The lexicon of emoji? Conventionality modulates processing of emoji

Benjamin Weissman[1], Jan Engelen[2], Elise Baas[2], & Neil Cohn[2]

[1] Department of Cognitive Science, Rensselaer Polytechnic Institute

[2] Department of Communication and Cognition, Tilburg University

Abstract

Emoji have been ubiquitous in communication for over a decade, yet how they derive meaning remains underexplored. Here we examine an aspect fundamental to linguistic meaning-making: the degree to which emoji have conventional lexicalized meanings and whether that conventionalization affects processing in real time. Experiment 1 establishes a range of meaning agreement levels across emoji within a population; Experiment 2 measures accuracy and response times to word-emoji pairings in a match/mismatch task. In this experiment, we found that accuracy and response time both correlated significantly with the level of population-wide meaning agreement from Experiment 1, suggesting that lexical access of single emoji may be comparable to that of words, even out of context. This is consistent with theories of a multimodal lexicon that stores links between meaning, structure, and modality in long term memory. Altogether, these findings suggest that emoji can allow a range of entrenched, lexicalized representations.

Keywords: emoji; visual language; lexicon; language processing

1. Introduction

As emoji continue to increase in prevalence, so has scholarly attention to the way they are used to communicate. One element of this is whether emoji have a consistent "vocabulary," either on their own or when combined in context with written language (e.g., Holtgraves & Robinson, 2020; Miller et al., 2017; Miller et al., 2016; Weissman, 2019a). Here we aim to investigate emoji lexicalization by measuring meaning agreement and response times to emoji presented out of context.

1.1 Lexicalization

Jackendoff, in an early description of the parallel architecture (2002), emphasizes the detachment of "lexical items" from "words." A lexical item, under this view, is a "unit stored in long term memory" (2002: 154), and could be one of a variety of forms, including a word, a full phrase (e.g., an idiom), or a productive morpheme. These items include not only semantic information but also correspondences to structural and phonological (or orthographic) information as well. This approach of the parallel architecture has subsequently been extended to lexical items in other modalities (Cohn, 2016; Cohn & Schilperoord, 2022); consistent with the original flexibility of the "lexical item," these same long-term memory representations can be established for items like gestural emblems (Ladewig, 2020; McNeill, 1992), traffic signs (Forceville, 2019), logos (Foroudi, Melewar, & Gupta, 2014), and emoji.

In this view, lexicalization can be understood as the process by which content from any modality (e.g., speech sounds, graphics) establishes a correspondence to a concept, and this association becomes stored in memory. In other words, "where a particular meaning component is found to be in regular association with a particular morpheme." (Talmy, 1985: 59). According to the parallel architecture, this information encoded in memory necessarily includes elements of modality, grammar, and meaning coindexed across structures (Jackendoff, 2002; Jackendoff & Audring, 2020). The lexicalization process involves the strengthening of these encodings, which is often called entrenchment (Langacker, 1987; Schmid, 2007, 2017): the more entrenched something is, the easier and more automatic it will be to access or generate or process (Blumenthal-Dramé, 2012; Divjak & Caldwell-Harris, 2015; Schmid, 2017).

Given the lexicon and lexicalization process described above, highly lexicalized items will be accessed and recognized more quickly (e.g., Blumenthal-Dramé, 2017), and thus can yield faster response times in tasks such as lexical decisions and match/mismatch assessments. Entrenchment depends on a variety of internal and external factors, including community-wide usage and frequency (Schmid, 2007; 2017); even though this process occurs on an individual-by-individual level, we may expect entrenchment for specific items (excepting highly specialized items e.g., jargon, slang, regionalisms) to be relatively similar across a similar population. Indeed, Schmid (2015) describes a model of language built around entrenchment and conventionalization, elaborating on the myriad ways the two processes can feed each other. The present study will investigate this link as it relates to emoji, exploring the relationship between population-wide meaning agreement and lexical access.

1.2 Emoji meanings

Since the standardized emoji set is still relatively recent and remains in flux, the lexicalization process is likely still underway and the picture we paint here is not a static one. We might expect some emoji to be relatively agreed upon, such as those with clear pictorial representations of a concrete concept (most foods, animals, objects, etc.) without much potential for ambiguity nor room for interpretation. Częstochowska et al. (2022) analyzed ambiguity rates for different categories of emoji based on participant descriptions of individual emoji presented without surrounding context. They found “food & drink”, “clothes & accessories”, “nature”, and “hearts” to be the categories with the least variation. The path to lexicalization for these items may be an abbreviated one, since the emoji in these categories are often iconic of their real-world counterparts.

An analysis of emoji usage rates (emojitracker, 2021), however, indicates that these emoji are relatively rare among all emoji used – the most commonly used emoji tend to be facial expressions and other expressive emoji. These emoji, in contrast with the previous set, have more room for ambiguity and differing interpretations. This is reflected in Emojipedia’s definition pages (Emojipedia, 2022), which often include multiple meanings for these face emoji, as well as in a number of empirical studies. The analysis by Częstochowska et al. found face emoji to yield significantly more variance in participant-provided descriptions than the categories mentioned earlier. Other work targeting emoji meanings (Sick et al., 2020) has indicated a range of agreement across these emoji whereby some seem to be significantly more agreed upon than others. Previous research (Miller et al., 2016) has indeed found that people demonstrate lower agreement rates when naming face emoji and can even differ in classifying their emotional valence (positive/neutral/negative). In addition, these emoji prove more useful in representing abstract semantic concepts than concrete ones (Wicke & Bolognesi, 2020), consistent with the idea that they are less tied to particular unambiguous meanings. In line with recent findings involving word-based polysemy (Lai & Chung, 2018), we may expect a correlation between the opaqueness of the emoji-meaning link and degree of lexicalization, whereby these more opaque, more abstract, and less rigidly defined emoji are not as strongly lexicalized as their less common, more concrete, more transparent counterparts.

Much of the research on emoji lexicalization has looked at emoji within sentence contexts. Emoji are readily integrated into reading of sentences when they substitute for matching words (*John loves eating* 🍕) compared to when they substitute for mismatching words (*John* 🍕 *eating pizza*) (Cohn et al., 2018). In addition, when substituting high naming agreement emoji (e.g., foods, animals, objects) for words in a sentence (e.g., *the chef carved the turkey with a* 🔪), unexpected and incongruous emoji were found to lead to the same neural responses as unexpected and incongruous words (Late Frontal Positivity and N400, respectively) (Weissman, 2019b). In another study testing face and expressive emoji (e.g., *our project eventually succeeded, and I felt very* 😊), incongruous emoji were found to elicit a slightly different neural response than incongruous words, taken as evidence of a processing cost in accessing these emoji in sentence contexts (Tang et al., 2020). This difference in findings may be consistent with a fundamental semantic difference between expressive and non-expressive emoji (Kaiser & Grosz, 2021), though there may be questions about the naturalness of these stimuli. These findings demonstrate that such emoji have at least prototypical correspondences to meanings that are violable.

Other work has investigated emoji-only productions in which an entire conversational turn is replaced by emoji. Holtgraves & Robinson (2020) designed an experiment in which a conversation included an indirect response to a prompt in the form of text, text + a face emoji, or only a face emoji. To the question "What did you think of my presentation?" the response could be "It's hard to give a good presentation," "It's hard to give a good presentation 😬," or "😬." They measured response times to a yes/no categorization task of a text-based paraphrase of that indirect response ("I didn't like your presentation"). Paraphrases of messages with an emoji (either on its own or in conjunction with text) were responded to quicker and more accurately than the same paraphrases of messages without any emoji included. The emoji-only response condition here does not directly replace words or phrases, but rather expresses similar affective information; in this experiment, the uptake of such a response was faster and more accurate compared to text-only.

More research has investigated the meaning contributions of emoji when added to sentences rather than replacing words in sentences. For example, non-face emoji added to the ends of sentences were fixated for shorter periods and re-fixated less often when the emoji were congruent with a target word in the sentence (Barach et al., 2021). Congruent emoji were also skipped more frequently than incongruent emoji, suggesting that semantic information can be gleaned from a parafoveal preview of emoji like it can with words (ibid.). Synonymous and inference-consistent emoji added to sentences also yielded higher perceived coherence and better emoji recall than incongruous emoji, demonstrating that these symbols are integrated readily and holistically into participants' utterance interpretations in memory (Christofalos, Feldman, & Sheridan, 2022). In addition, adding an irony-marking wink emoji to the ends of sentences elicited the same neural response pattern as "traditional," word-based irony (Weissman & Tanner, 2018). Congruence can also occur between the emotional valence of emoji and sentences: positive/negative emoji that matched the valence of sentences led to faster processing than when the word and emoji valence mismatched (Boutet at al., 2021). Neutral valence sentences, however, sponsored no processing effect of emoji valence (Robus et al., 2020), although sentences with sentence-final emoji yielded slower reading times and longer emoji fixations than those with sentence-initial emoji, a finding attributed to integration-related wrap-up effects. These studies all explore how emoji contribute meaning to utterances and how that meaning is processed, though they do not directly address lexicalization; our current work aims to build on this robust and growing body of literature by exploring emoji meaning access and consistency outside of sentence contexts.

1.3 Polysemy and homonymy

Insights from the linguistics literature regarding polysemy and homonymy are relevant to the emoji discussion here. While both fit under the broader umbrella of semantic ambiguity, homonyms have two or more unrelated senses (e.g., river *bank* vs. financial institution *bank*) while polysemes have two or more related senses (e.g., a *line* of text, a *line* drawing, wait in a *line*, a hockey *line*). In these cases, multiple entrenched meanings for the same form become disambiguated through context.

Though the dynamics are not quite as clear with emoji as they are with words, examples of both homonymy and polysemy exist in the emoji data collected by Częstochowska et al. (2022). The most frequent descriptions provided for 😈 were "evil" and "devil," and the most frequent for ☹️ were "sad" and "frown." These examples both seem to constitute classic examples of polysemy, with the latter acting as an embodiment or a manifestation of the former. On the other hand, the most frequent descriptions provided for 😋 were "tasty" and "silly," and 😗 was described as "kissing," "surprise," and "whistle." While these aren't quite as "accidental" (Weinreich, 1964) as some examples of word-based lexical ambiguity, since these concepts share facial features like the tongue sticking out or pursed lips, these meanings are not related enough to be considered polysemy and these would be better classified as unrelated, homonymous emoji.

How polysemy and homonymy affect processing has proven to be a convoluted question (see Eddington & Tokowicz, 2015 for review) with ample inconsistency. Many studies find a processing advantage for polysemous words (as measured by, for example, faster response times to polysemous vs unambiguous words in a lexical decision task) and a processing disadvantage for homonymous words. In addition, homonyms exhibit dominance effects – a dominant sense of an ambiguous word will be activated with more ease than a subordinate sense, and the magnitude of this effect is correlated with how dominant that primary meaning is (e.g., Klepousniotou et al., 2012; Meade & Coch, 2017); whether this effect also extends to polysemes has been a source of controversy (e.g., Brocher et al., 2018; Foraker & Murphy, 2012; Klepousniotou et al., 2012). While the bevy of processing results is not directly comparable to the present study of emoji meanings due to task differences, some relevant insights can nonetheless be gleaned.

Just as with ambiguous words, context is key for disambiguating homonymous emoji. Attempts at machine disambiguation of emoji in context have been successful (Shardlow et al., 2022; Wijeratne et al., 2016) and human emoji judgments are context-dependent as well (Miller et al., 2017; Weissman, 2019a). However, even though people can use context to figure out which sense of a word is meant in a sentence, the dominance effect that manifests as a processing advantage for one meaning over another demonstrates that context does not tell the entire story here. For this reason, exploring emoji out of context is valuable for tracking their potential lexicalization. If a dominance effect can indeed be found in the processing of out-of-context emoji, that suggests a similar lexicalization process as that of words whereby meaning(s) of these items gets entrenched in the lexicon, even without context.

1.4 Current study

Here, we aim to explore the extent to which emoji currently constitute lexicalized units. While emoji meaning can certainly change based on context, studying isolated emoji on their own will provide the clearest picture of whether emoji themselves can be lexicalized items. We hypothesize that emoji with higher agreement for their meaning will lead to faster response times than emoji with relatively low agreement. Such a relationship would suggest emoji are susceptible to lexical entrenchment, comparable to the lexicalization of words. If no such correlation is present between response time and agreement, it may be that the extent to which these meanings are shared across a broader population may not affect or reflect individuals' internal cognitive processes. Other possibilities are that emoji have no entrenched meanings and derive their meanings based only on their contextual relationship with language, or that the emoji differ in how sensitive they are to surrounding context.

2. Experiment 1: Emoji meaning agreement

We first turn to our analysis of offline emoji lexicalization: do people agree on the meanings of emoji? Intuitively and evidently (Częstochowska et al., 2022), some emoji should have high meaning agreement (e.g., many of the animals, foods, and objects). Others, however, are not as straightforward – some are polysemous with multiple widely-recognized meanings, like the notoriously euphemistic peach and eggplant, and others may not have any clearly-accepted and widely-agreed upon meaning. Surrounding context can help disambiguate the meanings of emoji (Miller et al., 2017; Weissman, 2019a), but without any surrounding context there is more room for varying interpretations (Miller et al., 2016). This experiment thus seeks to establish baseline agreement ratings for a set of emoji to then be compared to responses and reaction times during processing in Experiment 2.

2.1. Methods

2.1.1. Participants

There were 120 participants in the experiment (average age = 25.8 (SD = 8.90), 37 male, 81 female, 2 other). Participants confirmed their informed written consent according to approval from (*information removed for anonymization purposes*). Participants completed a basic demographic questionnaire as well as an Emoji Language Fluency questionnaire to assess participants' self-reported familiarity and expertise with emoji. Emoji Language Fluency, as assessed through this self-report questionnaire, did not significantly co-vary with participant agreement score in Experiment 1 nor participant accuracy/reaction time in Experiment 2 and thus will not receive further discussion here; the survey itself appears in the supplemental materials.

2.1.2. Stimuli

Thirty Apple iOS emoji were selected for use in this agreement study. Many of these (n = 25) had more than one meaning listed on Emojipedia and a small subset (n = 5) had only one meaning listed. This heterogeneity in the stimuli was desired to identify a potential entrenchment range. In search of emoji with potentially ambiguous meanings, we avoided seemingly straightforward ones (like many of the objects/animals/foods) and instead included more facial expressions and potentially euphemistic emoji. The full list of emoji used can be seen in Figure 1.

2.1.3. Procedure

In an online experiment (hosted on Qualtrics), participants were shown one emoji at a time in a randomized order and were instructed to type the meaning of that emoji into a text box. All participants were shown all 30 emoji. Text box input was limited to 30 characters to ensure participants' answers were short. In addition to this main task, participants filled out a short demographic survey and the aforementioned Emoji Language Fluency questionnaire.

2.1.4. Data Analysis

Responses were categorized into meaning bins by three coders (intercoder reliability: κ = .902) and those bins were sorted based on frequency. Coders were instructed to categorize responses with the same meaning into the same bin (e.g., “smile” and “smiling” into bin 1, “happy” into bin 2, “eager” into bin 3). For each emoji, if a response was given by five or more participants, it was treated as a category for that emoji; if a certain response was given by fewer than five participants, it was treated as “other.” If a participant typed in two responses, both were logged. This categorization was used to determine the most commonly-provided answer for each emoji, which will later be used in Experiment 2; it also provides the proportion of participants who gave that most commonly-provided answer as well as how many other answers were provided. This gives a well-rounded view of population-wide agreement for each emoji.

2.2. Results

Figure 1 shows the proportion of respondents that provided a given answer for each emoji. Over half of the emoji had over 50% agreement for the most agreed-upon answer. As evidenced by the yellow portion of each row representing the most commonly-provided answer for each emoji, there was a wide range of agreeability across the different emoji ranging from 25% to 98%. Secondary choices ranged from 1% to 42%.

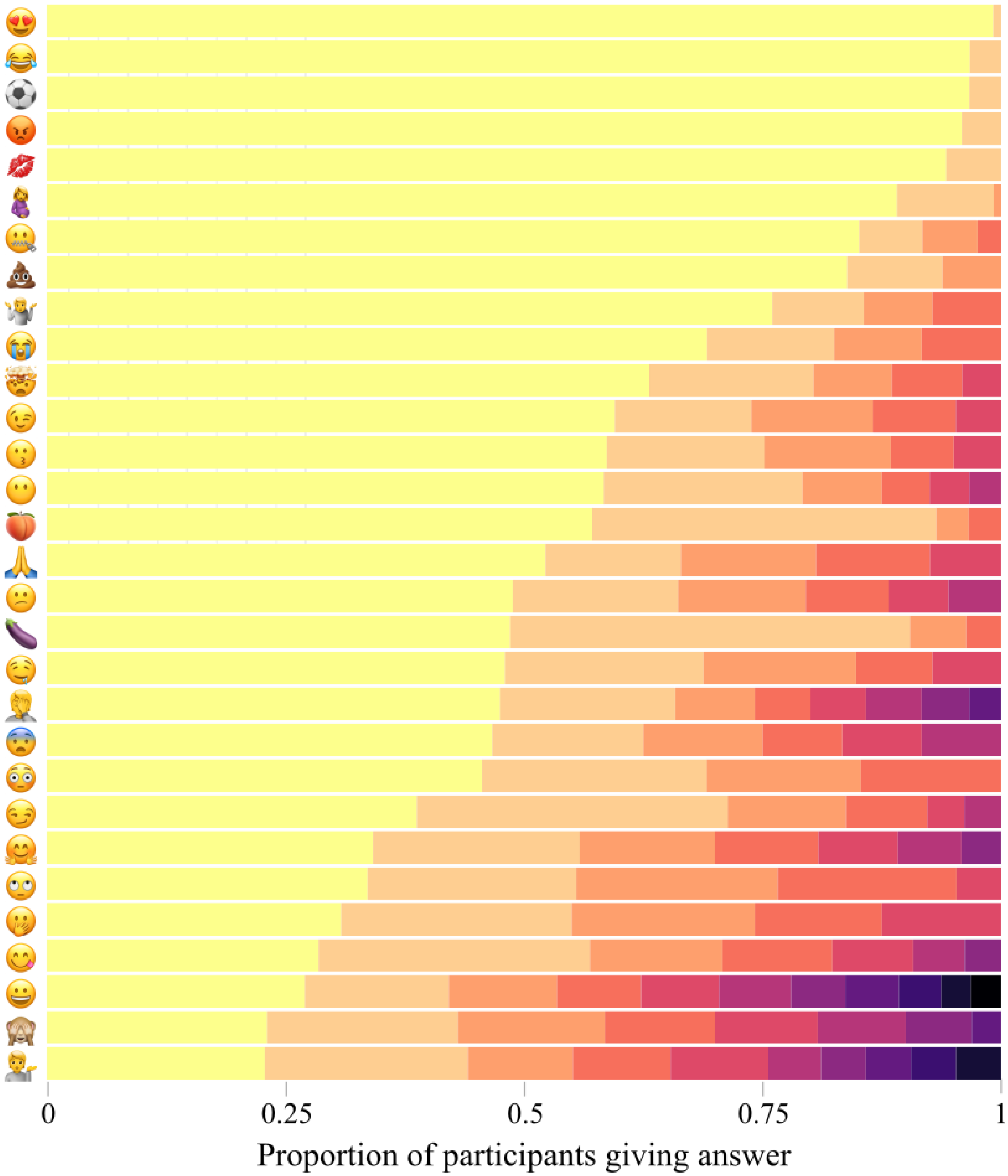

Figure 1. Responses given in the meaning agreement survey for each emoji tested, sorted by proportion of participants who gave the most common response. Within each row, the responses are sorted from most-commonly given (left side, lighter shade) to least-commonly given (right side, darker shade).

In order to test whether agreement rate varies significantly across this set of 30 emoji, the data was converted to a binary response (most commonly given answer = 1, any other answer = 0); a logistic regression model was fitted to the data with response as the dependent variable and emoji as an independent variable. This model indicates a significant effect of the specific emoji

($p < .001$), in line with visual observation of the graph above; the variance in the proportion of people who agree on the meaning of a given emoji differs significantly across this set of 30 emoji.

Across all emoji, 58% of responses given matched the primary listed meaning for the emoji on Emojipedia. There was a significant correlation between inter-participant agreement and participant-Emojipedia agreement ($r = 0.69$, $p < .001$); in other words, the responses given for high-agreement emoji in this experiment were likely to match the given emoji definitions on Emojipedia. This furthers the idea that broader, population-level lexicalization is occurring for at least some of these emoji.

An additional analysis investigated the relationship between relative emoji frequency and meaning agreement. Relative emoji frequencies were taken from Unicode's published 2019 data (Daniel, 2020) as it better reflects frequency patterns around the time of data collection. This data did not include actual frequency rates for each emoji but rather sorted them into frequency-ranked groups, from which the frequency rates are roughly estimable. There was a small but not statistically significant correlation between relative emoji frequency and meaning agreement ($r = 0.31$, $p = .09$). Upon observation, this relationship is likely driven by the two most frequent emoji used in this study (face with tears of joy and smiling face with heart eyes) also being the emoji with the highest agreement. Once the steep slope of the Zipfian frequency distribution is descended, any positive correlative trend disappears ($r = -.22$, $p = .25$ with those two data points omitted). The two high frequency high agreement emoji may lend mild support to the idea that increased frequency (and thus exposure) accompanies stronger lexicalization and agreement. Perhaps the nonlinear frequency distributions of the emoji cause this effect to emerge for only the very high frequency emoji.

2.3. Discussion

Our first study sought to answer whether people agree on the meanings of various emoji by asking them directly. Some emoji, like the smiling face with heart eyes, have one clear meaning, shared widely across the population. Others, such as the eggplant, appear to have two clear but distinct meanings. Others yet may be associated less strongly with two or three meanings, with only around a quarter of the participants providing the most-provided answer. Indeed, these possibilities form the ends of a spectrum that also includes degrees of agreement in between.

Offline agreement of emoji meanings suggests a likely-still-underway process of lexicalization for emoji, whereby they gain population-wide meaning agreement and standardization. Just as lexicons can develop for conventionalized form-meaning mappings in comics (Cohn, 2013), traffic signs (Forceville, 2019), memes (Dancygier & Vandelanotte, 2017; Schilperoord & Cohn, 2021) and other visual and multimodal representations, so too can emoji. Indeed, this follows recent views on the lexicon speculating such form-meaning correspondences proliferate across modalities (Jackendoff & Audring, 2020). We here find evidence that these emoji have a range of conventionalized meanings. In some cases, one definition clearly dominates, but in others, our data substantiate the ambiguity of emoji; the peach or eggplant, for example, have their top choices split between their literal and innuendo meanings. While these meanings would often be disambiguated in context, the fact that such polysemy exists for the emoji itself is significant here. To investigate whether these degrees of agreement actually manifest in processing and uptake, Experiment 2 measures entrenchment via reaction times to emoji-phrase pairings.

3. Experiment 2: Emoji entrenchment

Having established the range of agreement for various emoji, we next ask about entrenchment: are emoji meanings lexicalized in memory in a way that reflects their population-level conventionality? To test this, we presented emoji preceded by words that either matched or mismatched the emoji meanings (as indicated by Experiment 1). Participants had to indicate via a timed yes/no judgment task whether the word-emoji pairing matched; we will analyze both the response times and the congruence judgments themselves.

3.1. Methods

3.1.1. Participants

156 people who did not participate in Experiment 1 completed Experiment 2. Informed consent was confirmed by all participants according to approval from (*information removed for anonymization purposes*). Data from participants who began but did not complete the experiment was discarded. Three participants had exceedingly high proportions of outlier trials (> 50%); two of these consistently provided responses under 300 ms and one consistently provided responses over 20 seconds. Data from these participants was removed, leaving a final dataset of 153 participants (average age = 29.3 (SD = 8.85); 62 male, 85 female, 6 other). Participants also completed a basic demographic survey and the aforementioned Emoji Language Fluency questionnaire.

3.1.2. Stimuli

Based on the results of Experiment 1, the most commonly given word/phrase answer for each of the 30 emoji in the agreement survey was selected as the “match” condition answer for each emoji in the experiment. “Mismatch” condition answers were generated by scrambling the other match condition words, none of which were provided as answers for the emoji with which they mismatched. Experimental items were counterbalanced such that each participant would see each emoji only once, but all emoji appeared in both matching and mismatching relationships across lists. This experiment was run in conjunction with another experiment; the stimuli for that experiment act as fillers for the present study. The filler stimuli consisted of longer text phrases and two-emoji combinations but asked for the same match/mismatch judgment.

3.1.3. Procedure

Participants accessed the experiment in an online survey (hosted on Qualtrics), where the timed response experiment was presented using the jspsych plugin (De Leeuw, 2015). Each trial in the experiment consisted of two screens. The first screen was untimed and presented a text word or phrase. When ready, participants pressed a button to move to the second screen, which presented an emoji. On this timed screen, participants indicated whether the emoji matches or mismatches the phrase on the first screen by pressing the corresponding keyboard button. A sample trial is depicted in Figure 2. Trial orders were randomized for each participant.

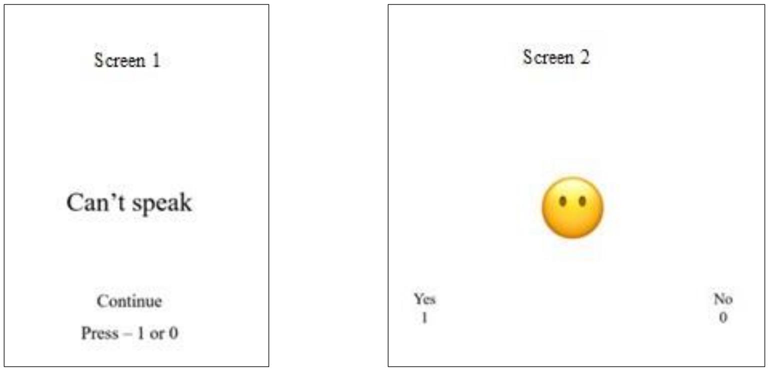

Figure 2. Sample trial from Experiment 2. Screen 2 was the timed response (YES/NO) to whether the emoji matches the text from Screen 1.

3.1.4. Data Analysis

A few extra-long responses were strongly influencing the data for outlier detection, so all responses over 30 seconds long (n=5) were removed from the data. After removing those, we followed a standard outlier removal procedure. Responses under 300 ms (n=6) or over 2.5 standard deviations from the grand mean (n=120) were tagged as outliers and removed from the dataset. In total, 2.8% of data points collected were removed.

3.2. Results

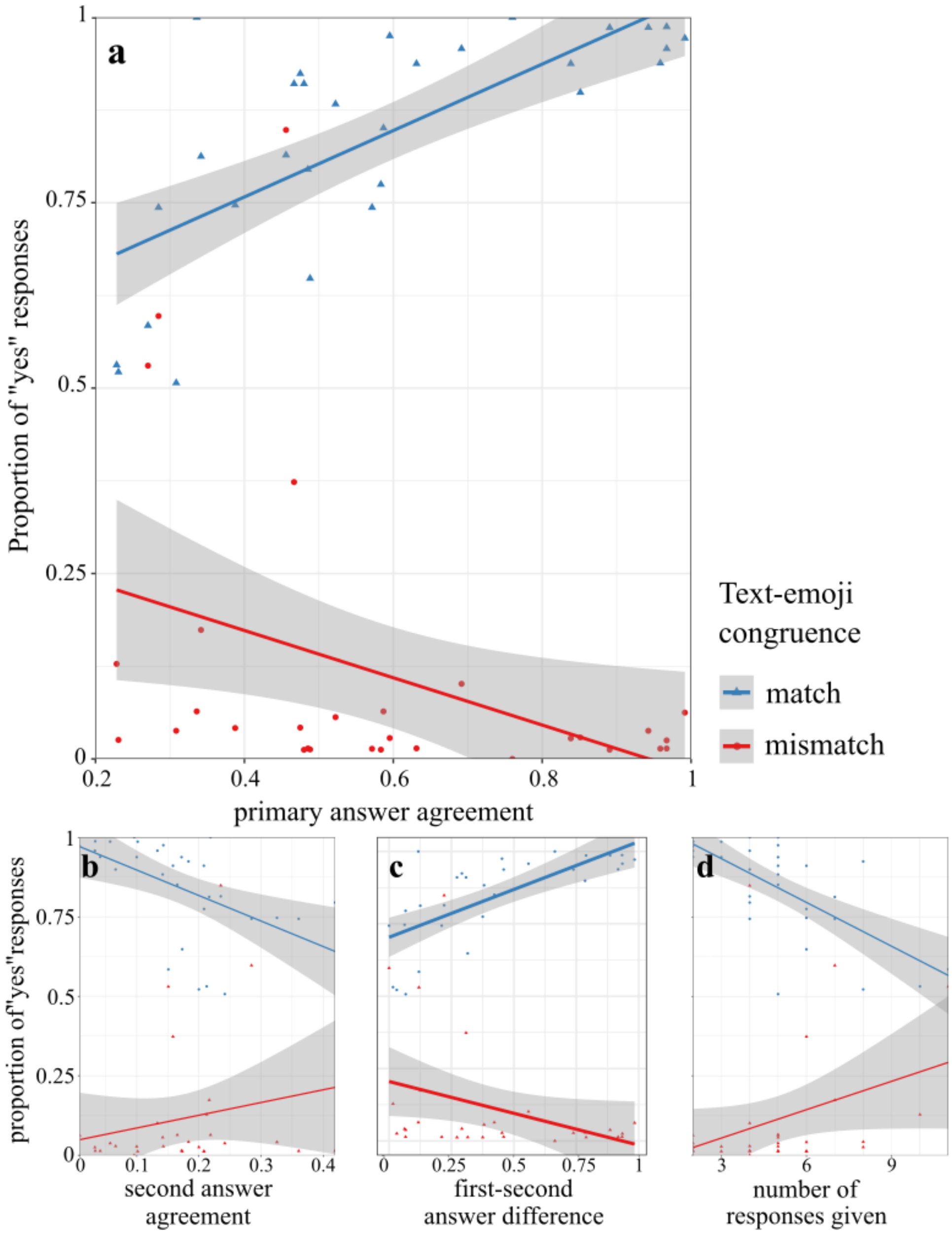

Figure 3. Proportion of “yes” responses for each emoji. Match condition appears in blue, and mismatch condition in red; grey shading represents 95% confidence interval. **a** depicts responses to emoji sorted on the x-axis by the proportion of the most-commonly given answer in Experiment 1; **b** depicts responses by proportion of the second-most-commonly given answer in Experiment 1; **c** depicts responses by difference between proportion of the most- and second-most-commonly given answers in Experiment 1; and **d** depicts responses by the number of different responses given for each emoji in Experiment 1.

Participant responses were analyzed using a logistic mixed effects model, which was run on the binary response data with Text-emoji congruence (match/mismatch), Emoji agreement (from Experiment 1, converted to z-scores), and the interaction as predictors and random intercepts for participant. The results of the model, with match set as the reference level for the congruence factor (dummy coding), are depicted in Figure 3 and presented in Table 1. The model showed significant effects for Emoji agreement, Text-emoji congruence and a significant interaction between them.

Table 1. Summary of logistic mixed effects regression modeling the likelihood of a "yes" response based on emoji agreement and text-emoji congruence. *** $p < .001$

	Estimate	**SE**	**z**	***p***
Intercept	2.04	0.08	24.20	< .001 ***
Emoji agreement	1.06	0.08	12.77	< .001 ***
Text-emoji congruence (Mismatch)	-4.45	0.12	-36.35	< .001 ***
Agreement:Congruence (Mismatch)	-2.02	0.12	-16.31	< .001 ***

As should be expected, text-emoji matches receive more "yes" responses than text-emoji mismatches throughout. As emoji agreement increases, participants get stronger in their responses in both conditions; high-agreement emoji have higher "yes" rates than low-agreement emoji in match condition and high-agreement emoji have higher "no" rates than low-agreement emoji in mismatch condition. In other words, as population-wide emoji agreement increases, participant accuracy increases as well, regardless of whether the item is a match or mismatch.

This trend persists across further analysis as well. When responses are compared to the proportion of participants who provided the second-most commonly provided answer in Experiment 1, a significant interaction ($z = 10.82$, $p < .001$) again emerges whereby a higher proportion for the second-most provided answer yields a decrease in "yes" responses in match condition and an increase in "yes" responses in mismatch condition. In addition, a larger difference between the first- and second-most provided emoji yields an increase in "yes" responses in match condition and a decrease in "yes" responses in mismatch condition ($z = -15.42$, $p < .001$). Lastly, having more answers provided during Experiment 1 leads to a decrease in "yes" responses in match condition and an increase in mismatch condition ($z = 14.89$, $p < .001$). Altogether, these results demonstrate that increased population-wide meaning ambiguity yields slower responses in this timed task, suggesting a slower match-analysis process.

Figure 4 shows the response time data for text-emoji pairs, plotted by Emoji agreement from Experiment 1.

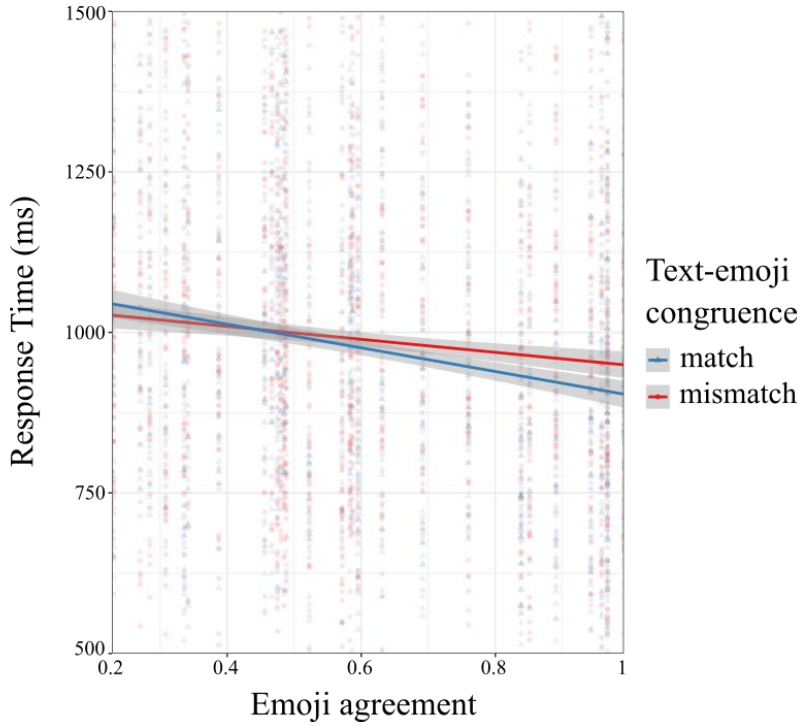

Figure 4. Scatterplot with regression lines for response times for each emoji by agreement tested by offline emoji agreement proportion from the pre-test. Match condition appears in blue, and mismatch condition in red; grey shading represents 95% confidence intervals. Only a portion of the y-axis is shown to ensure the regression lines are visibly clear.

Response times were analyzed using a linear mixed effects model, again with Text-emoji congruence (match/mismatch), Emoji agreement (from Experiment 1) and the interaction as predictors, and participant as a random effect. Again, significant effects appeared for both Emoji agreement and Text-emoji congruence as well as the interaction between them. The results of the model, with match set as the reference level for congruence (dummy coding), are shown in Table 2.

Table 2. Summary of linear mixed effects regression modeling response time based on emoji agreement and text-emoji congruence. *** $p < .001$

	Estimate	**SE**	**t**	***p***
Intercept	1415.86	37.93	37.32	< .001 ***
Emoji agreement	-200.85	14.82	-13.55	< .001 ***
Text-emoji congruence (Mismatch)	-72.17	20.55	-3.51	< .001 ***
Agreement:Congruence (Mismatch)	88.20	20.85	4.23	< .001 ***

Higher agreement emoji led to significantly faster response times overall, but a significant interaction modulates this effect across the congruence conditions. The inverse correlation between agreement and response time is significant for both match and mismatch conditions, though the correlation is significantly stronger in the match condition, as determined by a post-hoc test of estimated marginal means trends. That is to say, emoji agreement yields faster responses times in general, even more so when the text is a match. Visually evident in

Figure 4, larger match-mismatch differences are observed in the higher range of the agreement continuum.

The secondary analyses reveal a similar pattern for response times as well. Increased ambiguity/polysemy, as determined through responses given in Experiment 1 (secondary answer proportion ($t = -3.77$, $p < .001$), difference between first- and second-most commonly provided answer ($t = 4.35$, $p < .001$), number of responses given ($t = -4.74$, $p < .001$)), again leads to slower response times, especially in match condition.

3.3. Discussion

This experiment asked whether the degree of agreement for emoji meanings would manifest in real time processing, as suggested by response times. In other words, how entrenched are emoji meanings? It was evident that degree of agreement – established offline in the prior agreement survey of Experiment 1– manifests as entrenchment in online processing as well.

First, we observed that participants consistently responded that emoji that matched their text were congruous compared to those that mismatched. This reinforces the findings in Experiment 1, suggesting that participants have clear preferences for the meanings that align with different emoji. These responses correlated significantly with the offline agreement ratings established by Experiment 1: the more that emoji meanings were agreed upon, the more likely that participants responded that the matching texts were congruous with their emoji (or that mismatching texts were incongruent with their emoji). Thus, the degree to which participants recognize emoji as matching their meaning corresponds to the degree to which those meanings are agreed-upon by a broader population, a pattern consistent with the idea that these population-wide trends correspond to degrees of entrenchment in an individual's lexicon.

This idea is further supported by secondary analyses. We observe a significant interaction between responses in Experiment 2 and the proportion of participants who gave the second-most commonly provided answer in Experiment 1; the same goes for the difference between the proportion of participants who gave the first-most and second-most commonly provided answers in Experiment 2. For some emoji, there was near-equal representation of the top two answers. For these emoji, whose meanings carry more potential for ambiguity, we observe less strong matching in match condition and less strong mismatching in mismatch condition. The same pattern holds regarding how many different responses were given in Experiment 1. These relationships are particularly striking given that secondary meanings and the total number of answers provided in Experiment 1 are not present in the words preceding the emoji in Experiment 2; yet this knowledge implicitly affects participants' performance. All of these provide additional support for the idea that emoji are subject to population-wide meaning-related phenomena.

Second, the time it took participants to make these responses further supports the graded entrenchment of emoji meanings. A significant inverse relationship arose between emoji meaning agreement and response times to their congruence: emoji with higher agreement yielded faster responses. The secondary metrics of ambiguity reflect the same pattern, again supporting the idea that these meaning phenomena bear out in real-time processing as well.

The correlation between agreement and response time was present for both matching and mismatching labels, though the correlation was stronger when the verbal meaning matched. This agreement by congruence interaction perhaps suggests that participants are more likely to engage in a more thorough meaning verification search when faced with mismatching contents, even if they are relatively certain of the emoji's meaning already. Though not tested directly here, this could point to the general flexibility of emoji meaning and a tendency to search for compatibility, even when such compatibility is neither immediately detected nor ultimately determined. This would help characterize the steeper decrease in response time as agreement increases for the matching items, as a high agreement match is the scenario least warranting such a search. Low agreement emoji appear to yield this search regardless of congruence.

Together, these findings suggest that people can rapidly access consistent meanings for emoji – when such a consistent meaning is present. Emoji with less variable meanings are accessed quicker and more accurately than emoji with more variable meanings. This consistency seems to align with the notion of entrenchment, as evidenced by faster reaction times to emoji with higher naming agreement. Entrenchment here offers online evidence for the lexicalization process observed offline in Experiment 1. Individualized and population-wide factors can both affect lexical access (Schmid, 2007, 2015); the fact that there was a significant correlation between offline agreement and online processing points to the population-level dynamics here. We take these findings as evidence that emoji are entrenched in a visual lexicon in a manner similar to words, and which are in line with other studies showing familiarity-based entrenchment of other emoji-like visual representations (Cohn & Foulsham, 2022).

4. General Discussion

This study investigated the processing of emoji, specifically focusing on the fundamental aspect of linguistic structure of lexicalization. Overall, we found evidence that emoji in isolation can be entrenched as lexical items, though it is important to note this entrenchment varies across emoji. Emoji exhibit a fairly wide range of agreement levels, and the ones with less consistent agreement yield slower response times in processing. Indeed, the degree to which these emoji demonstrate polysemy and ambiguity correlates significantly with response consistency and response time, as measured by number of answers given, proportions of participants providing the second-most commonly provided answer, and difference between first- and second-most commonly provided answers. The graded response demonstrated in our studies thus suggests that entrenchment of emoji meaning varies on the basis of their conventionality. These findings support the idea that emoji can be entrenched as lexical items, though it is clear that this does not happen equally for every emoji.

These results also contribute to the already-complex understanding of polysemy and homonymy processing. Though the emoji here are likely better classified as examples of polysemy, they appear to show the opposite effect typically found to word-based polysemy, wherein increased polysemy yields easier lexical access and a processing advantage. Here we find evidence of a dominance effect like that observed in homonym words – emoji with a stronger, more-agreed-upon meaning yielded faster response times than emoji with a less dominant single meaning. The apparent difference between emoji polysemy and word polysemy processing may be due to task differences (Eddington & Tocowicz, 2015), as this direct match/mismatch task is not easily achievable with words as target items. Future research could aim to more precisely situate polysemy effects across a multimodal lexicon.

These emoji patterns are consistent with the multimodal parallel architecture of the lexicon, in which aspects of graphics are recognized as having conventionalized meanings, including the drawings in comics (Cohn, 2013), traffic signs (Forceville, 2019), memes (Dancygier & Vandelanotte, 2017; Schilperoord & Cohn, 2021), and many others. In addition, the bodily modality also has lexicalized expressions outside of full sign languages with "gestural emblems" (McNeill, 1992), such as thumbs up, peace signs, and the middle finger, many of which themselves have become instantiated in graphic form as emoji. To further invoke comparisons with gesture (Cohn, Engelen, & Schilperoord, 2019; Feldman et al., 2017; Grosz, Kaiser, & Pierini, 2021; McCulloch & Gawne, 2018), emoji may thus be comparable to graphic emblems, which have also been posited for common simple conventions of drawing, such as stick figures (Cohn, 2013). This analysis is consistent with ideas of a multifaceted lexicon that consists of various sorts of form-meaning mappings stored in long term memory. The lexical model of Jackendoff and Audring (2020), for example, supports this organization, positing uniform organizational setup in memory across domains.

The foundations of a multimodal parallel lexicon that includes emoji forms and meanings built here can guide further investigations into the processing of emoji meanings. Future work could more directly address the notions of polysemy/homonymy approached here, specifically how cross-context flexibility of emoji meanings affects processing of those emoji in contexts. Given the parallel architecture adopted here, future work could also aim to explore grammatical elements of this lexicalized relationship. The establishment of these lexical items in long term memory include not just the form-meaning link but also structural components. These investigations could include emoji-only sequences as well as more detailed looks at the combinatorial parameters affecting sequences that consist of content from multiple modalities.

Conflict of Interest

One author has advised in the creation of new emoji and in commercial use of emoji.

References

Barach, E., Feldman, L. B., & Sheridan, H. (2021). Are emojis processed like words?: Eye movements reveal the time course of semantic processing for emojified text. *Psychonomic Bulletin and Review, 28,* 978-991. https://doi.org/10.3758/s13423-020-01864-y,

Blumenthal-Dramé, A. (2012). *Entrenchment in usage-based theories*. Berlin: De Gruyter Mouton.

Blumenthal-Dramé, A. (2017). Entrenchment from a psycholinguistic and neurolinguistic perspective. In H.-J. Schmid (Ed.), *Entrenchment and the psychology of language learning: How we reorganize and adapt linguistic knowledge* (pp. 129–152). De Gruyter Mouton.

Boutet, I., LeBlanc, M., Chamberland, J., & Collin, C. (2021). Emojis influence emotional communication, social attributions, and information processing. *Computers in Human Behavior*, *119*(January), 106722. https://doi.org/10.1016/j.chb.2021.106722

Brocher, A., Koenig, J.-P., Mauner, G., & Foraker, S. (2018). About sharing and commitment:

the retrieval of biased and balanced irregular polysemes. *Language, Cognition and Neuroscience*, *33*(4), 443–466.

Christofalos, A., Feldman, L., & Sheridan, H. (2022). Semantic Congruency Facilitates Memory for Emojis. *Proceedings of the The Fifth International Workshop on Emoji Understanding and Applications in Social Media*, 63–68. https://doi.org/10.18653/v1/2022.emoji-1.7

Cohn, N. (2013). *The visual language of comics: Introduction to the structure and cognition of sequential images*. London: Bloomsbury.

Cohn, N. (2016). A multimodal parallel architecture : A cognitive framework for multimodal interactions. *Cognition*, *146*, 304–323. https://doi.org/10.1016/j.cognition.2015.10.007

Cohn, N., Engelen, J., & Schilperoord, J. (2019). The grammar of emoji? Constraints on communicative pictorial sequencing. *Cognitive Research: Principles and Implications*, *4*(1), 1-18. https://doi.org/10.1186/s41235-019-0177-0

Cohn, N., & Foulsham, T. (2022). Meaning above (and in) the head: Combinatorial visual morphology from comics and emoji. *Memory and Cognition*, *50*(7), 1381-1398. https://doi.org/10.3758/s13421-022-01294-2

Cohn, N., Roijackers, T., Schaap, R., & Engelen, J. (2018). Are emoji a poor substitute for words? Sentence processing with emoji substitutions. *Proceedings of the 40th Annual Conference of the Cognitive Science Society*, 1524–1529.

Cohn, N., & Schilperoord, J. (2022). Remarks on Multimodality: Grammatical Interactions in the Parallel Architecture. *Frontiers in Artificial Intelligence*, *4*(January), 1–21. https://doi.org/10.3389/frai.2021.778060

Częstochowska, J., Gligoric, K., Peyrard, M., Mentha, Y., Bien, M., Grutter, A., … West, R. (2022). *On the Context-Free Ambiguity of Emoji*. In C. Budak, M. Cha, & D. Quercia (Eds.), Proceedings of the Sixteenth International AAAI Conference on Web and Social Media (pp. 1388-1392). Retrieved from http://arxiv.org/abs/2201.06302

Dancygier, B., & Vandelanotte, L. (2017). Internet memes as multimodal constructions. *Cognitive Linguistics*, *28*(3), 565–598. https://doi.org/https://doi.org/10.1515/cog-2017-0074

Daniel, J. (2020). Emoji Frequency: The Most Frequent Emoji — 2019. Unicode. Retrieved from https://home.unicode.org/emoji/emoji-frequency/

De Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a Web browser. *Behavior Research Methods*, *47*(1), 1–12.

Divjak, D., & Caldwell-Harris, C. L. (2015). Frequency and entrenchment. In E. Dabrowska, & D. Divjak (Eds.), *Handbook of Cognitive Linguistics* (Vol. 39, pp. 53–75). https://doi.org/10.1515/9783110292022-004

Eddington, C. M., & Tokowicz, N. (2015). How meaning similarity influences ambiguous word processing: the current state of the literature. *Psychonomic Bulletin and Review*, *22*(1), 13–37. https://doi.org/10.3758/s13423-014-0665-7

Emojipedia. (2022). Retrieved from https://emojipedia.org/

emojitracker. (2021). Retrieved from 2021 website: https://emojitracker.com/

Feldman, L. B., Aragon, C. R., Chen, N.-C., & Kroll, J. F. (2017). Emoticons in text may function like gestures in spoken or signed communication. *Behavioral and Brain Sciences*, *40*.

Foraker, S., & Murphy, G. L. (2012). Polysemy in sentence comprehension: Effects of meaning dominance. *Journal of Memory and Language*, *67*(4), 407–425.

Forceville, C. (2019). Reflections on the Creative Use of Traffic Signs' "Micro-Language. In A.

Benedek & K. Nyíri (Eds.), *Perspectives on Visual Learning* (pp. 103–113). Budapest: Hungarian Academy of Sciences.
Foroudi, P., Melewar, T. C., & Gupta, S. (2014). Linking corporate logo, corporate image, and reputation: An examination of consumer perceptions in the financial setting. *Journal of Business Research*, *67*(11), 2269–2281. https://doi.org/10.1016/j.jbusres.2014.06.015
Grosz, P., Kaiser, E., & Pierini, F. (2021). Discourse anaphoricity and first-person indexicality in emoji resolution. *Proceedings of Sinn Und Bedeutung*, *25*.
Holtgraves, T., & Robinson, C. (2020). Emoji can facilitate recognition of conveyed indirect meaning. *PLoS ONE*, *15*(4), 1–13. https://doi.org/10.1371/journal.pone.0232361
Jackendoff, R. (2002). *Foundations of language: Brain, meaning, grammar, evolution*. Oxford: Oxford University Press.
Jackendoff, R., & Audring, J. (2020). *The texture of the lexicon: relational morphology and the parallel architecture*. Oxford University Press, USA.
Kaiser, E., & Grosz, P. G. (2021). Anaphoricity in emoji: An experimental investigation of face and non-face emoji. *Proceedings of the Linguistic Society of America*, *6*(1), 1009–1023.
Klepousniotou, E., Pike, G. B., Steinhauer, K., & Gracco, V. (2012). Not all ambiguous words are created equal: An EEG investigation of homonymy and polysemy. *Brain and Language*, *123*(1), 11–21.
Ladewig, S. (2020). *Integrating gestures: The dimension of multimodality in cognitive grammar*. Berlin: De Gruyter Mouton.
Lai, H. L., & Chung, S. F. (2018). Color polysemy: Black and White in Taiwanese languages. *Taiwan Journal of Linguistics*, *16*(1), 95–130. https://doi.org/10.6519/TJL.2018.16(1).4
Langacker, R. W. (1987). *Foundations of Cognitive Grammar: descriptive application. Volume 2* (Vol. 2). Stanford university press.
McCulloch, G., & Gawne, L. (2018). Emoji grammar as beat gestures. *CEUR Workshop Proceedings*, *2130*, 3–6.
McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago: University of Chicago Press.
Meade, G., & Coch, D. (2017). Word-pair priming with biased homonyms: N400 and LPC effects. *Journal of Neurolinguistics*, *41*, 24–37.
Miller, H., Kluver, D., Thebault-Spieker, J., Terveen, L., & Hecht, B. (2017). Understanding Emoji Ambiguity in Context: The Role of Text in Emoji-Related Miscommunication. *Proceedings of the Eleventh International AAAI Conference on Web and Social Media*, 152–161. Retrieved from https://www.aaai.org/ocs/index.php/ICWSM/ICWSM17/paper/view/15703
Miller, H., Thebault-Spieker, J., Chang, S., Johnson, I., Terveen, L., & Hecht, B. (2016). “Blissfully happy” or “ready to fight”: Varying Interpretations of Emoji. *International AAAI Conference on Web and Social Media*, (Icwsm), 259–268. https://doi.org/10.1089/cyber.2011.0179
Robus, C. M., Hand, C. J., Filik, R., & Pitchford, M. (2020). Investigating effects of emoji on neutral narrative text: Evidence from eye movements and perceived emotional valence. *Computers in Human Behavior*, *109*(April), 106361. https://doi.org/10.1016/j.chb.2020.106361
Schilperoord, J., & Cohn, N. (2021). Let there be.. visual optimal innovations: making visual meaning through Michelangelo’s The Creation of Adam. *Visual Communication*. https://doi.org/10.1177/14703572211004994

Schmid, H.-J. (2007). Entrenchment, salience, and basic levels. *The Oxford Handbook of Cognitive Linguistics*, *117*, 138.

Schmid, H.-J. (2015). A blueprint of the entrenchment-and-conventionalization model. *Yearbook of the German Cognitive Linguistics Association*, *3*(1), 3–26.

Schmid, H.-J. (2017). A framework for understanding linguistic entrenchment and its psychological foundations. In *Language and the Human Lifespan Series. Entrenchment and the psychology of language learning: How we reorganize and adapt linguistic knowledge.* (pp. 9–35). https://doi.org/10.1037/15969-002

Shardlow, M., Gerber, L., & Nawaz, R. (2022). One emoji, many meanings: A corpus for the prediction and disambiguation of emoji sense. *Expert Systems with Applications*, *198*(March), 116862. https://doi.org/10.1016/j.eswa.2022.116862

Sick, J., Monteleone, E., Pierguidi, L., Ares, G., & Spinelli, S. (2020). The meaning of emoji to describe food experiences in pre-adolescents. *Foods*, *9*(9), 1307.

Talmy, L. (1985). Lexicalization patterns: Semantic structure in lexical forms. I*n T. Shopen (Ed.), Language Typology and Syntactic Description Volume III: Grammatical Categories and the Lexicon (pp. 57-149). Cambridge: Cambridge University Press.*

Tang, M., Chen, B., Zhao, X., & Zhao, L. (2020). Processing network emojis in Chinese sentence context: An ERP study. *Neuroscience Letters*, *722*, 134815. https://doi.org/10.1016/j.neulet.2020.134815

Weinreich, U. (1964). Webster's third: A critique of its semantics. *International Journal of American Linguistics*, *30*(4), 405–409.

Weissman, B. (2019a). Peaches and eggplants or. . . something else? The role of context in emoji interpretations. *Proceedings of the Linguistic Society of America*, *4*(1), 29. https://doi.org/10.3765/plsa.v4i1.4533

Weissman, B. (2019b). Emojis in sentence processing: An electrophysiological approach. *The Web Conference 2019 - Companion of the World Wide Web Conference, WWW 2019*, *1*, 478–479. https://doi.org/10.1145/3308560.3316544

Weissman, B., & Tanner, D. (2018). A strong wink between verbal and emoji-based irony: How the brain processes ironic emojis during language comprehension. *PLoS ONE*, *13*(8). https://doi.org/https://doi.org/10.1371/journal.pone.0201727

Wicke, P., & Bolognesi, M. (2020). Emoji-based semantic representations for abstract and concrete concepts. *Cognitive Processing*, *21*(4), 615–635. https://doi.org/10.1007/s10339-020-00971-x

Wijeratne, S., Balasuriya, L., Sheth, A., & Doran, D. (2016). Emojinet: Building a machine readable sense inventory for emoji. *International Conference on Social Informatics*, 527–541.

Supplementary Materials

https://osf.io/3m2r5/?view_only=a5f99b3d5db84a28ba5e8791ddb85ece

This link contains data from the experiments, R code used in analysis, the stimulus set, and the Emoji Language Fluency questionnaire.