

Anaphoricity in emoji: An experimental investigation of face and non-face emoji

Elsi Kaiser & Patrick Georg Grosz*

Abstract. Emoji are widely used, but have received relatively little attention in psycholinguistic research. Upon encountering a message consisting of both text and emoji, readers presumably construct some link between emoji and text. Based on a psycholinguistic study on text-emoji relations, we argue for (at least) two types of emoji-text dependencies, related to referential dependencies known to exist in the linguistic domain, namely (i) the dependency between an expressive (e.g. *wow*, *damn*, *f*king*) and the individual whose opinion it expresses, and (ii) the dependency between a pronoun (or other pro-form) and its antecedent. We extend the discussion of these dependencies to emoji, and provide experimental data that face emoji (e.g. 😊😓😞) resemble expressives in that they tend to be interpreted as expressing the opinion of a salient experiencer, while action emoji (e.g. 🍌🏏🍰) are interpreted based on principles of discourse coherence (e.g. discourse relations like explanation), similar to what coherence-based accounts of pronoun resolution predict.

Keywords. emoji; digital communication; anaphora; expressives; discourse coherence; first-person orientation; perspective shift; reference resolution; psych verbs

1. Introduction. Emoji (e.g. 😊😓😞🏏🍰) are frequently used in present-day digital communication, including text messages, social media posts and email (see Bai et al. 2019 for a recent overview). Readers who encounter a message consisting of both text and an emoji, as in (1), presumably construct some kind of link between an emoji and the text that it accompanies.

(1) Ana sent a surprise birthday present to Betty 😊

Thus, emoji can be regarded as offering a new window into investigating dependency formation. The research presented here provides experimental data building on our theoretical proposal (Grosz, Kaiser & Pierini to appear) that there exist (at least) two distinct types of emoji-text dependencies, namely those involving *face emoji* (e.g. 😊😓😞) and those involving non-face emoji depicting actions or action-related objects (e.g. 🍌🏏🍰), which we call *action emoji*. We propose that both face and action emoji involve anaphoric dependencies, but of different types: Face emoji link to an experiencer (often the first-person author), akin to expressives (*wow*, *damn*), while action emoji are interpreted based on principles of discourse coherence, akin to (third-person) pronoun resolution. (On face emoji see also Grosz, Greenberg, De Leon & Kaiser 2021.)

1.1. FACE AND ACTION EMOJI. Before discussing our experiment, it is important to clarify the kinds of emoji we are investigating, namely (i) face emoji and (ii) action emoji. By face emoji, we mean the (mostly) yellow discs (‘smileys’) with stylized facial expressions, such as 😊😓😞. Emojipedia calls these “yellow balls of emotion.”¹ We are not investigating the ‘person emoji’

* This material is based upon work supported by the National Science Foundation under Grant No. BCS-1749612. Many thanks are due to Prim Phoolsombat for help with stimulus creation and implementation as well as Francesco Pierini for helpful discussion. We would also like to thank the audience at the Annual Meeting of the Linguistic Society of America for helpful comments and feedback. Authors: Elsi Kaiser, University of Southern California (emkaiser@usc.edu) & Patrick Grosz, University of Oslo (p.g.grosz@iln.uio.no).

¹ <https://blog.emojipedia.org/emoji-use-in-the-new-normal/> (accessed 5/5/21)

such as 🧑🏻🧑🏼🧑🏽🧑🏾. Furthermore, the present work only focuses on face emoji conveying emotional/affective states (e.g. happy, worried). The face emoji that we investigate can perhaps be viewed as a subset of a broader class of affective emoji, which could also include emoji such as 👍👉💕💕💕; we leave the precise scope of this class as a question for future research.

The second class of emoji included in our study are action emoji. These are emoji that look like objects (e.g. 🏀) or people (e.g. 🏄), and describe an activity, state or property related to the depiction. For example, 🏄 can mean ‘surfing, to surf’ or ‘to be a surfer.’ Semantically, action emoji are not all activities in the spirit of Vendler (1957); some can denote states or properties. We remain agnostic as to whether all action emoji denote eventualities in the spirit of Davidson (1967). These open issues are not crucial for the basic claims made in this paper.

The Unicode Emoji Standard includes other types beyond face and action emoji (e.g. weather emoji 🌞🌨️, animals 🐕🐶, pointing emoji 👉👈). For now, we put these aside.

Recent analyses show that face emoji are used more often than other types of emoji. For example, according to an analysis by Twitter, all of the top 10 most frequently tweeted emoji in 2020 were face emoji² (see also emojitracker.com for current updates). The ease of using face emoji presumably stems from their resemblance to human facial expressions (though see Weiß et al. 2020 for brain imaging work on differences in responses to real faces vs. emoji). However, although there exists a large literature on facial expressions (e.g. Tomkins & McCarter 1964, Ekman, Friesen & Ellsworth 1972, Russell & Fernández-Dols 1997, Keltner & Cordaro 2017), in this paper we make no specific claims about the precise nature of the relation between face emoji and human emotions (e.g. whether the relation is iconic or symbolic). Although the details of this mapping are an important issue, they are not central to the aims of the current work.

2. Emoji as an object of study: Prior research. At first glance, some might be surprised at emoji being considered a meaningful object of study for (psycho)linguistics, given that they are a human-created artifact. Present-day emoji evolved from the emoticons (e.g. :-)) of the 1980s (see e.g. Dresner & Herring 2010, Garrison et al. 2011). The creation of emoji (small pixel-based images) is attributed to Shigetaka Kurita in 1999. Apple added an emoji keyboard to iOS in 2011; Android did so in 2013. Despite (or because of) emoji being a human invention, their use has soared in the past decade: By some estimates, nearly one in five tweets in 2020 contained at least one emoji; on Facebook, over 700 million emoji are used in posts every day; on Instagram, half of comments included an emoji by 2015, and in 2017, an average of 5 billion emoji were sent daily on Facebook Messenger. Emoji are a recent but highly prevalent phenomenon.³

The sky-rocketing frequency of emoji use shows that they help satisfy communicative needs by filling in the gaps otherwise present in non-face-to-face communication (see e.g. Godin 1993, Evans 2017), and – we would argue – studying them can offer insights into the linguistic aspects of human communication. In recent years, emoji have been investigated from a wide variety of perspectives, including computer science (e.g. LeCompte & Chen, 2017), marketing and communication (e.g. Luangrath et al. 2017, Jaeger et al. 2019), psychology (e.g. Li et al. 2018), education (e.g. Dunlap et al. 2016) and health communication (e.g. Troiano & Nante 2018). A recent overview is provided by Bai et al. (2019). There is also a rapidly growing body of work on emoji in linguistics (e.g. Evans 2017 for recent discussion). One of the questions that linguists

² https://blog.twitter.com/en_us/topics/insights/2020/spending-2020-together-on-twitter.html (accessed 5/5/21)

³ <https://blog.emojipedia.org/emoji-use-in-the-new-normal/>, <https://emojipedia.org/stats/>, <https://worldemojiday.com/statistics>, <https://blog.emojipedia.org/5-billion-emojis-sent-daily-on-messenger/> (accessed 5/5/21)

have investigated has to do with the combinatorial properties of strings of emoji. For example, in recent work, Cohn et al. (2019) used experimental results to conclude that emoji have only restricted combinatorial properties and sequences of emoji lack the grammatical structure of sentences (However, see Gerke & Storoshenko 2018 for observations that people’s native language may influence their emoji ordering preferences.)

In the present paper, we focus not on the relations between emoji, but rather on the relation between emoji and the text that they accompany – thus, we are interested in the nature of the emoji-text relation, rather than emoji-emoji combinations. One approach to this relation comes from the work of Gawne & McCulloch (2019), Pasternak & Tieu (2020) and Pierini (to appear), who analyze some emoji as digital counterparts of the gestures and facial expressions that accompany spoken language. Relatedly, Maier (2020) proposes that both face emoji and facial expressions should be analyzed as expressives (see also Grosz, Greenberg, De Leon & Kaiser 2021). Related to the idea of emoji resembling facial expressions is the brain-imaging work by Weissman & Tanner (2018) and Weissman (2019) on how face emoji can be used to indicate illocutionary force (see also Dresner & Herring 2010), or signal irony.

While our work builds on these prior investigations, we approach the question from a somewhat different angle by zeroing in on the nature of the dependency between the emoji and the text, using insights from prior linguistic work about what kinds of reference-related dependencies exist between various linguistic elements, including the dependencies between (i) subjective expressions and the attitude-holders to which they are anchored and between (ii) pronouns and their antecedents. We outline our approach in more detail in the next section.

It is worth emphasizing that we make no claims about emoji being linguistic expressions. Although the present research investigates the relation between emoji and linguistic elements, it does not make any claims about whether emoji should be integrated into the linguistic system. Indeed, it may well be the case that the relation between emoji and language is ‘indirect’ in the sense that it is mediated by general human cognition.

3. Our proposal: Two types of dependencies. The psycholinguistic experiments reported in the present paper test the proposal, based on our earlier corpus-based work in collaboration with Francesco Pierini (Grosz, Kaiser & Pierini to appear) that both face and action emoji involve anaphoric dependencies (i.e., can be linked to preceding linguistic content) but in different ways: For face emoji, our proposal is that they resemble **expressives** (e.g. *wow*, *yay*, *damn*) which are typically interpreted as expressing the attitudes of the first-person speaker but can, in some contexts, shift to another salient experiencer (e.g. Amaral et al 2007, Harris & Potts 2009, Lasersohn 2005). For action emoji, our claim is that they are interpreted based on principles of discourse coherence (e.g. relations like Explanation, Elaboration), thus resembling **third-person pronouns** (e.g. Kehler 2002, Kehler & Rohde 2013, see also Asher & Lascarides 2003).

In the rest of this section, we provide an overview of relevant linguistic work on expressives (Section 3.1) and discourse coherence and pronouns (Section 3.2) that provides the foundation for our proposal, and then discuss our predictions for the interpretation of face and action emoji (Section 3.3), focusing especially on the configurations we tested experimentally (Section 4).

3.1. EXPRESSIVES. There exist various kinds of subjective expressions in language that convey someone’s opinion/feelings, including expressives such as *wow* and *damn* which are by default anchored to the first-person author (e.g. Potts 2007, Rett to appear; see Foolen 2015 for an overview). Other subjective expressions such as predicates of personal taste (PPTs, e.g. *fun*, *amazing*) are also by default interpreted as having the author as the attitude-holder (e.g. Lasersohn 2005). Examples are in (2). In these cases, the author can be inferred to be the one with the relevant

kind of first-hand experience that triggers the affective response or judgement reflected by the subjective expression (see e.g. Ninan 2014, Bylinina 2014, McNally & Stojanovic 2017 on the ‘experiencer = attitude-holder’ link).

- (2) a. Damn, I left my keys in the car. b. Elsi was on time, wow!
c. That rollercoaster was fun. d. She is amazing.

Although these kinds of subjective expressions typically express the opinion or attitude of the first-person author, this is not the only possibility. The attitude holder can shift away from the author in some contexts (e.g. Amaral et al 2007, Harris & Potts 2009, Kaiser 2015). What is particularly relevant for the data presented in this paper is the finding from Kaiser & Herron Lee (2017, 2018) that predicates of personal taste (e.g. *wonderful* in (3)) can shift to an *experiencer argument* if one is present in the linguistic context. In a series of experiments, Kaiser & Herron Lee found that in sentences like (3) – when presented as extracts from a narrative – predicates of personal taste (*wonderful*) can be interpreted as expressing the opinion of the referent realized with the thematic role of experiencer (here, David), in addition to the opinion of the author.

- (3) David_{EXPERIENCER} listened to John_{THEME}. He was wonderful.

Although these prior studies tested predicates of personal taste and did not investigate expressions like *wow* or *damn*, one might expect that – especially in light of prior work on the relation between experiencers and attitude-holders/judges – expressives might also exhibit this preference to ‘glom onto’ an experiencer as their potential attitude-holder. In sum, prior linguistic work (i) shows that subjective expressions are by default interpreted as expressing the first-person author’s attitude and (ii) gives reason to believe that a linguistically-expressed *experiencer argument* can also be interpreted as an attitude-holder.

3.2. THIRD-PERSON ANAPHORIC ELEMENTS. Interpretation of third-person pronouns is known to be guided by factors such as topicality/prominence (e.g. Ariel 1990, Gundel et al. 1993) – typically associated with subjecthood in English – as well as discourse coherence (e.g. Hobbs 1979, Kehler 2002, Kehler & Rohde 2013). More specifically, and most relevant for this paper, there exists a large literature showing that psych verbs (e.g. *impress*, *admire*, *annoy*) create an expectation for an upcoming explanation (e.g. Rohde 2008) that starts with mention of the *stimulus argument* (e.g. work on implicit causality by Garvey & Caramazza 1974, Hartshorne & Snedeker 2013, Bott & Solstad 2014, and many others). To see this, consider the examples in (4). Ex.(4a) uses *impress*, a Stimulus-Experiencer (Stim-Exp) psych verb (subject=stimulus, object=experiencer). Ex.(4b) uses *admire*, an Experiencer-Stimulus (Exp-Stim) psych verb (subject=experiencer, object=stimulus). Prior work shows that in (4a), people tend to interpret *she* as referring to Ana, while in (4b), *she* tends to be interpreted as referring to Betty – in other words, whichever referent is realized as the stimulus argument.

- (4) a. Ana_{STIM} impressed Betty_{EXP} because she... (Stim-Exp verb)
b. Ana_{EXP} admired Betty_{STIM} because she... (Exp-Stim verb)

3.3. ADDING EMOJI TO THE MIX. So far, we have seen that the interpretation of subjective linguistic expressions is sensitive to the presence of an *experiencer argument* (Section 3.1), while the interpretation of third-person pronouns after psych verbs can be sensitive to the presence of a *stimulus argument* (Section 3.2). Let’s now bring emoji back into the picture.

In our earlier corpus- and intuition-driven work (Grosz, Kaiser & Pierini to appear), we proposed that both face and action emoji involve anaphoric dependencies (i.e., can be linked to the

linguistic context) but in different ways. In particular, we hypothesized that face emoji resemble expressives (e.g. *wow*, *yay*, *damn*) in that they typically express the attitudes of the first-person speaker. If we also take seriously prior experimental work showing that presence of a linguistically-realized experiencer argument influences the interpretation of subjective expressions such as PPTs (Kaiser & Herron Lee 2017, 2018), this further leads to the hypothesis that face emoji can also be interpreted as conveying the attitude of an *experiencer* argument, in addition to the speaker. Thus, the prediction is that the face emoji in (5) can be interpreted as expressing the affective state of the author or of the linguistically-realized experiencer argument (Betty in (a) with an Stim-Exp verb, and Ana in (b) with an Exp-Stim verb).

- (5) a. Ana_{STIM} impressed Betty_{EXP} 😊 b. Ana_{EXP} admired Betty_{STIM} 😊

We also hypothesized that action emoji are interpreted based on principles of discourse coherence, thus resembling third-person pronouns. In the present paper, we focus on psych verbs which are known to create an expectation for an explanation relation; see Grosz, Kaiser & Pierini (to appear) for discussion of other coherence relations. The prediction is that after psych verbs, action emoji are interpreted as being linked to the *stimulus* argument, in line with third-person pronouns. To see this more concretely, let's consider (6). The prediction is that in (6), the soccer ball emoji is interpreted as providing information about the *stimulus* (Ana in (a) and Betty in (b)) – presumably that she is a talented soccer player and/or that she played soccer well.

- (6) a. Ana_{STIM} impressed Betty_{EXP} ⚽ b. Ana_{EXP} admired Betty_{STIM} ⚽

In sum, if face emoji are interpreted like expressives, they should show a preference for the experiencer argument (in addition to the first-person author). And if action emoji are interpreted like third-person pronouns, they should show a preference for the stimulus argument. In the rest of the paper, we report a psycholinguistic experiment that tests these complementary predictions.

4. Experiment.

4.1. INITIAL EXPERIMENT. We conducted an initial experiment where participants saw text messages with sentence-final emoji (Figure 1) and were asked to provide an answer to a statement with a blank: “The emoji provides information about _____” with three answer choices (the name of the subject character, the name of the object character, and the label ‘the sender of the message’). This same format was used with both action emoji and face emoji.

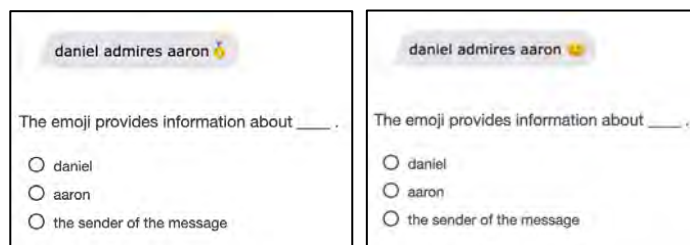


Figure 1. Sample items from the initial experiment

However, we subsequently realized that the wording ‘The emoji provides information about _____’ – in particular the noun *information* – is ambiguous in the face emoji conditions. The same question could be interpreted in different ways by different people, even in a situation where the actual interpretation that people construct for the text-emoji relation is the same. To see this, let's consider Figure 1. Here, under an interpretation where Daniel is feeling positive because he admires Aaron, someone might say that the smiling face *provides information about*

Daniel (the subject) because he is the one who is feeling pleased. But, under the same interpretation (Daniel is pleased because he admires Aaron), one could also say that the *emoji provides information about Aaron* because he is the one making Daniel pleased. While this second construal may be less likely, it is nevertheless possible. This is problematic because it means that participants' responses to this question in the face emoji conditions may not accurately reflect the information we are trying to tap into, namely whose emotions/feelings does the emoji represent.

Because of this wording ambiguity, we changed the question for the face emoji conditions to ensure that it clearly asks about whose emotions the face emoji expresses and conducted another experiment. Here, for reasons of brevity, we only report the results of this improved experiment – the ‘main experiment’ below. However, both experiments yielded similar results: The results of the initial experiment with the wording complication show the same basic patterns as the experiment reported below. The main difference is that certain patterns in the face emoji conditions are marginally significant ($0.05 < p < 0.1$) in the initial study (presumably due to the ambiguous question) and are significant ($p \leq 0.05$) in the main experiment (below).

4.2. MAIN EXPERIMENT. To test our hypothesis that face emoji are interpreted like expressives while action emoji are interpreted like third-person pronouns, we conducted a web-based experiment using psych verbs (e.g. *admire, annoy*) paired with face and action emoji. We also included transfer-of-possession verbs (e.g. *give, send*) as a control condition.

5. Method.

5.1. PARTICIPANTS. The experiment was conducted online using Qualtrics (Provo, Utah), and participants were recruited via Amazon MTurk. We report data from 56 adult native speakers of U.S. English. (None had done the initial experiment in Section 4.1). Participants received \$2.00.

5.2. MATERIALS AND DESIGN. We tested text messages with message-final emoji. We used this position because corpus data show emoji are frequently message-final (e.g. Novak et al. 2015, Na’aman et al. 2017). The messages were presented as images (made with iphonefaketext.com) and looked like messages sent with the iPhone message app. We chose the iPhone interface due to the widespread use of iPhones in the United States. The images were cropped so only the message was visible (Figure 2). All targets contained two same-gender names.

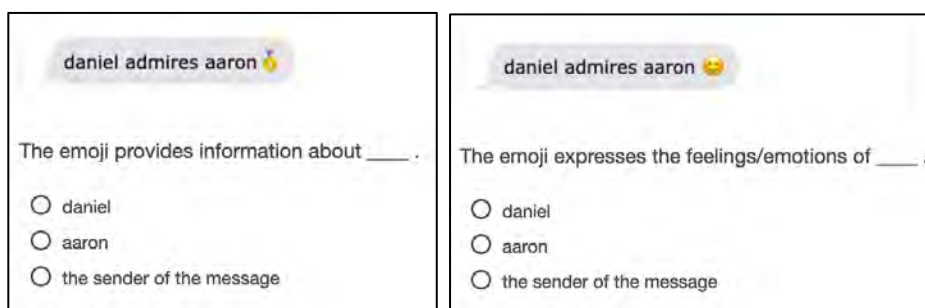


Figure 2. Sample items from the main experiment (psych verb condition)

Each message ended in an action or face emoji. Each emoji was used once, and we did not use action emoji depicting a full person (e.g. 🏃). In addition to emoji type (face vs. action), we manipulated verb type, and tested psych and transfer verbs (Sections 5.2.1-5.2.2). In addition to the targets (see below), the study had 18 fillers. Filler items did not contain psych or transfer verbs.

5.2.1 VERB TYPE: PSYCH VERBS. Psych verbs – which have one experiencer argument and one stimulus argument – are an ideal tool to test (i) whether face emoji tend to be linked to the expe-

riencer (like expressives), and (ii) action emoji to the stimulus (like third-person pronouns). We used both Stimulus-Experiencer (Stim-Exp) verbs (e.g. *impress*, *annoy*) and Experiencer-Stimulus (Exp-Stim) verbs (e.g. *admire*, *hate*). As illustrated in ex.(4-6), with Stim-Exp verbs the subject is the stimulus and the object is the experiencer, whereas the configuration is reversed with Exp-Stim verbs. This allowed us to assess whether the grammatical role of the experiencer or the stimulus argument plays a role (above and beyond the predicted effects of thematic role).

The Exp-Stim/Stim-Exp verbs were selected based on norms (Hartshorne & Snedeker 2013) to ensure that they have clear subject/object biases in implicit causality contexts (mean object bias of Exp-Stim verbs: 83%; mean subject bias of Stim-Exp verbs: 77%). Thus, the Exp-Stim verbs are NP2 verbs in implicit causality terms; the Stim-Exp verbs are NP1 verbs.⁴

On trials with psych verbs, the face emoji expressed positive or negative affect (e.g. 😊😞, see the right panel of Figure 2). The face emoji were chosen to be compatible – in principle – with all three potential candidates (the author of the message, the subject and the object). To ensure this, the valence of the face emoji matched the valence of the verb (e.g. *annoy* occurred with a negatively valenced face emoji and *admired* occurred with a positively valenced face emoji). This was intentional on our part: We wanted the face emoji to be potentially ambiguous, in order to be able to detect participants’ interpretational preferences in the absence of additional pragmatic cues.⁵ In other words: To test, in as fair a way as possible, whether face emoji tend to be linked to the first-person author or whether they can also be linked to a linguistically-expressed experiencer, both of these candidate antecedents should be equally available.

The action emoji paired with psych verbs (e.g. 🥁🏅🌸👟) represented objects related to actions or properties (e.g. playing drums, ice-skating), see Figure 2 for an example. The actions or properties expressed by the action emoji were chosen so that they could be interpreted as providing an explanation for the situation described in the sentence (e.g. *angela surprised tiffany* 🌸).⁶

In light of prior work on the phenomenon of interrogative flip with some subjective expressions, we also tested yes/no question versions of Exp-Stim and Stim-Exp verbs. However, as that phenomenon is not the focus of this paper, we do not discuss those results here.

5.2.2 VERB TYPE: TRANSFER VERBS. In addition to psych verbs, we tested transfer verbs (e.g. *bring*, *toss*, *give*; 16 verbs, mostly from Rohde 2008). Transfer verbs act as a control case relative to the psych verbs, since transfer verbs have source and goal arguments, instead of stimulus and experiencer arguments. They offer a means for testing the interpretational biases of face and action emoji in the absence of these thematic roles. Indeed, face emoji paired with transfer verbs

⁴ The experiment included eight NP1 verbs and eight NP2 verbs. However, due to experimenter oversight, three of the NP2 verbs had agentive subjects instead of experiencer subjects, e.g. *congratulate*, *praise*. These verbs were omitted from further analysis. However, inclusion of these additional items yields the same overall pattern of results.

⁵ If we had used a negative emoji with a positive verb, e.g. *Kate admired Mary* 😞, the valence mismatch would have provided an *additional pragmatic cue* that would have guided interpretations of who the emoji is linked to. E.g. *Kate admired Mary* 😞 would presumably have triggered a reading where the emoji reflects the feelings of the author, since it is unlikely that the admirer or the admiree would experience negative feelings.

⁶ It is worth emphasizing that with Stim-Exp/Exp-Stim verbs, there seems to be a strong preference to interpret virtually any action emoji as providing an explanation for the event described by the verb (and being linked to the stimulus). Consider the following: *daniel admires aaron* 🏅, *daniel admires aaron* 🌸, *daniel admires aaron* 🥁, *daniel annoyed aaron* 🏅, *daniel annoyed aaron* 🌸, *daniel annoyed aaron* 🥁. This preference to interpret the action emoji as providing an explanation linked to the stimulus (e.g. because aaron is a great tennis player) can be straightforwardly derived from claims about the argument structure of Stim-Exp/Exp-Stim verbs having an ‘empty slot’ for a propositional (explanation of the) stimulus that caused the experiencer’s mental state (e.g. Bott & Solstad 2014).

may exhibit a preference to be linked to the first-person author – in line with the default first-person orientation of expressives. Action emoji, on the other hand – if they pattern like third-person pronouns – may show a preference for salient/topical referents such as sentential subjects (e.g. Brennan, Friedman, & Pollard 1987, Crawley & Stevenson 1990, also Chafe 1976).

- (7) a. ava shipped a package to isabella 📦 / ava shipped a package to isabella 🤔
b. abigail brought dessert to emily 🍰 / abigail brought dessert to emily 😊

With transfer verbs, the face emoji expressed positive or negative affect and were selected to be sufficiently ambiguous such that they were compatible with all potential candidates (the author, the source and the recipient, ex.7 [right]). The action emoji depicted the transferred object (ex.7 [left]). In light of prior work on pronoun interpretation with transfer verbs (e.g. Rohde et al. 2006, Ferretti et al. 2009, Kehler & Rohde 2013), we also manipulated aspect of transfer verbs (e.g. *shipped/was shipping*). However, because our data revealed no clear effects of the aspect manipulation, the simple past and past progressive are collapsed in the subsequent discussion.⁷

5.3. PROCEDURE. Participation took place online. Participants saw each text message on a separate screen, and answered a multiple-choice question about it. Examples are in Figure 2. On action emoji trials, participants were asked “The emoji provides information about ____” and on face emoji trials, participants were asked “The emoji expresses the feelings/emotions of ____.” Each question was followed by three answer choices, as illustrated in Figure 2. Participants could only select one answer choice.⁸ Participants saw the text message and the question on the same screen, i.e., there was no memory load. Participants completed the experiment at their own pace.

5.4. PREDICTIONS. Before turning to the results, let us recap the predictions for both emoji types, for the psych verb and the transfer verb conditions.

If face emoji pattern like expressives (i.e., are sensitive to presence of experiencers), we may find that they have a baseline preference to be linked to the first-person author of the message. Indeed, this *first-person author preference* is what we expect to see in the transfer verb conditions. This prediction is built on prior linguistic work showing that the first-person speaker tends to be interpreted as the default attitude-holder with expressives (see Section 3.1). However, in the psych verb conditions (ES/Stim-Exp verbs), we also have a linguistically-realized experiencer argument. If face emoji are sensitive to this, we predict that they will tend to be linked to the *experiencer* argument (the object with SE, the subject with ES) if such an argument is pre-

⁷ Work on pronoun interpretation, e.g. Rohde et al. (2006), has found that with transfer-of-possession verbs, imperfective aspect (e.g. *was shipping*) increases the likelihood of participants interpreting a subject-position pronoun as referring to the source (subject) referent, as compared to perfective aspect (e.g. *shipped*) which boosted the proportion of goal (object) interpretations. This is attributed to (i) perfective aspect making the end-state more prominent, boosting goal interpretations, and (ii) imperfective aspect making the ongoing process of the event more prominent. In this second ongoing-event case, the source (subject) character can presumably be construed as relatively more prominent. In our stimuli the emoji depicted the transferred object itself, which could easily be construed as commenting on the entire event (and not only the goal or source character). In light of this, we think it is not surprising that we found no clear effects of the aspect manipulation: Focusing on the event itself (due to the emoji) could yield an effect in line with imperfective aspect – namely a preference to focus on the subject (source), which is indeed what we found (see Section 6.1 on transfer verbs). Thanks to Andy Kehler for discussion regarding this point.

⁸ A possible follow-up could allow people to choose more than one answer, thereby indicating that an emoji can reflect multiple opinion-holders or be ambiguous between two or more alternatives. In the current study, we can already get a sense of whether these issues are at play by aggregating responses. As we see in Section 6, our data suggest that each condition we tested has a ‘winner,’ one choice that was significantly preferred over others.

sent. This prediction is built on earlier psycholinguistic work by Kaiser & Herron Lee (2017, 2018) showing that other perspective-sensitive linguistic expressions (predicates of personal taste) tend to be interpreted as being anchored to linguistically-realized experiencer arguments.

If action emoji show discourse coherence effects and prominence effects similar to those seen with third-person pronoun resolution (see Section 3.2), we predict a preference for action emoji to be linked to the *stimulus* with psych verbs (the subject of Stim-Exp and the object of Exp-Stim verbs). In the case of transfer verbs, no stimulus argument is present. Here, we may find that action emoji tend to be interpreted as linked to the *subject* of the message, in light of the large body of prior work showing that agentive subjects are topical/prominent (see also fn 7).

	<i>Transfer verbs</i>	<i>Psych verbs</i>
Face emoji 😊	First-person author	Experiencer argument
Action emoji 🍷	Subject	Stimulus argument

Table 1. Predictions for face and action emoji as a function of verb type, based on our hypothesis

6. Results. In this section we first consider the results for face emoji and action emoji with transfer verbs (Section 6.1) and then turn to psych verbs (Section 6.2).

6.1. TRANSFER VERBS. The results for transfer verbs are shown in Figure 3. Here, the y axis shows the proportion of author, subject (source) and object (goal) character responses participants gave when asked about who the emoji provides information about (action emoji) and whose feelings/emotions the emoji expresses (face emoji).

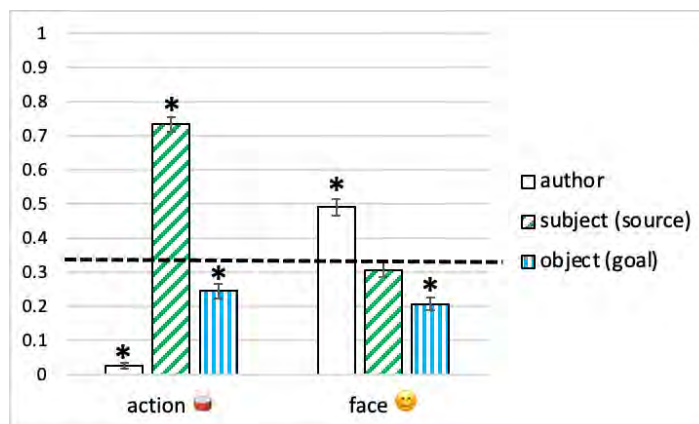


Figure 3. Transfer verbs: The proportion of author, subject and object responses. Here and in subsequent graphs, the dotted line denotes chance (0.333), the asterisk * indicates a significant difference ($p \leq 0.05$) from chance. Error bars show +/- 1 SE.

As can be seen in the figure, *face emoji* (the three bars on the right) are most likely to be interpreted as expressing the emotions of the author of the message: the proportion of author responses is significantly higher than chance (i.e., significantly higher than 0.333, one-sample t-test). Throughout the paper, we use * to indicate significant differences ($p \leq 0.05$) from chance. Chance (marked by the dotted horizontal line) is defined as 0.333, as there are three answer choices. This preference for the author is exactly in line with our predictions. The proportion of object responses is significantly below chance, while the proportion of subject responses does not differ from chance. The observed preference to interpret face emoji as expressing the feelings of the first-person author echoes the well-known bias exhibited by expressives and other per-

spective-sensitive elements in language for first-person attitude-holders, discussed in Section 3.1.

In contrast to face emoji, *action emoji* (the three bars on the left) tend to be interpreted as providing information about the subject (source) of the sentence. As shown in Figure 3, the proportion of subject (source) responses is significantly above chance, while the proportion of object (goal) and author choices are below chance. This subject preference is what we predict if the interpretation of action emoji resembles the interpretation of third-person pronouns, as we hypothesized in our earlier work.

6.2. PSYCH VERBS. Figure 4a shows the results for face emoji when they follow a psych verb, and Figure 4b shows the results for action emoji in the same context. Recall our prediction that face emoji will show a preference for the experiencer argument whereas action emoji will prefer the stimulus argument. Indeed, these predictions are borne out for both emoji types:

Face emoji tend to be interpreted as expressing the emotions of the linguistically-expressed experiencer: As shown in Figure 4a, with both Stim-Exp and Exp-Stim verbs, the proportion of experiencer choices is significantly above chance. Although the experiencer preference is numerically stronger when the experiencer is in subject position (Exp-Stim verbs; three right-side bars in Figure 4a), it is nevertheless present even with object-position experiencers (Stim-Exp verbs; three left-side bars). Thus, face emoji resemble perspective-sensitive adjectives in showing a preference to be interpreted as expressing the opinions of a linguistically-expressed experiencer argument (Kaiser & Herron Lee 2017, 2018). Indeed, in the presence of an experiencer argument, we no longer see a preference for the author – in other words, the first-person attitude holder preference that face emoji exhibited with transfer verbs is not seen with psych verbs.

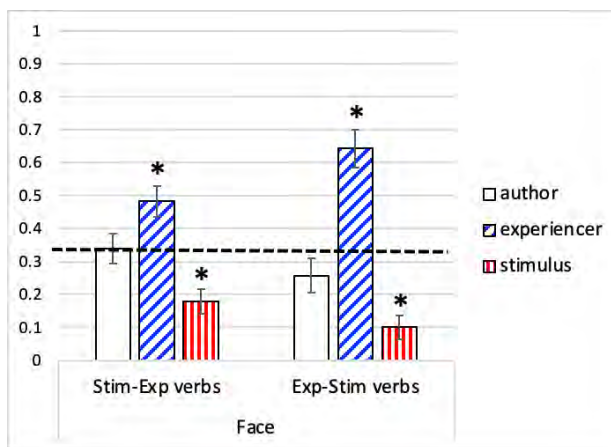


Figure 4a. Psych verbs and face emoji: Proportion of author, experiencer & stimulus responses

In contrast to face emoji, action emoji tend to be interpreted as providing information about the stimulus. As can be seen in Figure 4b, the proportion of stimulus choices is significantly above chance with both Exp-Stim and Stim-Exp verbs. Similar to what we saw with face emoji, grammatical role has a slight numerical effect: the stimulus preference is stronger when the stimulus is in subject position (Stim-Exp verbs), but, crucially, is still present even with object-position stimulus arguments (Exp-Stim verbs). Thus, action emoji resemble third-person pronouns in showing a preference to be interpreted as linked to the stimulus argument (see e.g. Garvey & Caramazza 1974, Hartshorne & Snedeker 2013, Bott & Solstad 2014 on pronoun resolution).

It is worth pointing out that Figures 4a and 4b present the results in terms of thematic, not grammatical role. Thus, the blue diagonal striped (experiencer) column corresponds to the grammatical *object* of Stim-Exp verbs and to the grammatical *subject* of Exp-Stim verbs, where-

as the red vertical striped (stimulus) columns corresponding to the grammatical *subject* of Stim-Exp verbs and to the grammatical *object* of Exp-Stim verbs. The graphs transparently show that grammatical role (which may be connected to effects from adjacency, recency, etc) is not decisive for the interpretation of face emoji. Although grammatical role has a modulating effect, the key factor is clearly thematic role.

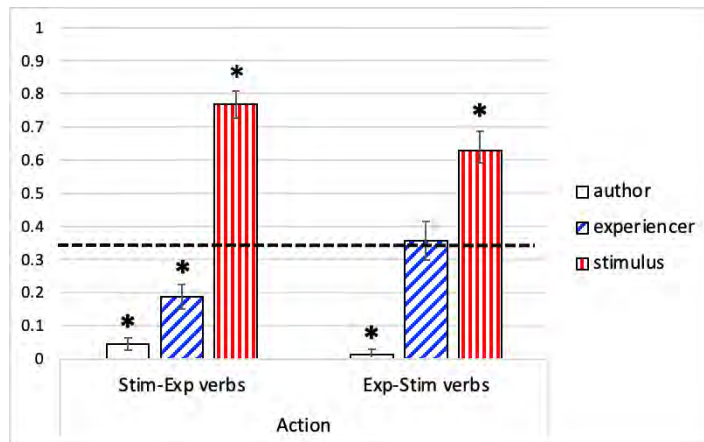


Figure 4b. Psych verbs and action emoji: Proportion of author, experiencer & stimulus responses.

7. Discussion and conclusions. Inspired by the prevalent use of emoji in digital communication, this paper investigates the nature of the relation between emoji and text. The psycholinguistic experiment reported in the present paper tests and extends a proposal that we developed in recent work with Francesco Pierini (Grosz, Kaiser & Pierini to appear); we point the interested reader to this paper for a more formal implementation of the ideas tested here, as well as discussion concerning different coherence relations and how they interact with scope-taking elements such as negation. That work also includes detailed discussion of corpus examples from Twitter. We regard the combined use of experimental and corpus-based work as very important in this domain.

The experiment reported here tests the idea that both face emoji (e.g. 😊😐😞) and action emoji (e.g. 🏀👉🍰) involve anaphoric dependencies (i.e., can be linked to preceding linguistic content) but that these dependencies are of different types: We hypothesize that face emoji resemble expressives (e.g. *wow*, *yay*, *damn*) which are typically interpreted as expressing the attitudes of the first-person speaker but can, in some contexts, shift to another salient experiencer (e.g. Amaral et al 2007, Harris & Potts 2009, Lasersohn 2005). We further hypothesize that action emoji are interpreted based on principles of discourse coherence (e.g. relations like Explanation, Elaboration), thus resembling third-person pronouns (e.g. Kehler 2002, Kehler & Rohde 2013, see also Asher & Lascarides 2003).

The experimental results confirm these predictions. We manipulated emoji type (face emoji vs. action emoji) as well as verb type (psych verbs with stimulus and experiencer arguments vs. transfer verbs without stimulus or experiencer arguments). We chose to focus on psych verbs because existing work shows that (i) subjective expressions in language are sensitive to the presence of experiencer arguments (e.g. Kaiser & Herron Lee 2017, 2018 on predicates of personal taste) and (ii) third person pronouns in implicit causality contexts are sensitive to the presence of stimulus arguments (e.g. Garvey & Caramazza 1974 and much subsequent work). Thus, psych verbs (with one stimulus argument and one experiencer argument) provide an ideal tool for testing *both* face emoji (to see if they tend to be interpreted as linked to experiencers) and action emoji (to see if they tend to be linked to stimulus).

The results show that face emoji indeed resemble expressives: If a linguistically-realized experiencer argument is present (psych verbs), it is preferred as the attitude-holder for the affective content expressed by the emoji, regardless of grammatical role. In contexts with no linguistically-realized experiencer (transfer verbs), the first-person author is the preferred attitude-holder.

Moreover, we find that action emoji show the same kind of sensitivity to discourse-coherence and topicality as third-person pronouns. When a linguistically-realized stimulus argument is present (psych verbs), action emoji tend to be interpreted as providing information about that referent. In contexts with no stimulus argument (transfer verbs), action emoji tend to be interpreted as providing information about the subject of the message – which fits with prior work showing that by default, subjects tend to be the most prominent/topical referents.

7.1. GENERAL CONSTRAINTS ON DISCOURSE COHERENCE. Although we find clear evidence for our hypothesis that face emoji pattern like expressives and action emoji pattern like third-person pronouns, other factors are also at play in guiding the use and interpretation of emoji. For example, not every emoji after a psych verb is easily interpretable as referring to the stimulus.

- (8) a. ?? richie annoyed adrian 🌲
b. ?? John took a train from Paris to Istanbul. He likes spinach. (Hobbs 1979)
c. Jane took a train from Paris to Istanbul. She had to attend a conference.
(Jurafsky & Martin 2020)

This is illustrated by incongruence of the text-emoji pairing in ex.(8a). One *could* come up with some interpretation (e.g. he chopped down Adrian’s favorite spruce), but this requires some ‘mental gymnastics.’ We suggest that these effects are nothing special about emoji, and simply reflect the fact that there are limits to how much inferencing we can easily accomplish. Similar effects occur in language (ex.8b,c). Both sentences in ex.(8b) are fully grammatical, but put together, they yield an incoherent discourse. One *could* always come up with an explanation (e.g. the spinach in Istanbul is better than in Paris), but the mental gymnastics are considerable. In contrast, ex.(8c) sounds fine; we can easily see how the two sentences are related. In sum, the fact that not all text-emoji pairings are judged congruent/coherent is, in our opinion, not a special property of emoji but best attributed to general principles of communicative coherence that also apply to purely linguistic communication (see also Weissman & Tanner 2018, Weissman 2019 for work on how text-emoji (in)congruence can guide perception of irony).

7.2. FUTURE DIRECTIONS. Building on our earlier work in Grosz, Kaiser and Pierini (to appear), this paper focused on face emoji and action emoji. An important aim for future work is to expand the scope of investigation to other emoji types. Face emoji can plausibly be classified as a subset of a more general class of affective emoji, which also includes certain body part emoji (👍👊👏) and heart emoji (💕💖💗💘); affective emoji can be further classified along criteria such as valence (e.g. positively 👍💕😊 vs. negatively valenced 👎💔😞). Perhaps the core distinction is thus not between face emoji and non-face emoji, but between affective and non-affective emoji. At the same time, faces are known to be privileged in perception and cognition, which should be taken into consideration and controlled for in relevant follow-up research. We are currently investigating if our findings for face emoji extend to a more varied set of emoji. Moreover, so far we have focused on message-final anaphoric emoji (e.g. daniel admires aaron 😊). A possible extension involves the probing of cataphoric relations – cases where emoji precede the sentence (e.g. 😊 daniel admires aaron). Initial intuitions suggest that pre-text emoji have more of a frame-setting role than post-text emoji. This might suggest that the pre-text emoji are more strongly oriented

towards the author of the message than a possible experiencer argument.

References

- Amaral, Patricia, Craige Roberts & Elizabeth A. Smith. 2007. Review of *The logic of conventional implicatures* by Chris Potts. *Linguistics and Philosophy* 30. 707–749.
- Ariel, Mira. 1990. *Accessing noun phrase antecedents*. London: Routledge.
- Asher, Nicholas & Alex Lascarides. 2003. *Logics of conversation*. Cambridge: Cambridge University Press.
- Bai, Qiyu, Qi Dan, Zhe Mu & Maokun Yang. 2019. A systematic review of emoji: Current research and future perspectives. *Frontiers in Psychology*. 10. 2221. <https://doi.org/10.3389/fpsyg.2019.02221>.
- Bott, Oliver & Torgrim Solstad. 2014. From verbs to discourse: A novel account of implicit causality. In Barbara Hemforth, Barbara Mertins & Cathrine Fabricius-Hansen (eds.), *Psycholinguistic approaches to meaning and understanding across languages*, 213–251. Dordrecht: Springer.
- Brennan, Susan E., Marilyn Friedman & Carl Pollard. 1987. A centering approach to pronouns. *Proceedings of the 25th Annual Meeting of the Association for Computational Linguistics*, 155–162. <https://doi.org/10.3115/981175.981197>.
- Bylinina, Lisa. 2014. *The grammar of standards: Judge-dependence, purpose-relativity, and comparison classes in degree constructions*. Utrecht: UiL OTS dissertation.
- Cohn, Neil, Jan Engelen & Joost Schilperoord. 2019. The grammar of emoji? Constraints on communicative pictorial sequencing. *Cognitive Research: Principles and Implications* 4. 33. <https://doi.org/10.1186/s41235-019-0177-0>.
- Crawley, Rosalind A. & Rosemary J. Stevenson. 1990. Reference in single sentences and in texts. *Journal of Psycholinguistic Research* 19. 191–210.
- Chafe, Wallace L. 1976. Givenness, contrastiveness, definiteness, subjects, topics and point of view. In Charles N. Li (ed.), *Subject and topic*, 27–55. New York: Academic Press.
- Davidson, Donald. 1967. The logical form of action sentences. In Nicholas Resher (ed.), *The logic of decision and action*, 81–95. Pittsburgh: University of Pittsburgh Press.
- Dresner, Eli & Susan C. Herring. 2010. Functions of the nonverbal in CMC: Emoticons and illocutionary force. *Communication Theory* 20(3). 249–268. <https://doi.org/10.1111/j.1468-2885.2010.01362.x>.
- Dunlap, Joanna C., Devshikha Bose, Patrick R. Lowenthal, Cindy S. York, Michael Atkinson, & Jim Murtagh. 2016. What sunshine is to flowers: A literature review on the use of emoticons to support online learning. In Sharon Y. Tettegah & Martin Gartmeier (eds.), *Emotions technology design & learning*, 163–182. London: Academic Press.
- Ekman, Paul, Wallace V. Friesen & Phoebe Ellsworth. 1972. *Emotion in the human face*. Elmsworth, NY: Pergamon Press.
- Evans, Vyvyan. 2017. *The emoji code: How smiley faces, love hearts and thumbs up are changing the way we communicate*. London: Michael O'Mara Books Limited.
- Ferretti Todd, R, Hannah Rohde, Andrew Kehler & Melanie Crutchley. 2009. Verb aspect, event structure, and co-referential processing. *Journal of Memory and Language*. 61. 191–205.
- Foolen, Ad. 2015. Expressives. In Nick Riemer (ed.), *The Routledge handbook of semantics*, 473–490. London: Routledge.
- Garrison, Anthony, Dirk Remley, Patrick Thomas & Emily Wierszewski. 2011. Conventional faces: Emoticons in instant messaging discourse. *Computer and Composition* 28. 112–125.

- Garvey, Catherine and Alfonso Caramazza. 1974. Implicit causality in verbs. *Linguistic Inquiry* 5. 459–464. <https://www.jstor.org/stable/4177835>.
- Gawne, Lauren & Gretchen McCulloch. 2019. Emoji as digital gestures. *Language@Internet*. 17. 2. <https://www.languageatinternet.org/articles/2019/gawne>.
- Gerke, Katherine & Dennis Ryan Storoshenko. 2018. Picturing syntax: Cross-linguistic variation in the interpretation of emoji strings. Presented at the 2018 Conference of the Canadian Linguistic Association.
- Godin, Seth. 1993. *The smiley dictionary*. Berkeley, CA: Peachpit.
- Grosz, Patrick, Elsi Kaiser & Francesco Pierini. To appear. Discourse anaphoricity and first-person indexicality in emoji resolution. *Proceedings of Sinn und Bedeutung* 25.
- Grosz, Patrick, Gabriel Greenberg, Christine De Leon & Elsi Kaiser. 2021. A semantics of face emoji in discourse. Manuscript, University of Oslo / UCLA / USC.
- Gundel, Jeanette. K., Nancy Hedberg & Ron Zacharski. 1993. Cognitive status and the form of referring expressions in discourse. *Language* 69. 274–307.
- Harris, Jesse & Chris Potts. 2009. Perspective-shifting with appositives and expressives. *Linguistics and Philosophy* 32. 523–552. <https://doi.org/10.1007/s10988-010-9070-5>.
- Hartshorne, Joshua & Jesse Snedeker. 2013. Verb argument structure predicts implicit causality: The advantages of finer-grained semantics. *Language and Cognitive Processes* 28. 1474–1508. <https://doi.org/10.1080/01690965.2012.689305>.
- Hobbs, Jerry R. 1979. Coherence and coreference. *Cognitive Science* 3. 67–90.
- Jaeger, Sara, Christina Roigard, David Jin, Leticia Vidal & Gastón Ares. 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. <https://doi.org/10.1016/j.foodres.2018.10.074>.
- Jurafsky, Daniel & James H. Martin. 2020. Speech and language processing. Ms., Stanford.
- Kaiser, Elsi & Jamie Herron Lee. 2017. Experience matters: A psycholinguistic investigation of predicates of personal taste. *Semantics and Linguistic Theory (SALT)* 27. 323–339. <https://doi.org/10.3765/salt.v27i0.4151>.
- Kaiser, Elsi & Jamie Herron Lee. 2018. Predicates of personal taste and multidimensional adjectives: An experimental investigation. *Proceedings of WCCFL* 35. 224–231. <http://www.lingref.com/cpp/wccfl/35/paper3393.pdf>.
- Kaiser, Elsi. 2015. Perspective-shifting and free indirect discourse: Experimental investigations. *Semantics and Linguistic Theory (SALT)* 25. 346–372. <https://doi.org/10.3765/salt.v25i0.3436>.
- Kehler, Andrew & Hannah Rohde. 2013. A probabilistic reconciliation of coherence-driven and centering-driven theories of pronoun interpretation. *Theoretical Linguistics* 39. 1–37. <https://doi.org/10.1515/tl-2013-0001>.
- Kehler, Andrew. 2002. *Coherence, reference, and the theory of grammar*. Stanford: CSLI.
- Keltner, Dacher & Daniel T. Cordaro. 2017. Understanding multimodal emotional expressions: Recent advances in basic emotion theory. In James A. Russel & José M. Fernández-Dols (eds.), *The science of facial expression*, 57–75. Oxford: Oxford University Press.
- Lasnik, Peter. 2005. Context dependence, disagreement, and predicates of personal taste. *Linguistics and Philosophy* 28. 643–686. <https://doi.org/10.1007/s10988-005-0596-x>.
- Lasnik, Peter. 2007. Expressives, perspective and presupposition. *Theoretical Linguistics*. 33. 223–230. <https://doi.org/10.1515/TL.2007.015>.
- LeCompte, Travis & Jianhua Chen. 2017. Sentiment analysis of tweets including emoji data. *International Conference on Computational Science and Computational Intelligence*. <https://doi.org/10.1109/CSCI.2017.137>.

- Li, Weijian, Yuxiao Chen, Tianran Hu & Jiebo Luo 2018. Mining the relationship between emoji usage patterns and personality. Paper presented at the Twelfth International AAAI Conference on Web and Social Media (ICWSM).
- Luangrath, Andrea W., Joann Peck & Victor A. Barger. 2017. Textual paralanguage and its implications for marketing communications. *Journal of Consumer Psychology* 27. 98–107.
- Maier, Emar. 2020. The semantics of smiles and smileys. Presentation, SuB 25 Special Session on Gestures and Natural Language Semantics.
- McNally, Louise & Isidora Stojanovic. 2017. Aesthetic adjectives. In James O. Young (ed.), *Semantics of aesthetic judgments*. 17–37. Oxford: Oxford University Press.
- Na'aman, Noa, Hannah Provenza & Orion Montoya. 2017. Varying linguistic purposes of emoji in (Twitter) context. *Proceedings of ACL 2017, Student Research Workshop*. 136–141.
- Ninan, Dilip. (2014). Taste predicates and the acquaintance inference. *Semantics and Linguistic Theory (SALT)* 24. 290–309. <https://doi.org/10.3765/salt.v24i0.2413>.
- Novak, Petra K., Jasmina Smailović, Borut Sluban & Igor Mozetič. 2015. Sentiment of emoji. *PLoS ONE* 10(12). e0144296. <https://doi.org/10.1371/journal.pone.0144296>.
- Pasternak, Robert & Lyn Tieu. 2020. Co-linguistic content projection: From gestures to sound effects and emoji. Ms. <https://ling.auf.net/lingbuzz/005082>.
- Pierini, Francesco. To appear. Emojis and gestures: a new typology. *Proceedings of Sinn und Bedeutung* 25.
- Potts, Chris. 2007. The expressive dimension. *Theoretical Linguistics*. 33 165–198. <https://doi.org/10.1515/TL.2007.011>.
- Rett, Jessica. To appear. The semantics of emotive markers and other illocutionary content. *J of Semantics* https://linguistics.ucla.edu/people/Rett/Rett_2021_emotive_markers.pdf.
- Rohde, Hannah. 2008. *Coherence-driven effects in sentence and discourse processing*. San Diego, CA: University of California dissertation.
- Rohde, Hannah, Andrew Kehler & Jeffrey L. Elman. 2006. Event structure and discourse coherence biases in pronoun interpretation. *Proceedings of the Annual Conference of the Cognitive Science Society* 28. 697–702.
- Russell, James A. & José M. Fernández-Dols (eds.). 1997. *The psychology of facial expression*. Cambridge: Cambridge University Press.
- Tomkins, Silvan S. & Robert McCarter. 1964. What and where are the primary affects? Some evidence for a theory. *Perceptual and Motor Skills* 18. 119–158.
- Troiano, Gianmarco & Nicola Nante. 2018. Emoji: What does the scientific literature say about them? – A new way to communicate in the 21th century. *Journal of Human Behavior in the Social Environment* 28. 528–533. <https://doi.org/10.1080/10911359.2018.1437103>.
- Vendler, Zeno. 1957. Verbs and times. *Philosophical Review* 66. 143–160.
- Weiß, Martin, Patrick Mussel & Johannes Hewig. 2020. The value of a real face: Differences between affective faces and emojis in neural processing and their social influence on decision-making. *Social Neuroscience* 15(3). 255–268.
- Weissman, Benjamin & Darren Tanner. 2018. A strong wink between verbal and emoji-based irony: How the brain processes ironic emojis during language comprehension. *PLoS ONE* 13(8). e0201727. <https://doi.org/10.1371/journal.pone.0201727>.
- Weissman, Benjamin. 2019. Emojis in sentence processing: An electrophysiological approach. *Companion Proceedings of the 2019 World Wide Web Conference (WWW'19 Companion)* 478–479. <https://doi.org/10.1145/3308560.3316544>.