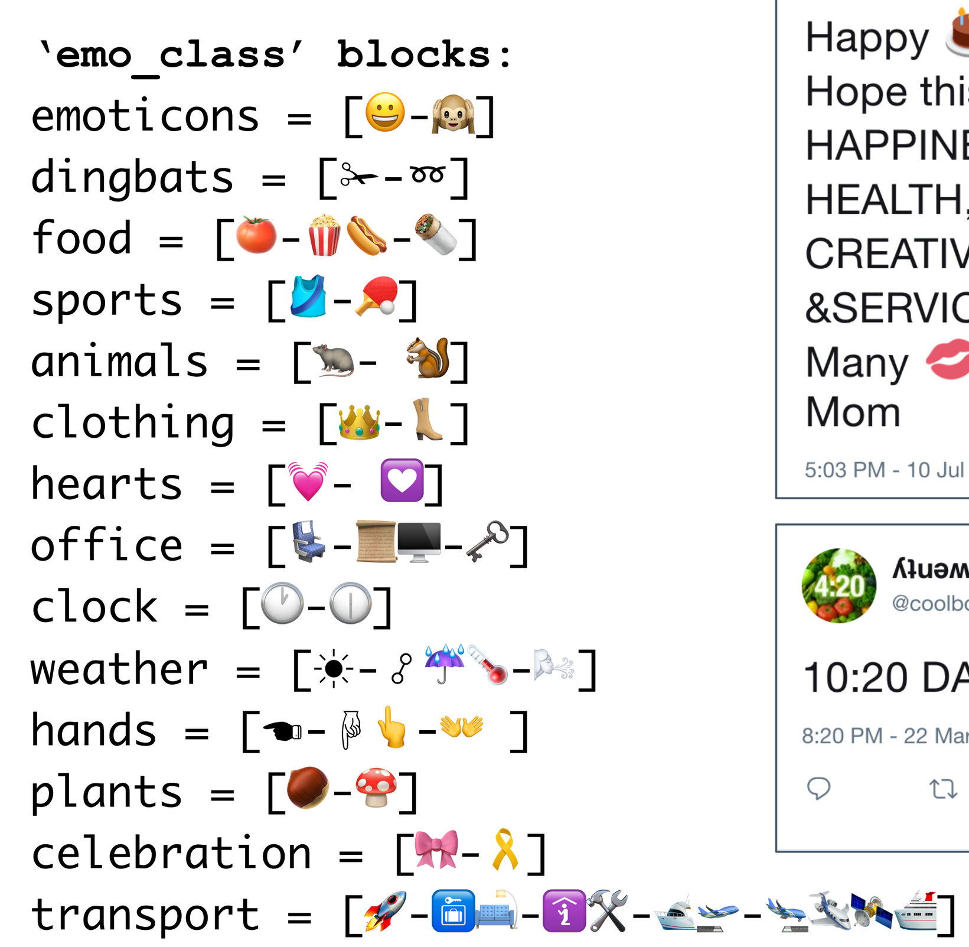
Using our gold-standard dataset, we trained a CRF tagger to assign linguistic-function labels to emoji characters. Due to the low agreement on the annotated sub-types of multimodal (`mm`) labels, and to the small number of `cont` and `func` labels assigned, we narrowed the focus of our classification task: simply **categorizing tokens correctly as either `mm` or `cont/func`**. After one iteration, we saw that the low number of `func` tokens was preventing us from finding any `func` emoji, so we combined the `cont` and `func` tokens into a single label of cont. Therefore our sequence tagger needed simply to decide whether a token was serving as a substitute for a textual word, or was a multimodal marker.

Feature engineering

Context helps; Unicode blocks can be a proxy for semantics; POS tagging is a nice hint

Features extracted for training

- The token itself
- 'emo?' — whether the token contains emoji characters (emo), or is purely word characters (txt).
- 'POS', a part-of-speech tag assigned by nltk.pos_tag
- 'position', a set of three positional features:
  - an integer 0–9 indicating a token's position in tenths of the way through the tweet;
  - a three-class BEGIN/MID/END to indicate tokens at the beginning or end of a tweet (different from the 0–9 feature in that multiple tokens may get 0 or 9, but only one token will get BEGIN or END);
  - the number of characters in the token.
- The 'contexty' feature is another set of three features, this time related to **context**:
  - A boolean TRUE if the previous token was a determiner, FALSE otherwise;
  - The previous and the next tokens' POS tags, paired with the current 'emo?' value
- The token's thematic **Unicode blocks**. The Unicode Consortium adds and lists emoji in semantically-related groups that tend to be contiguous within a range of codepoints. Blocks of characters with shared semantic attributes are matchable with a simple range regex. These provide a very inexpensive proxy to semantics, and the resulting 'emo_class' feature yielded a marked improvement in both precision and recall on content words (although the small number of cases in the test data make it hard to be sure of their true contribution).



Confounds for entertainment

Detail of three feature-extracted tweets

`content:`

```
word         It     txt   PRP     2   MID  2   False  :|txt       VBD|txt txt
word         cost   txt   VBD     3   MID  4   False  PRP|txt     -NONE-|txt     txt
word         $0     txt   -NONE-  4   MID  2   False  VBD|txt     .|txt   txt
punc         .      txt   .       5   MID  1   False  -NONE-|txt  CD|txt  txt
word         00     txt   CD      6   MID  2   False  .|txt       TO|txt  txt
word         to     txt   TO      7   MID  2   False  CD|txt      VB|txt  txt
word         be     txt   VB      8   MID  2   False  TO|txt      :|txt   txt
content_adj  💯     emo   :       9   END  2   False  VB|emo      END|emo other
```

`function:`

```
word     @RangersFC  txt  NN      4  MID  10  False  NN|txt      VBP|txt txt
word     have        txt  VBP     4  MID  4   False  NN|txt      VBN|txt txt
word     scored      txt  VBN     5  MID  6   False  VBP|txt     -NONE-|txt     txt
func_dt  ⚽⚽⚽⚽⚽⚽⚽⚽⚽  emo  -NONE-  5  MID  18  False  VBN|emo     NNS|emo other
word     goals       txt  NNS     6  MID  5   False  -NONE-|txt  IN|txt  txt
word     against     txt  IN      6  MID  7   False  NNS|txt     NNP|txt txt
word     Alloa       txt  NNP     6  MID  5   False  IN|txt      DT|txt  txt
word     this        txt  DT      7  MID  4   False  NNP|txt     NN|txt  txt
word     season      txt  NN      7  MID  6   True   DT|txt      IN|txt  txt
```

`a multimodal` menagerie:

```
word         thank       txt  NN     2  MID  5   False  :|txt     PRP|txt   txt
word         you         txt  PRP    3  MID  3   False  NN|txt    IN|txt    txt
word         for         txt  IN     3  MID  3   False  PRP|txt   PRP$|txt  txt
word         my          txt  PRP$   3  MID  2   False  IN|txt    NN|txt    txt
word         present     txt  NN     4  MID  7   False  PRP$|txt  :|txt     txt
mm_attitude  😺          emo  :      4  MID  2   False  NN|emo    RB|emo    emoticon
word         definitely  txt  RB     4  MID  10  False  :|txt     VBG|txt   txt
word         chilling    txt  VBG    5  MID  8   False  RB|txt    IN|txt    txt
word         in          txt  IN     5  MID  2   False  VBG|txt   DT|txt    txt
word         the         txt  DT     5  MID  3   False  IN|txt    NN|txt    txt
word         bath        txt  NN     6  MID  4   True   DT|txt    JJ|txt    txt
word         later       txt  JJ     6  MID  5   False  NN|txt    :|txt     txt
punc         ;           txt  :      6  MID  1   False  JJ|txt    PRP|txt   txt
word         they        txt  PRP    7  MID  4   False  :|txt     VBP|txt   txt
word         smell       txt  VBP    7  MID  5   False  PRP|txt   RB|txt    txt
word         insanely    txt  RB     7  MID  8   False  VBP|txt   JJ|txt    txt
word         good        txt  JJ     8  MID  4   False  RB|txt    NN|txt    txt
mm_topic     🛁          emo  NN     8  MID  2   False  JJ|emo    :|emo     transport
mm_attitude  💗          emo  :      8  MID  2   False  NN|emo    :|emo     heart
mm_gesture   👌          emo  :      9  MID  2   False  :|emo     :|emo     hand
word         🏼          emo  :      9  END  2   False  :|emo     END|emo   other
END          END
```

Metrics on CRF tagging

(at least recognizing `words` is easy)

feature	F1 `word`	F1 `mm`	P `cont`	R `cont`	F1 `cont`	Macro-avg F1
character	0.9721	0.7481	0.3571	0.3333	0.3448	0.8441
prev +emo?	0.9914	0.8649	0.4286	0.4000	0.4000	0.8783
prev +POS	0.9914	0.8784	0.5000	0.4667	0.4828	0.8921
prev +position	0.9914	0.8844	0.4667	0.4667	0.4667	0.9028
prev +contexty	0.9914	0.8831	0.6250	0.3333	0.4348	0.8848
prev +emo_class (best)	0.9914	**0.8933**	**0.7273**	0.5333	0.6154	**0.9168**
best − character	0.9906	0.8514	0.6429	**0.6000**	**0.6207**	0.9090
best − contexty	0.9922	0.8750	0.4706	0.5333	0.5000	0.8945
emo?+POS+emo_class	0.9914	0.8421	0.6000	0.4000	0.4800	0.8855

An encouraging start

89 examples of content and functional uses of emoji are not enough to reliably model the behavior of these categories. **More annotation** may yield much richer models of the variety of purposes of emoji, and will help get a better handle on the range of emoji polysemy.

Clustering of contexts based on observed features may induce **more empirically valid subtypes** than the ones defined by our specification.

Emoji's novel communicative functions must be attended to

- Some emoji senses may fall into ontological or onomasiological groupings of semantics
- Others clearly fall into the realm of pragmatics and its typologies
- Confusing them is likely a hindrance to insight
- "Emoji-sense disambiguation" could help

Anglophone Twitter users use emoji in their tweets for a wide range of purposes. Some emoji are clearly polysemous; few if any may be inherently monosemous.

Every emoji linguist notes the fascinating range of pragmatic and multimodal effects that emoji can have in electronic communication. If these effects are to be given lexicographical treatment and categorization, they must also be organized into functional and pragmatic categories that are not part of the typical range of classes used to talk about either printed or spoken words.

Emoji-sense disambiguation (ESD). ESD in the model of traditional WSD would seem to require an empirical inventory of emoji senses. Even our small sample has shown a number of characters that are used both as content words and as topical or gestural cues.

There can be little question that individuals use emoji differently, and this will certainly confound the study of emoji semantics in the immediate term. The study of community dialects will be essential to emoji semantics, and there is certain also to be strong variation on the level of idiolect. The categorizations may need refinement, but the phenomenon is undeniably worthy of further study.

References

Francesco Barbieri, Miguel Ballesteros, and Horacio Saggion. 2017. Are emojis predictable? *EACL 2017*, page 105.

Ben Eisner, Tim Rocktäschel, Isabelle Augenstein, Matko Bošnjak, and Sebastian Riedel. 2016. emoji2vec: Learning emoji representations from their description. In *Conference on Empirical Methods in Natural Language Processing*, page 48.

Nikola Ljubešic and Darja Fišer. 2016. A global analysis of emoji usage. *ACL 2016*, page 82.

Hannah Miller, Jacob Thebault-Spieker, Shuo Chang, Isaac Johnson, Loren Terveen, and Brent Hecht. 2016. "Blissfully happy" or "ready to fight": Varying interpretations of emoji. In *Proceedings of the Tenth International Conference on Web and Social Media, ICWSM 2016, Cologne, Germany, May 17– 20, 2016*. Association for the Advancement of Artificial Intelligence, May.

Tyler Schnoebelen. 2012. Do you smile with your nose? Stylistic variation in Twitter emoticons. In *University of Pennsylvania Working Papers in Linguistics*, volume 18, pages 117–125. University of Pennsylvania.

Jane Solomon, 2017. *Gun Emoji Pairings*, https://www.lexicalitems.com/blog/gun-emoji-pairings

Unicode Consortium, 2017. *Full Emoji List, v5.0*. http://unicode.org/emoji/charts/full-emoji-list.html

Thanks: James Pustejovsky, Keigh Rim, Marie Meteer; our annotators Anna Astori, Jake Freyer, Jose Molina, Annie Thorburn; Cherilyn Sarkisian.